Do health preferences differ among Asian populations? A comparison of EQ-5D-5L Discrete Choice Experiments data from 11 Asian studies

Zhihao Yang (✉ zhihao_yang_cn@126.com)

Guizhou Medical University <https://orcid.org/0000-0001-5468-0847>

Fredrick Purba

Universitas Padjadjaran

Asrul Akmal Shafie

Universiti Sains Malaysia

Ataru Igarashi

University of Tokyo Graduate School of Pharmaceutical Sciences Faculty of Pharmaceutical Sciences: Tokyo Daigaku Daigakuin Yakugakukei Kenkyuka Yakugakubu

Eliza Wong

The Chinese University of Hong Kong

Hilton Lam

University of the Philippines Manila

Hoang Van Minh

Haoni University of Public Health

Hsiang-Wen Lin

China Medical University

Jeonghoon Ahn

Ewha Womans University

Juntana Pattanaphesai

Ministry of Public Health Thailand

Min-Woo Jo

University of Ulsan College of Medicine

Vu Quynh Mai

Hanoi University of Public Health

Jan Busschbach

Erasmus Medical Centre: Erasmus MC

Nan Luo

National University of Singapore

Jie Jiang

Jinan University

Research Article

Keywords:

Posted Date: August 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-180271/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Quality of Life Research on February 18th, 2022. See the published version at <https://doi.org/10.1007/s11136-021-03075-x>.

Abstract

Introduction For establishing EQ-5D-5L value set, the EuroQol Group developed a standardized EQ-VT protocol. Both time-trade off (TTO) method and discrete choice experiment (DCE) method are used in the EQ-VT protocol. Published studies did not make use of the DCE data for comparison purpose. This study aims to compare the health preferences among 11 Asian studies using the DCE data collected in their EQ-5D-5L valuation studies. **Methods** In the EQ-VT protocol, 196 pairs of EQ-5D-5L health states were valued by a general population sample using DCE method for all 11 studies. DCE data was obtained from the study PI. Three different main-effects models were fitted for each study. Coefficients were first tested between studies. Next, the relative importance of dimensions and levels was calculated and compared. **Results** The number of statistically differed coefficients ranged from 5 to 16, out of 20 main-effects coefficients. For the relative importance, there is not a universal preference pattern that fits all studies, but with some common characteristics e.g. mobility is the most weighted dimension except for Vietnam; the relative importance of levels were approximately 20% for level 2, 30% for level 3, 70% for level 4 across all studies. **Discussion** This study confirmed that health preferences heterogeneity among Asian populations. It is therefore justifiable to announce that national/regional value sets should be used for calculating health utility.

Introduction

EQ-5D is a generic preference-based health related quality of life (HRQoL) questionnaire that is widely used around the world (1, 2). When value sets are available, EQ-5D data can be converted to health utility (2). Many countries have established their own EQ-5D value sets proceeding on the basis that health preferences differ among countries/populations (3, 4). Indeed, studies have found differences between value sets (3, 5, 6). In developing the value sets of three-level version of EQ-5D (EQ-5D-3L), published studies differed in terms of design, data collection protocol and the choice of model. By comparing these value sets, Norman et al concluded that these variations in methods could obscure true differences in values (6). For the latest five-level version of EQ-5D (EQ-5D-5L), the EuroQol Group developed a standardized protocol for data collection in valuation studies, which is named the EuroQol valuation technology protocol (EQ-VT) (7-9).

With application of the EQ-VT, EQ-5D-5L valuation data can be exploited to study whether important differences in health preferences across populations exist, as the method variations observed in the 3L studies are minimized. The EQ-VT data collection protocol uses both time trade-off (TTO) and discrete choice experiment (DCE) as preference elicitation methods (7). Currently, all comparison studies of EQ-5D value sets only used the TTO data from the valuation studies (5, 10, 11). This is partially because the TTO data is considered as the primary preference source in the EQ-VT protocol and some studies estimated their value set using TTO data only, for example, China and South Korea (12, 13). So far, the DCE data collected using the EQ-VT protocol has not been utilized for the purpose of identifying preference differences across studies. While the TTO valuation data could be subject to interviewer effects as the task relies on the good performance of the interviewers (14), there is minimal interviewer effect for the DCE data.

As a preference elicitation method, DCE has been increasingly used in health preference studies (15). Based on random utility theory, DCE is designed to ask respondents to choose a preferred multi-dimensional health state from two or more alternatives. The ordinal preference data can be modelled to predict health utility on a latent scale (16). This means that the coefficients of DCE are not directly comparable across studies and most studies assessed their difference by calculating and comparing the relative importance of five health dimensions (17, 18).

Further, differences in health preferences among Asian populations are not well understood. By comparing the multiplicative model coefficients of the EQ-5D-5L TTO valuation data from seven Asian studies, Wang et al noticed that there was no consensus about the rank ordering of the five dimensions (10). Additionally, statistical test suggested most coefficients differed among Asian studies. In the study of Roudijk et al, the authors found that cultural variables (i.e. traditional/rational-secular, survival/self-expression) did not explain the variations of value differences (defined as utility differences between the mild and severe states) among EQ-5D valuation studies, including 10 Asian studies (11). As stated before, these studies only explored the TTO valuation data.

Following the EQ-VT protocol, 11 studies (China, Indonesia, Japan, South Korea, Malaysia, Singapore, Thailand, Philippines, Vietnam, Hong Kong, Taiwan) have been completed in Asia. Of those studies, China and South Korea did not use the DCE data to model the value sets, and Singapore and Philippines have not yet published their value sets. The rest of the studies modelled the DCE data and TTO data jointly. Notably, no study has compared DCE-derived preference data among Asian populations. Given all studies used the standardized EQ-5D-5L instrument, DCE experimental design, and data collection protocol, it is possible to explore the variations of health preferences in Asia. We hypothesized that the health preferences differed in Asian populations. If this is true, the results of this study could further support the establishment of national/regional value sets for better guidance of healthcare decision making and resource allocation rather than using a unified value set designed merely for the continent. In this study, we aim to understand the similarities and differences in Asians' preferences for EQ-5D-5L health states in 11 Asian DCE datasets collected as part of EQ-5D-5L valuation studies.

Methods

2.1 EQ-5D-5L questionnaire & 11 Asian valuation studies

The EQ-5D-5L measures HRQoL using five dimensions (mobility, self-care, usual activities, pain/ discomfort, anxiety/depression), and under each dimension, there are five levels of severity (no problems, slight problems, moderate problems, severe problems, extreme problems/unable to). In total, the EQ-5D-5L defines $5^5=3,125$ health states (19). We obtained the DCE data from the principal investigators of 11 Asian EQ-5D-5L valuation studies, namely China (12), Indonesia (20), Japan (21), South Korea (13), Malaysia (22), Singapore, Thailand (23), the Philippines, Vietnam (24), Hong Kong (25), and Taiwan (26). The language(s) being used in each study is the official language and/or dialects that are spoken mostly by the population (see Table 1).

2.2 DCE design & tasks

All Asian studies included in this study used the standard EQ-VT protocol for data collection (9, 27). In general, the DCE design of the EQ-VT protocol consisted of a total of 196 pairs of health states including 186 pairs generated from a Bayesian efficient design algorithm and 10 pairs of mild states (27). The priors for the Bayesian efficient design algorithm were extracted from a main effects model of an EQ-5D-3L DCE study (28). The detailed experimental design development process and considerations were described in Oppe et al (27). The 196 pairs of EQ-5D-5L health states were distributed over 28 blocks, each consisting of 7 pairs of health states with similar severity. No dominant pairs were included. (27). In each study, each respondent was assigned one block of DCE tasks to complete. The 7 pairs were presented in random order, and the right-left presentation of the two health states was also randomized (8). Figure 1 shows the screenshot of one DCE task in EQ-VT software.

2.3 Data collection

Following the EQ-VT protocol, all respondents were interviewed face-to-face by a trained interviewer using the EQ-VT software. The data collection included four sections: The first section was for respondents to report their own health using the EQ-5D-5L descriptive system and the EQ-VAS. In the second section, respondents valued 10 different EQ-5D-5L health states using the composite time trade-off (cTTO) (8). In the third section, respondents completed 7 pairs of EQ-5D-5L discrete choice tasks (27). Finally, respondents reported their socio-economic and other background characteristics. We used the DCE data obtained from the third section for the analysis.

2.4 Analysis

To understand how the health preferences are different/similar with each other, the following analyses were done: 1) the statistical difference between the coefficients; 2) the relative importance of the five EQ-5D dimensions; 3) the utility decrements between each of the response levels.

For modelling, a 20-parameter main-effects conditional logit model was fitted for each study. In this model (Formula 1), utility was explained by 20 dummy variables and was on a latent scale (referred as latent utility). For each dimension (MO for mobility, SC for self-care, UA for usual activity, PD for pain/discomfort, AD for anxiety/depression), 4 dummy variables were used to represent the departure from level 1 to the other 4 levels, e.g. MO_3 was 1 if the health state being valued had “moderate problems with mobility” and 0 for any other level of mobility (29). To acknowledge the preference heterogeneity within each study, a heteroscedastic model and a mixed effect logit model were also fitted for each study (30). Model specifications of these two models are the same with the conditional logit model.

$$Latent\ utility = \beta_1 MO_2 + \beta_2 MO_3 + \beta_3 MO_4 + \beta_4 MO_5 + \beta_5 SC_2 + \dots + \beta_{20} AD_5 + \varepsilon \quad (1)$$

Next, the statistic difference between two studies' coefficients were explored using a pairwise comparison. For each pair, a dummy variable was generated as 0 for one region's data and as 1 for the other. Next, a 20-parameter main-effects model plus 20 interaction terms (multiplied the study dummy variable with 20 main-effects variables, see Formula 2). In this model with interaction terms, a significant interaction term suggests that the coefficient is statistically different between two studies. The number of statistically differed coefficients were summarized for each pair. Notably, the coefficient of a significant interaction term may not exceed the minimal important difference (MID) on the utility scale.

$$Latent\ utility = \beta_1 MO_2 + \beta_2 MO_3 + \beta_3 MO_4 + \beta_4 MO_5 + \beta_5 SC_2 + \dots + \beta_{20} AD_5 + \beta_{21} MO_2 * region\ dummy + \beta_{22} MO_3 * region\ dummy + \beta_{23} MO_4 * region\ dummy \dots + \beta_{40} AD_5 * region\ dummy + \varepsilon \quad (2)$$

Using the main-effects mixed effect logit model results (Formula 1), the relative importance of dimensions and levels were estimated for each study (17, 18, 31). The relative importance of the five dimensions were calculated in two steps. First, the dimension-level coefficient was divided by the mean of the same level from all the dimensions. For example, the adjusted coefficient for mobility level 3 was obtained by the MO_3 coefficient divided by the mean of all level 3 coefficients for each dimension $aMO_3 = MO_3 / [(MO_3 + SC_3 + UA_3 + PD_3 + AD_3) / 5]$. This step resulted in adjusted coefficients for the last four levels (level 1 is the

reference level) of every dimension. Second, the means of all adjusted coefficients for each dimension were calculated. Continuing the mobility example, the relative dimension importance of mobility for a study would be estimated as $(aMO_2 + aMO_3 + aMO_4 + aMO_5)/4$.

The relative importance of levels was also obtained in two steps: first, the sum of each level coefficient from all dimensions was calculated. Second, the sum of each level coefficient was divided by the sum of level 5 coefficients: e.g., the relative importance of level 2 was the sum coefficient of level 2 divided by the sum coefficient of level 5. In practice, relative importance for level 2 sum for a study would be calculated as follows: $(MO_2 + SC_2 + UA_2 + PD_2 + AD_2) / (MO_5 + SC_5 + UA_5 + PD_5 + AD_5)$. To compare the relative importance results across studies, two figures were plotted, one for the relative importance of the dimensions and one for the relative importance of levels. If five dimensions are equally weighted by a population, all five dimensions should have a relative importance of 0.20 (i.e., 1 divided by 5). The relative importance of levels are interpreted as the percentage of the weight attached to level five problems. Analyses were performed using STATA 14 (Stata Corp LLC) (32).

Results

3.1 Data descriptions

Table 1 summarizes the key information from the 11 valuation studies. Based on the EQ-VT protocol, all studies recruited at least 1,000 respondents. Quota sampling was the most used sampling strategy, but the quota differed. All studies were conducted between 2012 and 2017.

3.2 Modelling results

Table 2 shows the conditional logit modelling results. All coefficients for all studies were significant at 0.05 level except for the second level of self-care in South Korea. Thailand, Vietnam and Philippines each had 1, 1 and 2 inconsistent coefficients, respectively. All inconsistencies occurred on the third level of self-care, mobility and usual activities. Within each study, the standard errors of the coefficients generally increased with severity levels. Table 3 shows the number of coefficients that differed statistically between two studies. Overall, coefficients differed significantly among these studies. Almost all studies had at least 5 coefficient differences with others except for Taiwan versus Hong Kong, Taiwan versus Malaysia, and Japan versus China. Malaysia and Singapore differed the most with 16 statistically different coefficients. An example of this comparison between China and Indonesia can be found in Appendix 1.

Compared with the conditional logit model results, the heteroscedastic conditional logit model did not improve the coefficient non-significance and inconsistency I. The mixed effect logit model corrected one inconsistency for Thailand and improved one non-significance for South Korea, but led to one inconsistency for Taiwan. The modelling results of these two models can be found in Appendix 2 & 3.

3.3 Relative weight results

Figure 2 shows the relative importance of five dimensions across 11 studies. Overall, a universal rank order does not exist across 11 Asian populations. Mobility was the most important dimension for every study except for Vietnam. The lowest important dimension was either usual activity or self-care except for Philippines and Indonesia. Notably, these two functional dimensions had similar weights in China, Indonesia, Japan and Vietnam (their lines were parallel and close to each other), and only Korea had larger relative weight for usual activities. Pain/discomfort was the second most important dimension for 6 studies and it was valued higher than or equal to anxiety/depression in almost all studies except for Thailand. Singapore, Japan, Philippines and Indonesia placed similar weights on pain/discomfort and anxiety/depression. The sum of the first three functional dimensions were larger than the sum of the two symptom dimensions across all studies.

Some individual characteristics can be spotted from Table 2. Korea showed the largest difference between the dimensions of mobility and self-care. Japan had similar weights for dimensions other than mobility. Hong Kong, Malaysia and Taiwan showed similar rank order, i.e., Mobility > Pain/discomfort > Anxiety/depression > Self-care > Usual activities. China differed with these three studies by placing usual activities more important than self-care. Indonesia showed a different pattern by weighing more on usual activities and self-care over pain/discomfort and anxiety/depression. Both Vietnam and Singapore had similar weights three dimensions. Thailand and Vietnam were unique in the sense that Thailand valued anxiety/depression as the second most important dimension and Vietnam valued pain/discomfort as the most important dimension.

Compared with the large variations among the relative importance of health dimensions, the relative importance of levels were more comparable across studies. The weights of mild (L2) and moderate problems (L3) were more similar across regions as compared to the weights of severe (L4). The L2 ranges from 0.156 for Taiwan to 0.322 for the Philippines; the L3 ranges from 0.211 for Thailand to 0.367 for Indonesia; the L4 ranges from 0.600 for South Korea to 0.837 for the Philippines. In the Philippines and Thailand, the difference between level 2 and level 3 were minimal. On average, level 2 accounted for 20% of the weight of level 5, level 3 accounted for approximately 30% of the weight of level 5 and level 4 accounted for 70% of the weight of the level 5. The smallest relative importance was 0.156 of L2 from Taiwan, which represents having a mild problem accounted for about 15.6% weight of having an extreme problem. The smallest L3 was from Thailand (0.217), and this value was smaller than L2 from some studies.

Discussion

The present study compared 11 regional health preferences for EQ-5D-5L health states elicited using DCE method. By comparing preferences elicited using standardized health description system and data collection protocol, our results provide insights on the variations of health preference in Asia (10). Based on our results, it is fair to declare that there does not exist a single preference pattern for all 11 studies. This is in line with a previous study comparing TTO model for 7 Asian populations (10). A clear distinction between our DCE results and the TTO results is the relative weights for level 3 and level 4 are larger in that the TTO study.

In general, the coefficients of the DCE models differed across populations. In the conditional logit model, the number of differed coefficients ranged between 3 (Malaysia vs Taiwan, Hong Kong vs Taiwan) to 16 (Singapore vs Malaysia). In terms of the relative importance, both the relative importance of levels and the relative importance of dimensions varied across studies. Only Hong Kong, Taiwan and Malaysia showed same order of five dimensions. Here we concluded some common patterns, which, however, always come with exceptions. First, among the five dimensions, mobility is the most important dimension for every population except for Vietnam. This is similar to the results from a comparison of TTO-only preference data from 7 Asian regions (10). However, western countries do not value mobility as highly; the Dutch, German, and US populations view mobility as third, fourth, and second most important dimension, respectively (33-35). Purba et al argued that in the western developed countries, problems with mobility had less influence due to better infrastructure provision and less emphasis on manual labour (20). However, in high income and developed regions such as Singapore and Japan, mobility is still the most valued dimension. Second, the sum of three function dimensions (mobility, self-care and usual activities) were higher than the sum of two symptom dimensions (pain/discomfort and anxiety/depression). Also, either usual activities or self-care is the least important dimension. Indonesia and Philippines are exceptions. This result agrees with the previous study comparing the TTO data results among 7 Asian populations. In that study, Indonesia was the only one who valued pain/discomfort and anxiety/depression the lowest. This result highlights the importance of functional dimensions among Asian populations. Third, pain/discomfort was valued more important than anxiety/depression and is the second most important dimensions for 6 studies. These characteristics mark some notable difference between preference pattern from most European, American and African populations (5, 36-38).

Despite these similarities, it is clear that a singular preference pattern does not exist for all Asian populations. For example, there is no agreement on the least important dimension in our comparison—3 studies valued self-care, 2 studies valued anxiety/depression, and 6 studies valued usual activities as the least important. This contrasts to a previous study of comparing health preference pattern for Canada, England, the Netherlands and Spain. In that study, Olsen et al found a clear pattern existed for these four western countries and named it western preference pattern (WePP) (5). In the WePP, four general characteristics were noticed in terms of the relative importance: 1) $(PD + AD) \approx (MO + SC + UA)$; 2) $PD \approx AD$; 3) $MO \approx SC$; 4) $UA < SC$. However, no Asian preferences fit well with these four characteristics. In fact, the sum of pain/discomfort and anxiety/depression was less than the weight of the other three dimensions in all Asian studies: $(PD + AD) < (MO + SC + UA)$, suggesting that compared with the western countries, the Asian placed more weights on the functional dimensions. The second characteristic of ' $PD \approx AD$ ' was only observed in the results from Indonesia, Singapore and Malaysia. The third characteristic was clearly invalid in Asia as mobility was valued as the most important dimension while self-care had less relative importance in 4 studies. For the last characteristic, four Asian populations put similar or higher values for usual activities.

From our results, we have seen the variations of health preference across Asian populations. We hypothesize some reasons for these differences. First and probably the most important reason is: culture influences people's health preferences. The EQ-5D dimensions may not be equally important in all cultures. Some difference could be attributed to the study design. Though a standard data collection protocol was followed by each study, there were variations in sampling methods. Quota sampling method was the most used sampling strategy and different quota were set for different study. For example, ethnicity was used in some studies like Malaysia and Singapore, but not in China and South Korea. Similarly, some studies only recruited participants from urban areas, which may not be representative for the whole target population. Studies have shown that the demographic of respondents could influence the health preferences (39, 40). Hence, different respondents recruited for each study may contribute to the observed differences. Last but not least, the EQ-5D-5L descriptive system was translated into different official languages from English. Though a standardized translation process was conducted to maintain equivalence between the translated questionnaire and its source version, different cultures have different ways of expression which maybe inadequately captured (41).

This study has some limitations. First, the relative weights of dimensions and levels were used to present the differences in a straightforward way. The relative weights were only point estimates without consideration of uncertainties around the coefficients, and the differences among the dimensions were not statistically assessed. However, this method is widely used for comparing DCE results (17, 18, 30) and we tested the coefficients difference for all two-by-two study pairs prior to the relative weight results. Second, even though a standardized protocol was used, the demographic questions used for each study was customized by each local study team. Due to these sampling variations, we did not further test how these variations affect preferences. Only the heteroscedastic model shown in Appendix 2 demonstrates that the variances was constant for respondents with different age and gender.

In an earlier study comparing EQ-5D-3L value sets, Norman et al pointed out that differences in methods obscured the true differences in health preferences across countries (6). In this study, we have shown that by using a standardized data collection protocol, study design and modelling choice, there still remained differences in EQ-5D-5L modelling results, dimension weights and level weights across Asian populations. There has been an effort to estimate a joint value set by merging valuation data in European and Western countries (5, 42). Based on our results, we do not recommend such an approach for Asia.

Conclusion

By comparing the DCE data modelling results, we found that the DCE modelling results differed among Asian populations. The rank order of EQ-5D-5L dimensions and the relative weight of levels also differed. Our results confirmed the health preference heterogeneity among Asian populations that was observed in previous studies using TTO data. All the evidence suggested the necessity of using local value set for estimating health utility.

Declarations

This study was funded by the EuroQol Research Foundation (EQ Project 20180230). The views expressed by the authors in the manuscript do not necessarily reflect the views of the EuroQol Group.

ZY, FP, AI, MJ, JB and NL are EuroQol members, other co-authors do not have conflicts of interest to declare. We thank Elly Stolk of the EuroQol Research Foundation, Kim Rand of the Akershus University Hospital for helping in data analysis. We also thank Ruixuan Jiang from Merck and Tianjiao Zeng from Guizhou Medical University for proofreading our manuscript.

ZY performed the analysis. ZY, FP prepared the results and drafted the manuscript. JB conceptualized the study and facilitated the data acquisition process. NL, JJ as corresponding authors, critically revised the manuscript and oversee the whole project. Other co-authors prepared the data in specific format and revised the manuscript. The data of each valuation study should be requested from the PI of each study.

References

1. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: Past, Present and Future. *Appl Health Econ Health Policy*. 2017;15(2):127-37.
2. Kennedy-Martin M, Slaap B, Herdman M, van Reenen M, Kennedy-Martin T, Greiner W, et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. *The European Journal of Health Economics*. 2020.
3. Knies S, Evers SM, Candel MJ, Severens JL, Ament AJ. Utilities of the EQ-5D: transferable or not? *Pharmacoeconomics*. 2009;27(9):767-79.
4. Xie F, Pullenayegum E, Pickard AS, Ramos Goni JM, Jo MW, Igarashi A. Transforming Latent Utilities to Health Utilities: East Does Not Meet West. *Health Econ*. 2017;26(12):1524-33.
5. Olsen JA, Lamu AN, Cairns J. In search of a common currency: A comparison of seven EQ-5D-5L value sets. *Health Econ*. 2018;27(1):39-49.
6. Norman R, Cronin P, Viney R, King M, Street D, Ratcliffe J. International comparisons in valuing EQ-5D health states: a review and analysis. *Value Health*. 2009;12(8):1194-200.
7. 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.
8. 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.
9. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goni JM. Overview, Update, and Lessons Learned From the International EQ-5D-5L Valuation Work: Version 2 of the EQ-5D-5L Valuation Protocol. *Value Health*. 2019;22(1):23-30.
10. Wang P, Liu GG, Jo MW, Purba FD, Yang Z, Gandhi M, et al. Valuation of EQ-5D-5L health states: a comparison of seven Asian populations. *Expert Rev Pharmacoecon Outcomes Res*. 2019;19(4):445-51.
11. Roudijk B, Donders ART, Stalmeier PFM. Cultural Values: Can They Explain Differences in Health Utilities between Countries? *Medical Decision Making*. 2019;39(5):605-16.
12. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L Value Set for China. *Value Health*. 2017;20(4):662-9.
13. 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.
14. Yang Z, van Busschbach J, Timman R, Janssen MF, Luo N. Logical inconsistencies in time trade-off valuation of EQ-5D-5L health states: Whose fault is it? *PLoS One*. 2017;12(9):e0184883.

15. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete Choice Experiments in Health Economics: Past, Present and Future. *Pharmacoeconomics*. 2019;37(2):201-26.
16. Krabbe PF, Devlin NJ, Stolk EA, Shah KK, Oppe M, van Hout B, et al. Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-state values. *Med Care*. 2014;52(11):935-43.
17. Purba FD, Hunfeld JAM, Timman R, Iskandarsyah A, Fitriana TS, Sadarjoen SS, et al. Test-Retest Reliability of EQ-5D-5L Valuation Techniques: The Composite Time Trade-Off and Discrete Choice Experiments. *Value Health*. 2018;21(10):1243-9.
18. Mott DJ, Shah KK, Ramos-Goni JM, Devlin NJ, Rivero-Arias O. Valuing EQ-5D-Y-3L Health States Using a Discrete Choice Experiment: Do Adult and Adolescent Preferences Differ? *Med Decis Making*. 2021;272989X21999607.
19. 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.
20. Purba FD, Hunfeld JAM, Iskandarsyah A, Fitriana TS, Sadarjoen SS, Ramos-Goni JM, et al. The Indonesian EQ-5D-5L Value Set. *Pharmacoeconomics*. 2017;35(11):1153-65.
21. Shirowa T, Ikeda S, Noto S, Igarashi A, Fukuda T, Saito S, et al. Comparison of Value Set Based on DCE and/or TTO Data: Scoring for EQ-5D-5L Health States in Japan. *Value Health*. 2016;19(5):648-54.
22. Shafie AA, Vasan Thakumar A, Lim CJ, Luo N, Rand-Hendriksen K, Md Yusof FA. EQ-5D-5L Valuation for the Malaysian Population. *Pharmacoeconomics*. 2019;37(5):715-25.
23. Pattanaphesaj J, Thavorncharoensap M, Ramos-Goni JM, Tongsiri S, Ingsrisawang L, Teerawattananon Y. The EQ-5D-5L Valuation study in Thailand. *Expert Rev Pharmacoecon Outcomes Res*. 2018;18(5):551-8.
24. Mai VQ, Sun S, Minh HV, Luo N, Giang KB, Lindholm L, et al. An EQ-5D-5L Value Set for Vietnam. *Qual Life Res*. 2020;29(7):1923-33.
25. Wong ELY, Ramos-Goni JM, Cheung AWL, Wong AYK, Rivero-Arias O. Assessing the Use of a Feedback Module to Model EQ-5D-5L Health States Values in Hong Kong. *Patient*. 2018;11(2):235-47.
26. Lin HW, Li CI, Lin FJ, Chang JY, Gau CS, Luo N, et al. Valuation of the EQ-5D-5L in Taiwan. *PLoS One*. 2018;13(12):e0209344.
27. Oppe M, Hout B. The 'power' of eliciting EQ-5D-5L values: the experimental design of the EQ-VT. *EuroQol Working Paper Series [Internet]*. 2017; (17003). Available from: <https://euroqol.org/wp-content/uploads/2016/10/EuroQol-Working-Paper-Series-Manuscript-17003-Mark-Oppe.pdf>.
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. Yang Z, Luo N, Oppe M, Bonsel G, Busschbach J, Stolk E. Toward a Smaller Design for EQ-5D-5L Valuation Studies. *Value in Health*. 2019;22(11):1295-302.
30. Hauber AB, Gonzalez JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. *Value Health*. 2016;19(4):300-15.
31. Rand-Hendriksen K, Augestad LA, Kristiansen IS, Stavem K. Comparison of hypothetical and experienced EQ-5D valuations: relative weights of the five dimensions. *Qual Life Res*. 2012;21(6):1005-12.
32. StataCorp. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LLC. 2015.
33. 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.
34. Ludwig K, Graf von der Schulenburg J-M, Greiner W. German Value Set for the EQ-5D-5L. *Pharmacoeconomics*. 2018.
35. Pickard AS, Law EH, Jiang R, Pullenayegum E, Shaw JW, Xie F, et al. United States Valuation of EQ-5D-5L Health States Using an International Protocol. *Value Health*.
36. Golicki D, Jakubczyk M, Graczyk K, Niewada M. Valuation of EQ-5D-5L Health States in Poland: the First EQ-VT-Based Study in Central and Eastern Europe. *Pharmacoeconomics*. 2019;37(9):1165-76.
37. Welie AG, Gebretekle GB, Stolk E, Mukuria C, Krahn MD, Enquoselassie F, et al. Valuing Health State: An EQ-5D-5L Value Set for Ethiopians. *Value in Health Regional Issues*. 2020;22:7-14.
38. Al Shabasy SA, Abbassi MM, Finch AP, Baines D, Farid SF. The EQ-5D-5L Valuation Study in Egypt. *Pharmacoeconomics*. 2021;39(5):549-61.

39. Craig BM, Reeve BB, Cella D, Hays RD, Pickard AS, Revicki DA. Demographic differences in health preferences in the United States. *Med Care*. 2014;52(4):307-13.
40. Sayah FA, Bansback N, Bryan S, Ohinmaa A, Poissant L, Pullenayegum E, et al. Determinants of time trade-off valuations for EQ-5D-5L health states: data from the Canadian EQ-5D-5L valuation study. *Qual Life Res*. 2016;25(7):1679-85.
41. Luo N, Wang Y, How CH, Tay EG, Thumboo J, Herdman M. Interpretation and use of the 5-level EQ-5D response labels varied with survey language among Asians in Singapore. *J Clin Epidemiol*. 2015;68(10):1195-204.
42. Greiner W, Weijnen T, Nieuwenhuizen M, Oppe S, Badia X, Busschbach J, et al. A single European currency for EQ-5D health states. Results from a six-country study. *Eur J Health Econ*. 2003;4(3):222-31.

Tables

Table 1. Basic information of the 11 studies

Region	Sample Size*	Sample age range	Sampling strategy	Sampling determinants	Language(s)**	Data collection time
China	1,300	18-85	Quota	Metropolitan areas, age, sex, and educational attainment	Mandarin	2012
Indonesia	1,056	17-75	Quota	Urbanicity, sex, age, educational attainment, religion, and ethnicity	Bahasa Indonesia	2015-2016
Japan	1,262	20-69	Stratified	Region, sex, and age	Japanese	2014
South Korea	1,080	19-87	Quota	Region, sex, age, and educational attainment	Korean	2013
Malaysia	1,143	18-88	Quota	Region, urbanicity, sex, age, and ethnicity	Malay and English	2016
Singapore	1,849	10-99	Quota	Age, gender and ethnicity	English and Mandarin	2014-2015
Thailand	1,209	18-84	Quota	Region, age and sex	Thai	2013-2014
Philippines	1,000	18-80	Quota	Age, sex, region, urbanicity, educational attainment, income, and ethno-linguistic	Tagalog, Cebuano, and English	2017
Vietnam	1,201	18-88	Quota	Region, urbanicity, age, and sex	Vietnamese	2017
Hong Kong	1,014	18-88	Quota	Sex, age, and educational attainment	Cantonese	2014-2015
Taiwan	1,000	20-90	Quota	Region, age, sex, and educational attainment	Traditional Chinese	2017

*The final sample size for estimating the value set may differ with the sample size shown above.

** The language(s) being used in each study is the official language and/or dialects that are spoken mostly by the population

Table 2. Modelling results of the conditional logit model (coefficient, standard error)											
	China	Indonesia	Japan	South Korea	Malaysia	Singapore	Thailand	Philippines	Vietnam	Hong Kong	Taiwan
mo2	-0.436, 0.051	-0.514, 0.056	-0.509, 0.052	-0.571, 0.055	-0.503, 0.056	-0.413, 0.041	-0.432, 0.054	-0.602, 0.058	-0.589, 0.055	-0.562, 0.059	-0.579, 0.061
mo3	-0.727, 0.060	-0.812, 0.066	-0.605, 0.058	-0.690, 0.063	-0.757, 0.067	-0.549, 0.048	-0.441, 0.063	-0.516 , 0.067	-0.617, 0.064	-0.859, 0.069	-0.867, 0.073
mo4	-1.247, 0.060	-1.539, 0.068	-1.092, 0.061	-1.180, 0.064	-1.542, 0.069	-1.149, 0.049	-1.204, 0.064	-1.309, 0.068	-1.411, 0.066	-1.621, 0.071	-1.600, 0.074
mo5	-1.845, 0.068	-2.348, 0.079	-1.575, 0.065	-2.102, 0.075	-2.221, 0.079	-1.648, 0.054	-2.364, 0.078	-1.788, 0.076	-2.438, 0.079	-2.284, 0.082	-2.268, 0.084
sc2	-0.155, 0.055	-0.376, 0.061	-0.278, 0.056	-0.101 , 0.058	-0.389, 0.062	-0.216, 0.044	-0.348, 0.059	-0.448, 0.062	-0.318, 0.060	-0.264, 0.063	-0.239, 0.067
sc3	-0.382, 0.060	-0.515, 0.067	-0.368, 0.060	-0.145, 0.062	-0.478, 0.067	-0.373, 0.048	-0.332 , 0.064	-0.483, 0.067	-0.294 , 0.066	-0.370, 0.067	-0.410, 0.072
sc4	-0.870, 0.061	-0.831, 0.067	-0.691, 0.061	-0.401, 0.063	-1.328, 0.069	-0.874, 0.049	-1.038, 0.065	-1.175, 0.068	-0.919, 0.067	-1.095, 0.070	-1.103, 0.073
sc5	-1.185, 0.059	-1.190, 0.064	-0.986, 0.058	-0.729, 0.060	-1.566, 0.068	-1.101, 0.047	-1.404, 0.064	-1.323, 0.065	-1.412, 0.065	-1.352, 0.068	-1.426, 0.072
ua2	-0.313, 0.053	-0.341, 0.058	-0.335, 0.053	-0.228, 0.055	-0.286, 0.059	-0.202, 0.042	-0.272, 0.056	-0.398, 0.059	-0.268, 0.057	-0.246, 0.060	-0.137, 0.063
ua3	-0.376, 0.059	-0.530, 0.066	-0.430, 0.059	-0.253, 0.061	-0.325, 0.066	-0.241, 0.047	-0.312, 0.062	-0.394 , 0.065	-0.326, 0.062	-0.357, 0.067	-0.250, 0.070
ua4	-0.796, 0.059	-1.102, 0.066	-0.855, 0.060	-0.657, 0.063	-0.931, 0.068	-0.587, 0.047	-0.848, 0.064	-0.951, 0.066	-1.073, 0.066	-0.884, 0.067	-1.025, 0.073
ua5	-1.320, 0.061	-1.439, 0.067	-1.164, 0.060	-1.161, 0.064	-1.194, 0.069	-0.749, 0.047	-1.461, 0.067	-1.157, 0.067	-1.813, 0.070	-1.120, 0.070	-1.462, 0.076
pd2	-0.426, 0.055	-0.301, 0.061	-0.294, 0.056	-0.288, 0.059	-0.508, 0.062	-0.234, 0.044	-0.243, 0.059	-0.494, 0.062	-0.653, 0.061	-0.374, 0.063	-0.401, 0.066
pd3	-0.621, 0.059	-0.343, 0.064	-0.436, 0.060	-0.435, 0.063	-0.570, 0.068	-0.246, 0.047	-0.286, 0.064	-0.524, 0.066	-1.005, 0.067	-0.587, 0.068	-0.547, 0.072
pd4	-1.366, 0.063	-0.660, 0.065	-0.848, 0.060	-1.094, 0.066	-1.478, 0.071	-0.901, 0.048	-1.106, 0.065	-1.280, 0.069	-1.624, 0.071	-1.268, 0.070	-1.408, 0.076
pd5	-1.634, 0.063	-0.856, 0.063	-1.380, 0.061	-1.528, 0.067	-2.083, 0.076	-1.338, 0.049	-1.405, 0.066	-1.419, 0.068	-2.186, 0.073	-1.429, 0.070	-2.157, 0.082
ad2	-0.204, 0.057	-0.185, 0.062	-0.250, 0.057	-0.124, 0.060	-0.397, 0.065	-0.207, 0.046	-0.281, 0.062	-0.241, 0.063	-0.309, 0.062	-0.307, 0.064	-0.149, 0.069
ad3	-0.570, 0.058	-0.410, 0.063	-0.513, 0.058	-0.325, 0.062	-0.586, 0.066	-0.349, 0.046	-0.487, 0.063	-0.387, 0.065	-0.641, 0.064	-0.504, 0.067	-0.621, 0.071
ad4	-1.171, 0.063	-0.831, 0.065	-0.942, 0.061	-0.702, 0.064	-1.352, 0.073	-0.917, 0.049	-1.361, 0.069	-0.884, 0.069	-1.046, 0.069	-1.154, 0.071	-1.325, 0.079
ad5	-1.478, 0.064	-1.145, 0.069	-1.391, 0.064	-0.903, 0.065	-1.863, 0.074	-1.188, 0.050	-1.917, 0.074	-0.892, 0.068	-1.385, 0.069	-1.347, 0.071	-1.749, 0.080
Number of observations	18,200	14,756	16,838	15,120	16,002	25,886	16,898	14,000	16,814	14,196	14,000
Log-likelihood	-4723.24	-3958.62	-4659.15	-4148.61	-3787.40	-7195.06	-4114.02	-3796.76	-4017.79	-3652.12	-3285.30
AIC	9486.48	7957.24	9358.30	8337.22	7614.79	14430.12	8268.05	7633.53	8075.57	7344.23	6610.59
BIC	9642.66	8109.23	9512.92	8489.69	7768.40	14593.35	8422.75	7784.46	8230.17	7495.45	6761.53

* Italic and bold font suggests the coefficient is not significant at 0.05 level; bold font suggests the coefficient is inconsistent.

Table 3. Number of coefficients differed statistically between two regions										
	China	Indonesia	Japan	Korea	Malaysia	Singapore	Thailand	Philippines	Vietnam	Hong Kong
Indonesia	9									
Japan	4	8								
Korea	9	11	13							
Malaysia	8	11	10	13						
Singapore	11	8	11	10	16					
Thailand	10	11	10	9	8	7				
Philippine	7	10	8	8	9	11	7			
Vietnam	10	12	10	15	5	12	10	9		
Hong Kong	5	8	7	9	5	11	6	6	8	
Taiwan	8	10	12	11	3	15	5	9	7	3

Figures

Which is better, state A or state B?

- severe problems in walking about
- no problems washing or dressing myself
- moderate problems doing my usual activities
- slight pain or discomfort
- not anxious or depressed

A

- slight problems in walking about
- no problems washing or dressing myself
- moderate problems doing my usual activities
- extreme pain or discomfort
- severely anxious or depressed

B

Figure 1

An example of Discrete Choice Experiments (DCE) in English

Page 10/11

Figure 2

Relative importance of five dimensions

Figure 3

Relative importance of the severity levels

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendixrevisionv3.docx](#)