




The Indonesian EQ-5D-5L Value Set

Fredrick Dermawan Purba^{1,2}  · Joke A. M. Hunfeld¹ · Aulia Iskandarsyah³ ·
Titi Sahidah Fitriana⁴ · Sawitri Supardi Sadarjoen³ · Juan Manuel Ramos-Goñi⁵ ·
Jan Passchier⁶ · Jan J. V. Busschbach¹

Published online: 10 July 2017

© The Author(s) 2017. This article is an open access publication

Abstract

Background The EQ-5D is one of the most used generic health-related quality-of-life (HRQOL) instruments worldwide. To make the EQ-5D suitable for use in economic evaluations, a societal-based value set is needed. Indonesia does not have such a value set.

Objective The aim of this study was to derive an EQ-5D-5L value set from the Indonesian general population.

Methods A representative sample aged 17 years and over was recruited from the Indonesian general population. A multi-stage stratified quota method with respect to residence, gender, age, level of education, religion and ethnicity was utilized. Two elicitation techniques, the composite time trade-off (C-TTO) and discrete choice experiments (DCE)

were applied. Interviews were undertaken by trained interviewers using computer-assisted face-to-face interviews with the EuroQol Valuation Technology (EQ-VT) platform. To estimate the value set, a hybrid regression model combining C-TTO and DCE data was used.

Results A total of 1054 respondents who completed the interview formed the sample for the analysis. Their characteristics were similar to those of the Indonesian population. Most self-reported health problems were observed in the pain/discomfort dimension (39.66%) and least in the self-care dimension (1.89%). In the value set, the maximum value was 1.000 for full health (health state ‘11111’) followed by the health state ‘11112’ with value 0.921. The minimum value was –0.865 for the worst state (‘55555’). Preference values were most affected by mobility and least by pain/discomfort.

Conclusions We now have a representative EQ-5D-5L value set for Indonesia. We expect our results will promote and facilitate health economic evaluations and HRQOL research in Indonesia.

Electronic supplementary material The online version of this article (doi:[10.1007/s40273-017-0538-9](https://doi.org/10.1007/s40273-017-0538-9)) contains supplementary material, which is available to authorized users.

✉ Fredrick Dermawan Purba
f.purba@erasmusmc.nl; fredrick.purba@unpad.ac.id

¹ Section Medical Psychology and Psychotherapy, Department of Psychiatry, Erasmus MC University Medical Center, Wytemaweg 80, Room Na-2019, 3015 CN Rotterdam, The Netherlands

² Department of Developmental Psychology, Faculty of Psychology, Padjadjaran University, Jatinangor, Indonesia

³ Department of Clinical Psychology, Faculty of Psychology, Padjadjaran University, Jatinangor, Indonesia

⁴ Center of Applied Psychometrics, Faculty of Psychology, YARSI University, Jakarta, Indonesia

⁵ Executive Office, EuroQol Research Foundation, Rotterdam, The Netherlands

⁶ Department of Clinical, Neuro and Developmental Psychology, VU University, Amsterdam, The Netherlands

Key Points for Decision Makers

Indonesia does not have an EQ-5D value set.

An EQ-5D-5L value set was derived from a highly representative sample of the Indonesian general population.

Data were collected using a rigorous quality control procedure which led to logical and significant models.

This Indonesian EQ-5D-5L value set is now becoming available and will be used by all health economic evaluations and health-related quality-of-life studies in Indonesia that use EQ-5D.

1 Introduction

The Indonesian government wishes to improve equal access to healthcare by introducing universal health insurance. To ensure health technology assessment (HTA) can be undertaken for such an insurance scheme, Indonesia intends to employ cost-effectiveness analysis for new and existing medical interventions. To value the outcomes of a medical intervention in quality-adjusted life-years (QALYs) requires a quality-of-life instrument that can value the health states of patients using societal preferences, such as the EQ-5D instrument. At present, no Indonesian EQ-5D value set is available for the calculation of QALYs. There exists a standardized valuation protocol for the 5-level version of EQ-5D. We employed this protocol with over 1000 respondents representative of the Indonesian population. Below we describe in more detail (1) the social, economic and organizational HTA setting that determines the demand and specifications for an Indonesian valuation study; (2) a brief introduction to the EQ-5D-5L, its valuation protocol and the place of the EQ-5D in HTA; and (3) why we cannot rely on values set from European countries and/or neighbouring countries.

Indonesia is located in South East Asia, with 255.5 million inhabitants in 2015 [1]. Commencing in January 2014, Indonesia has implemented universal healthcare coverage organized by the ‘Badan Penyelenggara Jaminan Sosial Kesehatan’ or BPJS Kesehatan: the Healthcare and Social Security Agency. The aim of the BPJS Kesehatan is to include all Indonesian citizens in the National Health Insurance system to enable them to obtain access to healthcare benefits and to provide protection with respect to basic health needs [2]. The decision-making process related to the implementation of this national health coverage and the adoption of new technologies can benefit from an evidence-based strategy and the application of HTA, a decision-making process involving economic evaluation and other considerations such as those of an ethical and organizational nature, to ensure the optimal use of health technologies for the population. In 2015, the Ministry of Health of Indonesia formed a national HTA committee (Komite Penilaian Teknologi Kesehatan). The committee’s expected output is a policy recommendation to the Minister on the feasibility of the health service(s) to be included in the National Health Insurance benefit package [3, 4].

Economic evaluation uses clinical evidence to provide systematic consideration of the effects of all available alternatives regarding health, healthcare costs, and other effects regarded as valuable [5]. Cost-utility analysis (CUA) is used to evaluate health-related quality-of-life (HRQOL) outcomes and to compare costs and outcomes between different

healthcare programmes in terms of cost per QALY [5, 6]. A QALY is obtained by integrating a health state utilities function, measured by multi-attribute utility instruments (MAUIs), differentiated over a lifetime. The three most widely used MAUIs are the EQ-5D, the Health Utility Index (HUI), and the Short Form 6D (SF-6D) [5–8]. Several national HTA organizations, for example in the UK and Thailand, have recommended EQ-5D as the preferred method for deriving utilities [9, 10]. Developed by the EuroQol Group, EQ-5D is a standardized generic instrument that collects descriptive HRQOL data on five dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression); followed by a self-rating of overall health status on a visual analogue scale (EQ VAS) ranging from 0 (‘worst imaginable health state’) to 100 (‘best imaginable health state’) [11, 12]. In 2011, the EuroQol Group expanded the levels of severity of the classic version of EQ-5D, renamed EQ-5D-3L, from three to five levels. This new instrument is designated ‘EQ-5D-5L’ [12]. Recent studies have shown that EQ-5D-5L produces a richer description of health states, a higher discriminatory power, and a lower ceiling effect compared with EQ-5D-3L [13–18]. The EuroQol Group has also developed a valuation protocol for EQ-5D-5L [19], and the EuroQol Group Valuation Technology (EQ-VT) template computerized the interview method to standardize EQ-5D-5L valuation studies in different countries. This protocol provides a value set for the calculation of QALYs using a societal perspective, the preferred perspective in health economics [5].

Indonesia does not have an EQ-5D value set, either for the 3-level or for the new 5-level version. Previous EQ-5D studies conducted in Indonesia measured health preferences using the Malaysian value set or values derived from citizens of the UK [20, 21]. However, for a value set to be valid for Indonesia it should represent the culture and living standards of Indonesia [22]. Moreover, the values should match the particular wording of the Indonesian instrument: for instance, if ‘cukup’ (i.e. ‘moderate’) is less worse in Bahasa Indonesia than in the Malaysian language (‘sederhana’) or in English, then the values should match that difference. For these reasons the aim of our study was to obtain preferences from the general population in order to derive a national EQ-5D-5L value set for the calculation of QALYs from a societal, Indonesian perspective.

2 Methods

2.1 Respondents

A representative sample was recruited from the Indonesian general population, with a minimum of 1000 respondents aged 17 and over, based on the work of Ramos-Gofñi et al., to obtain a 0.01 standard error (SE) of the observed mean

composite time trade-off (C-TTO), 9735 C-TTO responses were needed. Therefore, the 1000 respondents interviewed provided 10,000 C-TTO and 7000 discrete choice responses to estimate the models [23]. The adult population was defined as aged 17 and over, because in Indonesia, the legal age to obtain an ID card, a driving license, and access to voting is 17. To ensure the representativeness of the final sample for the Indonesian general population, we used a multi-stage stratified quota method with respect to residence (urban/rural, as registered by the official national register); gender (male/female); age (17–30/31–50/>50 years); and level of education: basic (primary school and below), middle (primary school plus at least 1 year of high school) and high (all others). This resulted in the first stage of 36 quota groups. Two other categories, religion (Islam/Christian/Others) and ethnicity (own-declared ethnicity: Jawa/Sunda/Sumatera/Sulawesi/Madura-Bali/Others), were considered important as well. However, including them in the same way as residence, gender, age, and education would result in $36 \times 3 \times 6 = 3888$ quota groups. We therefore used religion and ethnicity quotas independently from the other factors. So religion and ethnicity are representative over the whole sample, but within the individual 36 quota groups this might not be the case. To take account of this second layer of sampling, we called this a ‘multi-stage stratified quota’. The predefined quotas were based on updated data from the Indonesian Bureau of Statistics [1].

We designed and used an online tool to ensure that the recruitment of respondents was in accordance with predefined quotas while the sampling was employed in different parts of the country. Interviews were conducted in the following six cities and their surroundings, located in different parts of Indonesia: Jakarta, Bandung, Jogjakarta, Surabaya, Medan, and Makassar. Respondents were recruited through a mixed strategy, i.e. through personal contact, local leader assistance, and from public places such as mosques and shopping streets. We also asked respondents to introduce us to other potential respondents. Interviews were conducted at the respondents’ or interviewers’ homes. For their participation, all respondents received a mug or a t-shirt specifically designed for the valuation study. Informed consent was obtained from all respondents included in the study. The study was approved by the Health Research Ethics Committee, Faculty of Medicine, Padjadjaran University, Indonesia.

2.2 Instruments

2.2.1 EQ-5D-5L

We used the official EQ-5D-5L Bahasa Indonesia version provided by the EuroQol Group. This translation of EQ-

5D-5L was produced using a standardized translation protocol that followed international recommendations [24]. As briefly mentioned in the introduction, EQ-5D-5L consists of five dimensions: mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD). Each dimension has five levels: no problems, slight problems, moderate problems, severe problems, and unable/ extreme problems [12]. The EQ-5D-5L instrument describes 3125 (5^5) unique health states. A 1-digit number expresses the level selected for that specific dimension. Hence, combining a 5-digit number for five dimensions will describe a specific health state. For example, state ‘11111’ indicates ‘no problems on any of the five dimensions’, while state ‘54321’ indicates ‘unable to walk about, severe problems washing or dressing, moderate problems doing usual activities, slight pain or discomfort, and no anxiety or depression’ [12]. Each health state has a so-called ‘sum score of the level digits’, which means the sum of the levels across domains; for example, ‘11111’ sum score of the level digits is 5 and ‘54321’ is 15. This EQ-5D descriptive system is followed by self-rating of overall health status on a visual analogue scale (EQ VAS) ranging from 0 (‘worst health you can imagine’) to 100 (‘best health you can imagine’).

2.2.2 Valuation Protocol

The EQ-5D-5L valuation protocol consists of five sections [19]:

1. A general welcome, where the interviewer explains the objectives of the research, followed by filling in the informed consent when the individuals agree to participate.
2. Introduction to and completion of the descriptive system, VAS and background questions (age, sex, experience of illness, religion, ethnicity and education).
3. C-TTO (see Sect. 2.2.3 below) tasks followed by a ‘Feedback Module’ task. Each respondent has to complete one example (health state: being in a wheelchair), three practice health states (mild: ‘21121’; severe: ‘35554’; and moderate but difficult to imagine: ‘15411’) and ten ‘real’ C-TTO tasks valuing hypothetical EQ-5D-5L health states. In the Feedback Module task, the respondents check whether they agree with the order of the health states they valued before. The EQ-VT screen shows health states for 10 C-TTO tasks arranged based on their value given by the respondents: from the lowest value at the bottom to the highest value at the top. Respondents are allowed to ‘flag’ the health state(s) for which they do not agree with the previously given relative position to

other health states, but they are not allowed to alter their initial values. Three debriefing questions regarding the difficulties of the C-TTO tasks are added at the end of this section.

4. A discrete choice experiment (DCE, see Sect. 2.2.3 below) followed by three debriefing questions regarding the DCE. Each respondent has to complete seven forced-pair comparisons.
5. A round-up, where respondents can comment on the valuation tasks.
6. Country-specific questionnaire(s) (if any).

All sections were administered utilizing computer-assisted face-to-face interviews employing the EQ-VT platform version 2.0.

2.2.3 Preference Elicitation Methods

Time trade-off (TTO) has been widely used as a standard method to elicit preferences [25, 26]. C-TTO uses conventional TTO to elicit better-than-dead (BTD) values, and lead-time TTO to elicit worse-than-dead (WTD) values. Details regarding C-TTO can be found in the study by Oppe et al. [27]. In summary, respondents were first faced with ‘conventional’ TTO where they had to choose between 10 years in an impaired health state (Life B) and 10 years of full health (Life A). After a series of choice-based iterations, respondents achieved a point of equivalence between the length of time in full health (Life A): ‘ x ’ and a period of time (10 years) in the impaired health state (Life B). The impaired health state value is defined as $x/10$. For example, if a respondent could not differentiate between 3 years of full health in Life A and 10 years living in Life B, then that health state value would be 0.3 (3/10). For a really poor health state, respondents might prefer to die immediately; that is, the value for that specific health state is <0 (death value = 0). In this case, the lead-time TTO approach was introduced to allow respondents to express a value below the value of death; that is, below 0. The two lives in the lead-time TTO are 10 years of full health (Life A) and 10 years of full health followed by 10 years in the impaired health state (Life B). When respondents reach an indifference point between the amount of time ‘ x ’ in Life A and Life B, the health state value is defined as $(x - 10)/10$. Hence, -1 is the lowest possible value of a given health state, generated from trading the full 10 years of Life A in a lead-time TTO.

The EQ-5D-5L valuation protocol included 86 EQ-5D-5L health states to be valued using C-TTO. The 86 health states were distributed into ten blocks with a similar level of severity. Eighty unique health states were selected using Monte Carlo simulation (eight unique health states included

in each block), five very mild states (only one dimension at level 2 and all others at level 1, e.g. ‘11112’) (each included in two blocks) and the most severe/‘pits’ state (‘55555’) (included in all blocks) [19]. Respondents were randomly assigned to one of the ten C-TTO blocks. Each state of the block was presented in random order to respondents using the EQ-VT platform.

However, it was realized that TTO has its limitations. EuroQol Group considered different valuation techniques to be used in conjunction with TTO to make the valuation studies more robust and valid. Previous experiments with DCE, like the study by Stolk et al. using EQ-5D-3L [28] or Ramos-Goñi et al. using EQ-5D-5L [29], showed that the DCE is a valid valuation technique to get health preferences from respondents. Since both TTO and DCE try to measure the same concept, it was anticipated that DCE could be used in combination with TTO [30]. In the light of this reasoning, DCE was included in the EuroQol VT protocol.

Each DCE task was conducted by presenting two health states and asking the respondent to select the preferred state for him/her. The DCE design consisted of 196 pairs of EQ-5D-5L health states distributed over 28 blocks, each consisting of seven pairs with a similar severity [19]. The seven paired comparisons were presented in random order by the EQ-VT; in addition, the right–left order of the two health states presented was also randomized.

2.3 Data Collection

At the outset, 13 interviewers were recruited and trained intensively in a 1-day workshop at two locations: (1) Jakarta for interviewers who worked in Jakarta, Bandung and Makassar; and (2) Jogjakarta for interviewers who worked in Jogjakarta, Surabaya and Medan. Each interviewer performed at least five pilot interviews in the week after training. Their experiences were discussed and feedback was given by the daily supervisor. Only after this were they permitted to conduct real data interviews. Three additional interviewers were hired during the data collection and they received similar training and met similar requirements to the first 13. Interviews were performed between March 9, 2015 and January 24, 2016. After 102 interviews we evaluated the quality of the interviews (see Sect. 2.5 below) and we concluded that their quality was not yet sufficient. Hence we retrained the interviewers and treated the 102 interviews collected thus far as pilot interviews, excluding the 102 interviews in the data analysis. A detailed description of this decision-making process and the retraining of the interviewers is provided elsewhere [30].

2.4 Exclusion Criteria

There were two main criteria for data exclusion: lack of completion of an interview and characteristics of respondents' answers that related to poor understanding of the task or to errors [31]. Note that the first criterion concerns excluding respondents and the second excludes respondent answers/responses.

With respect to the first criterion, interviews were excluded when respondents did not finish the interview for the following reasons: (1) the respondent indicated that he/she did not want to continue the interview process, (2) interviewers concluded that the respondent was unable to differentiate between the different dimensions and levels of EQ-5D-5L, (3) interviewers concluded that the respondent was not able to comprehend the C-TTO task during the practice session. When an interview had to be stopped during the C-TTO task it was excluded from the study.

With respect to the second criterion, completed interview responses were excluded on account of any of the following characteristics: (1) a respondent had a positive slope on the regression between his/her values on C-TTO and the 'sum score of the level digits', as this would indicate that the respondent provided higher utility values for poorer health states on average—the slope of the regression between C-TTO and the 'sum score of the level digits' was generated as part of the standard quality control report; (2) when a response in the C-TTO tasks was judged to be irrational: for instance, preferring life B (10 years in the corresponding health state) to life A (10 years in full health) and not shifting after his/her initial response was reconfirmed by the interviewer; (3) responses that were marked by the respondents in the Feedback Module task, which was a sign that the respondents disagreed with the valuation of those responses.

2.5 Quality Control

To ensure data quality, we followed the quality control (QC) process described by Ramos-Goñi et al. [32], which consisted of minimum quality criteria and cyclical feedback to improve interviewers' skills. The EuroQol Group facilitates use of the EQ-VT QC tool, which is a software programme that automates the production of QC reports based on data from EQ-VT studies. Bi-weekly meetings (teleconference-based) were organized to discuss the QC reports with the EQ-VT support team. The aim of these meetings was to evaluate and improve the interviewers' performance and to check for possible non-compliance to the interview protocol.

2.5.1 Minimum Quality Criteria

The QC reports provided a number of statistics related to the quality of the data collected thus far, differentiated by interviewer.

1. Wheelchair time: when the duration of time an interviewer used to explain the 'wheelchair example' preceding the actual C-TTO tasks was <3 min.
2. Wheelchair lead-time: when the interviewer did not explain the WTD element of the wheelchair example.
3. C-TTO duration: if completing the ten C-TTO tasks took <5 min.
4. 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 signs are observed, the interview is 'flagged' as being of suspicious quality. If four or more of the interviews are flagged as being of poor quality, all ten interviews thus far conducted by that specific interviewer are removed and retraining of that interviewer is conducted. After a further ten interviews, the performance and compliance are re-evaluated. If again four or more interviews are flagged, the next set of ten interviews will also be removed and the interviewer is removed from the data collection process. Quality control focused on the interviewer; responses in flagged interviews were not removed from the data that was analysed.

The DCE part of the valuation study was also monitored to detect suspicious response patterns. Assuming that A is the health state at the left of the screen and B is the health state at the right of the screen, then a consistent preference for the left (A) would be suspicious (AAAAAAA). The same would apply for the response pattern BBBBBBB, ABABABA, BABABAB. This was also reported in the QC report.

2.5.2 Cyclical Feedback

The retraining programme conducted by the daily supervisor was held in 2 locations: (1) Jakarta for interviewers who worked in Jakarta, Bandung and Makassar; and (2) Jogjakarta for interviewers who worked in Jogjakarta, Surabaya and Medan. The QC reports for their interviews were presented, discussions were held to address non-compliance problems, and suitable solutions were agreed upon among the interviewers. After the retraining programme, the daily supervisor continuously created QC reports, made notes at the group and individual levels, and sent feedback to the interviewers, so that they were able to learn from their own and other interviewers' performance.

2.6 Data Analysis

We describe the sample characteristics including self-reported health on the EQ-5D-5L descriptive system and the EQ-VAS using percentages for discrete variables and means and standard deviations for continuous variables in comparison with the Indonesian population. A general z test was used to investigate whether the proportions in the sample were similar to, or different from, the general population.

In this investigation we used TTO (specifically C-TTO) and DCE. TTO has limitations such as loss of aversion [33], but also has advantages as the TTO-based value sets are anchored on a scale of (0) death to (1) full health. DCE is not exempt from limitations, as lexicographic behaviour from respondents has been widely reported in the literature [34]. It is also noticeable that DCE, in its present form, where time is not incorporated in health state presentations, does not anchor value sets on a (0) death to (1) full health scale. Therefore, DCE produces value sets on an arbitrary scale based on the relative distances between health states.

However, both techniques attempt to measure health states preference, but using different underlying assumptions, and seem to not share the same limitations. Therefore, the data obtained from these two elicitation methods could be seen as complementary, not necessarily competing with each other. Hence, we chose the solution presented by Oppe and van Hout [35], who combined DCE with C-TTO in a ‘hybrid model’, imposing the (0) death to (1) full health scale as determined by C-TTO.

To illustrate how the hybrid model combined C-TTO and DCE responses in this study, we also present the results from the models estimated from each C-TTO and DCE separately, with the same assumptions as those used for the hybrid model. We used the 20-parameter main effects model, which estimates four parameters for the five levels of each of the five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each coefficient represents the additional utility decrement of moving from one level to another. Hence, the overall decrement of moving from ‘no problems’ to ‘unable/extreme problems’ is calculated as the sum of the coefficients of ‘no problems to slight problems’, ‘slight problems to moderate problems’, ‘moderate problems to severe problems’, and ‘severe problems to unable/extreme problems’.

Presenting the TTO, the DCE and the hybrid model also allows us to compare the value distribution in the form of the correlations between the predicted values of the models, and we can compare the weights of the individual dimensions. This gives information about construct validity in the form of ‘convergent validity’, or ‘concordance’.

Modelling was undertaken using the STATA statistical package. C-TTO data were modelled using the response

values as dependent variables and the health states as explanatory variables. This was achieved by the implementation of a Tobit model (hyreg with ll() option), which assumes a latent variable Y^* underlying the observed Y of C-TTO values when there is either left- or right-censoring in the dependent variable. The C-TTO data, in particular the lead-time C-TTO for WTD health states, is by nature censored at -1 [ll(-1) option on hyreg command]. This means that observed preference values were valued by the C-TTO method at -1 , despite the latent preferences of respondents possibly including values lower than -1 [36]. The Tobit model accounts for this censoring by estimating the latent variable Y^* , which can take on predicted preference values extrapolated beyond the range of the observed values. Variance of C-TTO data is not homogeneous among health states; this led us to model C-TTO data as heteroskedastic data. We used the hetcont() option of the hyreg command as suggested by Ramos-Goñi et al. [37]. The dummy variables included in the hetcont() option were the same as those included in the main model, that is, the 20 dummies that specified the main effects model.

DCE (forced pair comparisons in our case) responses were modelled as a conditional logistic regression model including the same 20 dummy parameters as those used for the C-TTO data. Nevertheless, we did not use the coefficients estimated from a conditional logit model because they were expressed on a latent arbitrary utility scale. We rescaled the DCE coefficients using the same parameter θ that was estimated in the hybrid model. This rescaling assumes that the C-TTO model coefficients are proportional to the DCE model coefficients. For more details on the modelling see Ramos-Goñi et al. [23, 37].

Pearson product-moment correlation analysis was applied to measure the strength and direction of association that exists between C-TTO, DCE rescaled and hybrid predicted values for 3125 health states.

3 Results

3.1 Respondent Characteristics

In total, 1056 of 1117 respondents who were approached after the retraining of the interviewers completed the interview. Reasons for interview failure were refusal to participate (36, 3.2%), conflicting schedules (14, 1.25%), discontinuation of the interview at the respondent’s request (10, 0.89%), and discontinuation of the interview by the interviewer’s decision because of the respondent’s lack of understanding (1, 0.09%). From the remaining 1056 respondents, we excluded two respondents who had a positive slope on the regression between their values on

C-TTO and the sum score of the level digits of the health states, indicating that the respondent provided higher utility values for poorer health states on average, leaving 1054 respondents in the final sample. No interviewers were removed because of persistent low-quality data.

Characteristics of the respondents in the final sample were similar to those of the Indonesian population in terms of residence, gender, and religion. There were some statistically significant differences in some of the age groups, education levels, and ethnicities, but the absolute differences are small as these are <4% (Table 1).

Table 1 Characteristics of the study respondents/general population

Characteristics	Study sample (<i>N</i> = 1054) <i>n</i> (%)	Indonesian general population ^a (%)	Differences (%)
Residence			
Urban	549 (52.09)	53.30	-1.21
Rural	505 (47.91)	46.70	+1.21
Gender			
Female	526 (49.91)	49.65	+0.26
Male	528 (50.09)	50.35	-0.26
Age			
17-19	159 (15.09)*	12.35	+2.74
20-29	236 (22.39)	24.37	-1.98
30-39	264 (25.05)	22.68	+2.37
40-49	180 (17.08)	18.08	-1.00
50-59	164 (15.56)*	11.84	+3.72
60-69	43 (4.08)*	6.36	-2.28
70+	8 (0.76)*	4.31	-3.55
Education			
Low	339 (32.16)*	35.18	-3.02
Middle	550 (52.18)	51.72	+0.46
High	165 (15.65)*	13.10	+2.55
Religion			
Islam	920 (87.29)	87.18	+0.11
Christian	103 (9.77)	9.86	-0.09
Others	31 (2.94)	2.96	-0.02
Ethnicity			
Jawa	441 (41.84)	40.22	+1.62
Sunda	199 (18.88)*	15.50	+3.38
Sumatera	128 (12.14)*	15.02	-2.88
Sulawesi	63 (5.98)*	8.09	-2.11
Madura—Bali	52 (4.93)	4.70	+0.23
Others	171 (16.22)	16.47	-0.25

* Significant difference at $\alpha = 0.05$ from *z* test

^a Data from Indonesian Bureau of Statistics (BPS)

3.2 Self-Reported Health Problems

Table 2 shows that the highest proportion of health problems was reported in the pain/discomfort dimension (39.66% reported ‘any problems’) and the lowest in the self-care dimension (1.9%). From the final sample, 464 (44.02%) reported no health problems on any dimension (‘11111’).

3.3 Data Characteristics

The 1054 respondents provided 10540 C-TTO observations (respondents valued 10 health states each). We excluded 45 observations because they were ‘irrational answers’: preferring life B (10 years in the corresponding health state, which is worse than full health) to life A (10 years in full health) and not shifting after his/her initial response was reconfirmed by the interviewer. Furthermore, 1033 observations that were pointed out by the respondents in the Feedback Module task were removed. Accordingly, the C-TTO dataset contained 9462 observations. Of these, 187 (1.97%) observations relayed the value 0, and another 3349 (35.39%) were negative values (see Fig. 1 for the histogram of the observed C-TTO values). The 86 observed mean C-TTO values ranged from -0.719 for state ‘55555’ to 0.909 for state ‘12111’. The mean observed values were negative for 29 health states out of 86 used in the C-TTO design (see Online resource 1 in the electronic supplementary material).

The DCE dataset comprised 7378 observations (all respondents completed seven paired comparisons). Twenty respondents (1.89%) answered with suspicious patterns: AAAAAAA (always chose the health state at the left of the screen),BBBBBBB (always chose the health state at the right of the screen), ABABABA or BABABAB; however, their responses were not excluded from the final dataset.

3.4 Modelling Results

There were 657 (6.92%) left-censored C-TTO observations: when respondent gave the lowest possible value (-1) for a health state in the C-TTO task. The Tobit C-TTO model results were logically consistent. Conditional logistic regression was used to model the DCE responses that were also logically consistent (we used the rescaled DCE coefficients). C-TTO and rescaled DCE predicted values for 3125 health states were correlated, as Fig. 2a shows ($r = 0.9881$, p value <0.0001). Table 3 shows that both sets of coefficients were in relative agreement; that is, the most important dimension was mobility and the least important was pain/discomfort. The hybrid model, which utilized both C-TTO and DCE data, was also in relative agreement with both C-TTO and DCE models. Figure 2b, c

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

	EQ-5D-5L descriptive system with scores in %				
	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression
No problems	92.03	98.11	89.18	60.34	65.75
Slight problems	6.74	1.71	9.68	36.53	28.18
Moderate problems	1.04	0.09	1.14	2.56	5.50
Severe problems	0.19	0.09	0.00	0.57	0.38
Unable/extreme problems	0.00	0.00	0.00	0.00	0.19
	Mean	SD	25th percentile	Median	75th percentile
VAS score	79.38	14.01	70.00	80.00	90.00

EQ EuroQol, VAS visual analogue scale

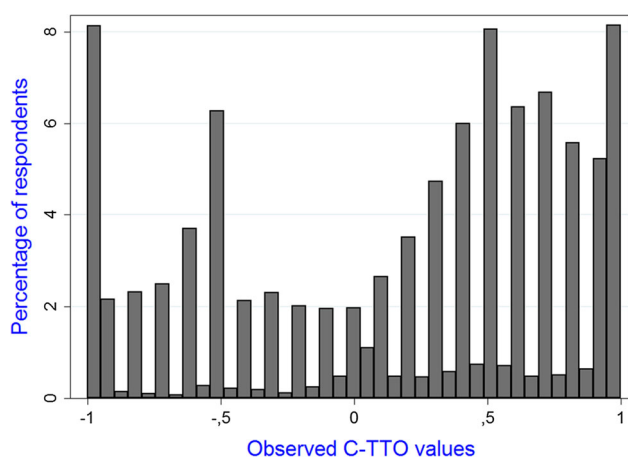


Fig. 1 Observed C-TTO values. C-TTO composite time trade-off

show a high correlation of hybrid predicted utility with models predicted from C-TTO ($r = 0.995$, $p < 0.0001$) and rescaled DCE ($r = 0.997$, $p < 0.0001$).

The hybrid model with main effects was logically consistent (Table 3). Using this as the final model to obtain 3125 EQ-5D-5L health states, the maximum value was 1.000 for full health (health state '11111') followed by the health state '11112' with value 0.921. The minimum value was -0.865 for the '55555' state. Of the 3125 health states, 1108 (35.46%) had negative values using the hybrid model. The coefficients from the hybrid model were also in agreement with the previous two models regarding mobility appearing as the most important dimension and pain/discomfort as the least important.

To obtain utility for an EQ-5D-5L health state, for instance '12345', the following calculation based on the hybrid model (final value set) is needed:

Utility weight ('12345') = 1 – no problems in MO (0) – no problems to slight problems in SC (0.101) – no problems to slight problems in UA (0.090) – slight problems to moderate problems in UA (0.066) – no problems to slight problems in PD (0.086) – slight problems to moderate problems in PD (0.009) – moderate problems to

severe problems in PD (0.103) – no problems to slight problems in AD (0.079) – slight problems to moderate problems in AD (0.055) – moderate problems to severe problems in AD (0.093) – severe problems to extreme problems in AD (0.078) = 0.240.

Note that each coefficient represents the additional utility decrement of moving from one level to another.

4 Discussion

The aim of this study was to obtain social preferences and thus derive an EQ-5D-5L value set from the Indonesian general population. To obtain values for 3125 EQ-5D-5L health states, 1054 respondents were interviewed using the computer-assisted EuroQol Group valuation protocol. C-TTO and DCE were part of the protocol employed in six cities and their surrounding areas. We used an iterative quality control approach in order to obtain high-quality data. The socio-demographic characteristics of the respondents were similar to those of the Indonesian population with respect to residence, gender, age, level of education, ethnicity, and religion. This makes EQ-5D-5L suitable for health economic evaluations that will benefit the national health insurance scheme. Furthermore, non-HTA studies in Indonesia such as those using patient-reported outcome measures (PROMs), clinical trials or improvements in hospital care could use EQ-5D-5L as an instrument to measure HRQOL, with the notion that the values attached to the health states are societal values.

Several limitations of this study should be considered. It could be argued that there are still statistically significant differences in the distribution of background variables in the sample compared with the data provided by the National Bureau of Statistics. There are statistically significant differences, but these are small, and limited to some age groups, some education levels, and some ethnicity groups. As a check to see if such small differences were of importance, we compared observed C-TTO values

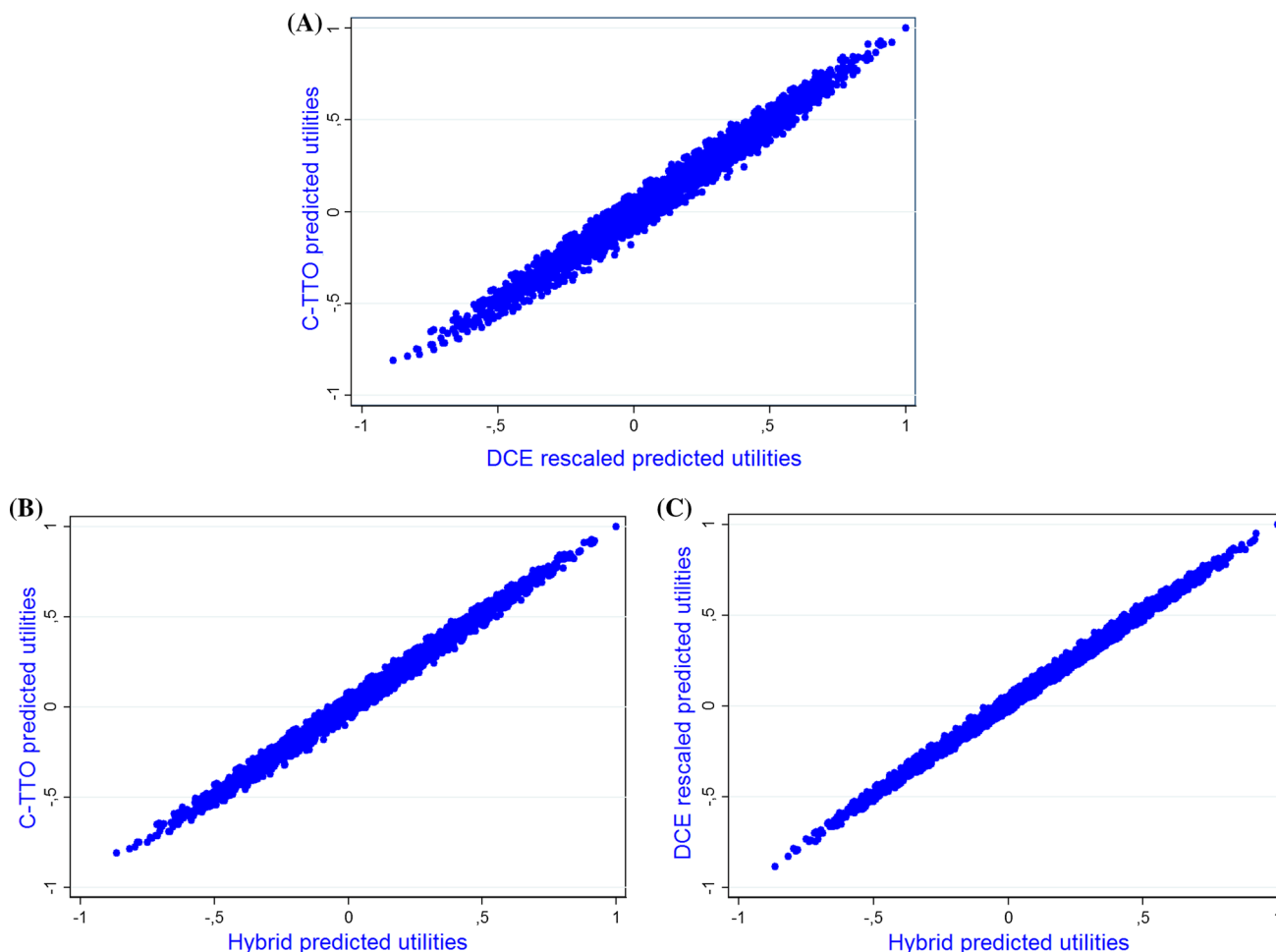


Fig. 2 a Comparison of C-TTO and DCE rescaled predicted utilities. b Comparison of C-TTO and hybrid predicted utilities. c Comparison of DCE rescaled and hybrid predicted utilities. *C-TTO* composite time trade-off, *DCE* discrete choice experiment

for each health state between respondents with different levels of age, education, and ethnicity. There was no clear pattern of differences in the health state values. Moreover, as can be seen in Table 1, the percentage deviations were small and statistical significance should be seen in the light of the statistical power of more than 1000 respondents. Given these observations, and given that weighting for background variables would add additional complexity, we chose not to introduce weighting for these small deviations from full representativeness.

The strategy of finding respondents using personal networks of the interviewers and the respondents could raise questions about the objectivity/representativeness of the study sample. Yet we preferred this way of recruitment in order to find respondents who fit into the pre-determined quota groups because we judged it to be a lesser problem than insufficiently filled categories in the quota sampling. The quota groups were determined on the variables that were considered to be important in defining representativeness. In that respect, we have constructed a

representative sample based on pre-determined variables: rural/urban, gender, age, level of education, religion and ethnicity. A further investigation could be conducted to find out whether recruiting respondents via personal networks of interviewers and/or respondents is not preferable or acceptable.

Indonesia has five major islands that are inhabited by 93.5% of the population [1]. However, 92.9% of respondents interviewed in this study were living on Java Island. This might raise questions about the representativeness of the study sample. However, we focused the data collection on Java island because it is the most populous island (57% of the population) and the main target of migration from all over Indonesia. The diversity of its residents in terms of ethnicity helps to fulfil all the categories in our quota sampling in a cost-effective way. We do not know whether the values obtained in Java from these migrants would have differed from the values should the interviews have been conducted on their original islands. One way to investigate whether location is indeed an issue in valuing health in

Table 3 Estimation results for C-TTO model, DCE rescaled model, and hybrid model

Independent variables of the model	C-TTO Tobit model censored at -1			DCE conditional logistic model rescaled			Hybrid model censored C-TTO values at -1 (final value set)		
	Coeff.	(SE)	<i>p</i> value	Coeff.	(SE)	<i>p</i> value	Coeff.	(SE)	<i>p</i> value
Mobility (MO)									
No problems to slight problems	0.088	(0.015)	0.000	0.139	(0.015)	0.000	0.119	(0.008)	0.000
Slight problems to moderate problems	0.086	(0.017)	0.000	0.080	(0.017)	0.000	0.073	(0.011)	0.000
Moderate problems to severe problems	0.250	(0.019)	0.000	0.196	(0.016)	0.000	0.218	(0.013)	0.000
Severe problems to unable	0.170	(0.018)	0.000	0.219	(0.018)	0.000	0.203	(0.012)	0.000
Self-care (SC)									
No problems to slight problems	0.085	(0.014)	0.000	0.101	(0.016)	0.000	0.101	(0.007)	0.000
Slight problems to moderate problems	0.056	(0.018)	0.002	0.038	(0.018)	0.032	0.039	(0.010)	0.000
Moderate problems to severe problems	0.128	(0.018)	0.000	0.085	(0.019)	0.000	0.108	(0.013)	0.000
Severe problems to unable	0.035	(0.016)	0.030	0.097	(0.017)	0.000	0.068	(0.012)	0.000
Usual activities (UA)									
No problems to slight problems	0.071	(0.015)	0.000	0.092	(0.016)	0.000	0.090	(0.006)	0.000
Slight problems to moderate problems	0.106	(0.017)	0.000	0.051	(0.017)	0.003	0.066	(0.011)	0.000
Moderate problems to severe problems	0.137	(0.019)	0.000	0.154	(0.017)	0.000	0.145	(0.013)	0.000
Severe problems to unable	0.061	(0.018)	0.001	0.091	(0.017)	0.000	0.084	(0.013)	0.000
Pain/discomfort (PD)									
No problems to slight problems	0.089	(0.013)	0.000	0.081	(0.016)	0.000	0.086	(0.006)	0.000
Slight problems to moderate problems	0.007	(0.019)	0.721	0.012	(0.018)	0.513	0.009	(0.011)	0.395
Moderate problems to severe problems	0.135	(0.018)	0.000	0.085	(0.017)	0.000	0.103	(0.013)	0.000
Severe problems to extreme problems	0.024	(0.019)	0.211	0.053	(0.018)	0.003	0.048	(0.013)	0.000
Anxiety/depression (AD)									
No problems to slight problems	0.079	(0.014)	0.000	0.050	(0.017)	0.003	0.079	(0.006)	0.000
Slight problems to moderate problems	0.055	(0.018)	0.002	0.061	(0.017)	0.000	0.055	(0.011)	0.000
Moderate problems to severe problems	0.086	(0.017)	0.000	0.114	(0.018)	0.000	0.093	(0.012)	0.000
Severe problems to extreme problems	0.062	(0.016)	0.000	0.085	(0.018)	0.000	0.078	(0.012)	0.000
Log likelihood	-6189.97			-3958.62			-9325.84		
AIC	12,421.93			7957.24			18,735.69		
BIC	12,572.19			8109.23			19,060.41		
Examples of estimated utility values									
U(21111)	0.912			0.861			0.881		
U(31111)	0.826			0.781			0.808		
U(41111)	0.576			0.585			0.590		
U(51111)	0.406			0.366			0.387		
U(12345)	0.225			0.268			0.240		
U(21231)	0.745			0.676			0.696		
U(55555)	-0.810			-0.884			-0.865		

AIC Akaike information criteria, *BIC* Bayesian information criteria, *C-TTO* composite time trade-off, *DCE* discrete choice experiments, SE standard error

Indonesia would be to sample values for health states at different places/islands in the republic. For instance, the same health states could be valued in Aceh (west), Java (middle) and Papua (east). Such a study could then be used to provide the motivation for additional studies that sample the values for people living in other parts of the

archipelago. For the time being, we conclude that the present value set is the best representative values set for the EQ-5D-5L now available for Indonesia.

Several study findings are worth highlighting. First, this is the first study in Asia to have used the hybrid model to maximize information obtained from C-TTO and DCE.

The models demonstrated logical consistency and significant regression coefficients. Two possible reasons that the data led to logical and significant models could be that the data were of high quality, which was assured by (1) the extensive use of the QC report provided by the EuroQol Group, and (2) the retraining programme conducted after dropping the first 102 interviews owing to their poor quality [30]. The QC report identified the first 102 interviews as problematic; indeed, further analysis using the hybrid model demonstrated that the results of these interviews showed logical inconsistencies in self-care and pain/discomfort dimensions, together with a regression coefficient that was not significant for pain/discomfort level 4 ($p = 0.179$). The lesson learned here is that even sophisticated models profit from high-quality data.

Second, the Indonesian results present more negative values than any other EQ-5D-5L valuation study undertaken so far (i.e. in the UK, the Netherlands, Canada, Uruguay, Japan and Korea [38–43]). It could be argued that the high number of negative values is the result of interaction between a process-related factor—quality control process and a cultural-related factor—interdependence among the members of a society (collectivism vs individualism). This study implemented the quality control process rigorously. It is possible that this quality control process provides the interviewer with better feedback and therefore better skills to administer the complex WTD trade-offs. Therefore, the more valid administration of the C-TTO means that more interviewers follow the protocol, which could have led to a higher proportion of negative values. The cultural factor, namely collectivism, might play a role. People from collectivistic cultures, such as Indonesia, are more concerned with how their illness might affect their closest circles such as family and friends [44]. Moreover, they are more reluctant to explicitly ask for help [45]. Some comments from our respondents support this: having severe or extreme/unable problems in EQ-5D dimensions was very bad for them individually, but will also be a burden for their closest circles (family and friends). For other respondents, they preferred to die than to bother their families and friends when they have a severe illness. The EQ-5D-3L value set of Singapore, a neighbouring country of Indonesia and a collectivistic country as well, showed the all-worst state ‘33333’ was -0.769 [46]. When more national valuation studies are published, it will be possible for a further investigation to disentangle the effect of these factors on proportion of WTD values in an EQ-5D-5L valuation study.

Third, we had a low level of non-response: only 61 of the 1117 respondents. Our recruitment strategy, which involved local leaders and asking respondents to recommend our study to other people, contributed to this low number.

Fourth, this study was performed according to the EuroQol Group’s EQ-5D-5L valuation protocol. Hence, the results are comparable to findings obtained in other countries. The final Indonesian value set shows that the mobility dimension influenced utility estimates the most, similar to EQ-5D-5L valuation study results from Uruguay and South Korea [41, 43]. The pain/discomfort dimension had the least influence on utility estimates, quite the opposite of the EQ-5D-5L value sets of England and the Netherlands where this dimension was in the top two most influential, after anxiety/depression [38, 39]. Perhaps this was because, in countries such as the Netherlands and the UK, problems with mobility had less influence due to better infrastructure provision and less emphasis on manual labour. It could also be argued that Indonesian people have adapted to mild levels of pain or discomfort, or perhaps they considered a mild level of pain or discomfort something they have to live with. The same line of reasoning applies to anxiety/depression. Indonesian people report more problems with anxiety/depression and have adapted to these mild levels of anxiety/depression, or they consider this as part of normal life. It could also be a result of small differences in translation. If the translated Indonesian words for depression and anxiety refer to a lighter problem, then it makes sense that the prevalence was higher and the disutility lower. Indeed, there are some indications that this was the case. In the Indonesian EQ-5D translation, the word ‘sedih’, which might also be translated as ‘sadness’, is added to the description of the anxiety/depression dimension. These kinds of interactions between the description of the dimensions and the values attached justify attempts to utilize local and linguistically matched value sets for utility questionnaires such as EQ-5D. If not, value sets based on other languages might apply the wrong (higher) utilities to the descriptors. For instance, it is now clear that one cannot use the UK value for anxiety/depression for the Indonesian descriptor with an additional word ‘sadness’.

Several policy implications of the present study can be considered. The finding that the mobility dimension most affects utility could be implemented in Indonesian government policies, such as allocating more funds to the prevention of diabetic foot ulcers or other interventions that improve mobility like better wheelchairs. Moreover, the anxiety and depression problems reported should be addressed. If so, the discussion concerning the translation of the anxiety and depression dimension mentioned in the paragraph above should be taken into account. If indeed anxiety and depression are such common afflictions in Indonesia, mental health treatment by professionals such as psychologists and psychiatrists within the national health insurance scheme should be considered.

Indonesia is endeavouring to implement HTA comprehensively. The present research shows that in measuring

and valuing quality of life, Indonesia bears comparison with the leading countries employing HTA. Evidently Indonesia still has ground to conquer when dealing with models, cost data and decision making. Nevertheless, this research shows that it is possible to arrive at an established level of HTA methodology in a short time-span when cooperating at an international level.

5 Conclusion

This paper contains the EQ-5D-5L value set for Indonesia based on Indonesian citizens' health preferences. We expect our results to promote and facilitate health economic evaluations in Indonesia which can help to inform decision makers concerning resource allocation decisions.

Data Availability Statement The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements The Dutch and Indonesian non-profit organization 'Van Deventer-Maas Stichting (VDMS)' and the 'Indonesia Bhadra Utama (IBU) Foundation' respectively, helped us with the data collection. The paper benefited from the efforts, insights and annotations of Elly Stolk and Arnd Jan Prause. Any errors or omissions are the responsibility of the authors only.

Author contribution JB, JP and JH designed the study and acquired the funding. SS and AI provided detailed information regarding data collection processes in Indonesia during the study design phase. FDP coordinated the data acquisition, assisted by TSF and AI. FDP and JMR conceived the analysis. FDP prepared the first draft of the manuscript. All authors played a role in the review of the analysis, interpretation of the results, and reviewed and recommended revisions to the final submitted manuscript to ensure accuracy and fair balance.

Compliance with Ethical Standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee [Padjadjaran University Health Research Ethics Committee (no: 065/UN6.C1.3.2/KEPK/PN/2015)].

Funding The research was financed by the EuroQol Group and Directorate General of Higher Education of Indonesia (<http://www.dikti.go.id>). Both grants were unrestricted.

Conflicts of interest The following authors are members of the EuroQol Research Foundation (the copyright holders of EQ-5D-5L): JB and JMR. There are no other conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Badan Pusat Statistik. Statistik 70 Tahun Indonesia Merdeka. Badan Pusat Statistik, Jakarta 2015. <http://bps.go.id/index.php/publikasi/1053>. Accessed 3 Nov 2015.
2. bpjs-kesehatan.go.id. Vision and Mission. BPJS Kesehatan. 2010. <http://bpjs-kesehatan.go.id/bpjs/index.php/pages/detail/2010/2>. Accessed 4 Nov 2015.
3. President of Republik Indonesia. Peraturan Presiden Republik Indonesia Nomor 12 Tahun 2013 Tentang Jaminan Kesehatan Jakarta. 2013. <http://bpjs-kesehatan.go.id/bpjs/index.php/unduh/index/3>. Accessed 14 Nov 2015.
4. World Health Organization. Countries with National agency/unit/committee that produces HTA reports for the Ministry of Health. World Health Organization. 2015. <http://www.who.int/health-technology-assessment/NationalAgencieHTA.pdf>. Accessed 5 Nov 2015.
5. Drummond M, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. 4th ed. Oxford: Oxford University Press; 2015.
6. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. *Qual Life Res.* 2015;24(8):2045–53.
7. Richardson J, Khan MA, Iezzi A, Maxwell A. Comparing and explaining differences in the magnitude, content, and sensitivity of utilities predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D multiattribute utility instruments. *Med Decis Mak.* 2015;35(3):276–91.
8. Chen G, Khan MA, Iezzi A, Ratcliffe J, Richardson J. Mapping between 6 multiattribute utility instruments. *Med Decis Mak.* 2016;36(2):160–75.
9. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. *BMJ.* 2004;329(7459):224–7.
10. Sakthong P. Measurement of clinical-effect: utility. *J Med Assoc Thai.* 2008;91(Suppl 2):S43–52.
11. Brooks R. EuroQol: the current state of play. *Health Policy.* 1996;37(1):53–72.
12. 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(10):1727–36. doi:10.1007/s11136-011-9903-x.
13. Feng Y, Devlin N, Herdman M. Assessing the health of the general population in England: how do the three- and five-level versions of EQ-5D compare? *Health Qual Life Outcomes.* 2015;13(1):171.
14. Pattanaphesaj J, Thavorncharoensap M. Measurement properties of the EQ-5D-5L compared to EQ-5D-3L in the Thai diabetes patients. *Health Qual Life Outcomes.* 2015;13:14.
15. Conner-Spady BL, Marshall DA, Bohm E, Dunbar MJ, Loucks L, Khudairy AA, et al. Reliability and validity of the EQ-5D-5L compared to the EQ-5D-3L in patients with osteoarthritis referred for hip and knee replacement. *Qual Life Res.* 2015;24(7):1775–84.
16. Golicki D, Niewada M, Karlinska A, Buczek J, Kobayashi A, Janssen MF, et al. Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQ VAS in stroke patients. *Qual Life Res.* 2015;24(6):1555–63.
17. Greene ME, Rader KA, Garellick G, Malchau H, Freiberg AA, Rolfson O. The EQ-5D-5L improves on the EQ-5D-3L for health-related quality-of-life assessment in patients undergoing total hip arthroplasty. *Clin Orthop Relat Res.* 2015;473(11):3383–90.
18. Agborsangaya CB, Lahtinen M, Cooke T, Johnson JA. Comparing the EQ-5D 3L and 5L: measurement properties and

- association with chronic conditions and multimorbidity in the general population. *Health Qual Life Outcomes*. 2014;12:74.
19. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health*. 2014;17(4):445–53. doi:10.1016/j.jval.2014.04.002.
 20. Endarti D, Riewpaiboon A, Thavorncharoensap M, Praditsithikorn N, Hutubessy R, Kristina SA. Evaluation of health-related quality of life among patients with cervical cancer in Indonesia. *Asian Pac J Cancer Prev*. 2015;16(8):3345–50.
 21. Iskandar S, van Crevel R, Hidayat T, Siregar IM, Achmad TH, van der Ven AJ, et al. Severity of psychiatric and physical problems is associated with lower quality of life in methadone patients in Indonesia. *Am J Addict*. 2013;22(5):425–31.
 22. Goudarzi R, Zeraati H, Akbari Sari A, Rashidian A, Mohammad K. Population-based preference weights for the EQ-5D health states using the visual analogue scale (VAS) in Iran. *Iran Red Crescent Med J*. 2016;18(2):e21584.
 23. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serano-Aguilar P, Rivero-Arias O. Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med Care*. 2017;55(7):e51–e8.
 24. Rabin R, Gudex C, Selai C, Herdman M. From translation to version management: a history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. *Value Health*. 2014;17(1):70–6.
 25. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35(11):1095–108.
 26. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005;43(3):203–20.
 27. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goni JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. 2016;34(10):993–1004.
 28. Stolk EA, Oppe M, Scalone L, Krabbe PF. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. *Value Health*. 2010;13(8):1005–13.
 29. Ramos-Goñi JM, Rivero-Arias O, Errea M, Stolk EA, Herdman M, Cabasés JM. Dealing with the health state 'dead' when using discrete choice experiments to obtain values for EQ-5D-5L health states. *Eur J Health Econ*. 2013;14(Suppl 1):S33–42. doi:10.1007/s10198-013-0511-2.
 30. 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(5):1197–208. doi:10.1007/s11136-016-1445-9.
 31. Busschbach JJ, McDonnell J, Essink-Bot ML, van Hout BA. Estimating parametric relationships between health description and health valuation with an application to the EuroQol EQ-5D. *J Health Econ*. 1999;18(5):551–71.
 32. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJV, Stolk E. Quality control process for EQ-5D-5L valuation studies. *Value Health*. 2016. doi:10.1016/j.jval.2016.10.012.
 33. Bleichrodt H, Pinto JL. Loss aversion and scale compatibility in two-attribute trade-offs. *J Math Psychol*. 2002;46(3):315–37. doi:10.1006/jmps.2001.1390.
 34. Hawkins SA. Information processing strategies in riskless preference reversals: the prominence effect. *Organ Behav Hum Decis Process*. 1994;59(1):1–26. doi:10.1006/obhd.1994.1048.
 35. Oppe M, van Hout B. The optimal hybrid: Experimental design and modeling of a combination of TTO and DCE. In: 27th scientific plenary meeting of the EuroQol Group-Proceedings. EuroQol Group Executive Office, Athens, Greece; 2010.
 36. Attema AE, Versteegh MM, Oppe M, Brouwer WB, Stolk EA. Lead time TTO: leading to better health state valuations? *Health Econ*. 2013;22(4):376–92.
 37. Ramos-Goñi JM, Craig BM, Oppe M, van Hout B. Combining continuous and dichotomous responses in a hybrid model. EuroQol Research Foundation. 2016. http://www.euroqol.org/eq-5d-publications/working-paper-series/file/EuroQol_Working_Paper_Series_Manuscript_16002_-_Juan_Ramos-Goni_01.pdf.html Accessed 2 Jan 2017.
 38. Devlin N, Shah K, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. OHE Research Paper. 2015. <https://www.ohe.org/publications/valuing-healthrelated-quality-life-eq-5d-5l-value-set-england>. Accessed 4 Nov 2015.
 39. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19(4):343–52. doi:10.1016/j.jval.2016.01.003.
 40. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. *Med Care*. 2016;54(1):98–105.
 41. Augustovski F, Rey-Ares L, Irazola V, Garay OU, Gianneo O, Fernandez G, et al. An EQ-5D-5L value set based on Uruguayan population preferences. *Qual Life Res*. 2016;25(2):323–33.
 42. Ikeda S, Shiroya T, Igarashi A, Noto S, Fukuda T, Saito S, et al. Developing a Japanese version of the EQ-5D-5L value set. *J Natl Inst Public Health*. 2015;64:47–55.
 43. Kim SH, Ahn J, Ock M, Shin S, Park J, Luo N, et al. The EQ-5D-5L valuation study in Korea. *Qual Life Res*. 2016;25(7):1845–52. doi:10.1007/s11136-015-1205-2.
 44. Uskul AK, Hynie M. Self-construal and concerns elicited by imagined and real health problems. *J Appl Soc Psychol*. 2007;37(9):2156–89.
 45. Kim HS, Sherman DK, Taylor SE. Culture and social support. *Am Psychol*. 2008;63(6):518–26.
 46. Luo N, Wang P, Thumboo J, Lim YW, Vrijhoef HJ. Valuation of EQ-5D-3L health states in Singapore: modeling of time trade-off values for 80 empirically observed health states. *Pharmacoeconomics*. 2014;32(5):495–507. doi:10.1007/s40273-014-0142-1.