การเปรียบเทียบ 5 วิธีในการวัดค่าอรรถประโยชน์ของคุณภาพชีวิตด้านสุขภาพของนักศึกษาปริญญาตรี
A Comparison of Five Approaches for Measuring Utility Values of Health-related Quality of Life among Undergraduate Students

นิพนธ์ต้นฉบับ

Original Article

ชัยพัฒน์ ฉัตรศิรินพคุณ[1], ภัทร ลาภเลิศลอย[1], ณัฏฐธิดา ปัญญาศรี[1] และ กฤตภาส กังวานรัตนกุล[2*]

[1] นักศึกษาเภสัชศาสตรบัณฑิต สาขาบริบาลเภสัชกรรม คณะเภสัชศาสตร์ มหาวิทยาลัยบูรพา จ.ชลบุรี 20131
[2] สาขาวิชาเภสัชกรรมสังคมและบริหารเภสัชกิจ คณะเภสัชศาสตร์ มหาวิทยาลัยบูรพา จ.ชลบุรี 20131

*** Corresponding author:** Krittaphas@buu.ac.th

วารสารไทยเภสัชศาสตร์และวิทยาการสุขภาพ2566;18(3):272-279.

Chaiphat Chatsirinoppakhun[1], Pattara Laplerdloy[1], Nutthathida Panyasri[1] and Krittaphas Kangwanrattanakul[2*]

[1] Bachelor's degree students, Faculty of Pharmaceutical Sciences, Burapha University, Chonburi, Thailand
[2] Division of Social and Administrative Pharmacy, Faculty of Pharmaceutical Sciences, Burapha University, Chonburi, 20131, Thailand

*** Corresponding author:** Krittaphas@buu.ac.th

Thai Pharmaceutical and Health Science Journal 2022;18(3):272-279.

บทคัดย่อ

วัตถุประสงค์: เพื่อทดสอบคุณสมบัติและความสอดคล้องกันของคะแนนอรรถประโยชน์จาก 5 วิธี คือ EQ-5D-3L, EQ-5D-5L (cTTO model), EQ-5D-5L (DCE model), EQ-5D-5L (Hybrid model) และ VAS ของนักศึกษาระดับปริญญาตรี **วิธีการศึกษา:** การศึกษาเชิงสำรวจแบบภาพตัดขวางเก็บข้อมูลระหว่างมีนาคมถึงเมษายน 2565 กับกลุ่มตัวอย่าง 393 คน ใช้สถิติสัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient; ICC) เพื่อทดสอบความสอดคล้องของค่าอรรถประโยชน์ 5 วิธี การวิเคราะห์การถดถอยเชิงเส้นพหุคูณเพื่อเปรียบเทียบค่าอรรถประโยชน์ตามกลุ่มย่อยตามลักษณะทางประชากร ได้แก่ เพศ อายุ โรคประจำตัว ประวัติการสูบบุหรี่ และประวัติการดื่มเหล้าหรือเครื่องดื่มผสมแอลกอฮอล์ **ผลการศึกษา:** ค่าอรรถประโยชน์เฉลี่ย (SD) จากน้อยไปหามาก คือ 0.79 ± 0.13 (VAS), 0.84 ± 0.18 (EQ-5D-3L), 0.92 ± 0.11 (Hybrid model) 0.93 ± 0.12 (cTTO model) และ 0.94 ± 0.10 (DCE model) ค่า ICC แสดงคู่วิธีที่สอดคล้องกันในระดับยอดเยี่ยม ได้แก่ cTTO model-Hybrid model, cTTO model-DCE model, DCE model-Hybrid model แต่พบว่า VAS และ EQ-5D-3L และ EQ-5D-5L มีความสอดคล้องระดับแย่ พบว่าเพศหญิงและผู้สูบบุหรี่มีค่าอรรถประโยชน์ EQ-5D-3L, cTTO model, DCE model และ Hybrid model ต่ำกว่าเพศชายและผู้ที่ไม่สูบบุหรี่ (P-value < 0.05) **สรุป:** EQ-5D-5L (DCE model) เป็นวิธีที่ดีที่สุดในการหาประเมินอรรถประโยชน์สำหรับนักศึกษาปริญญาตรีเพราะสามารถจำแนกค่าอรรถประโยชน์ได้ระหว่างกลุ่มย่อย

คำสำคัญ: คุณภาพชีวิตทางด้านสุขภาพ, ค่าอรรถประโยชน์, แบบสอบถามอีคิวไฟว์ดีทรีแอล, แบบสอบถามอีคิวไฟว์ดีไฟว์แอล, ความสอดคล้องกัน

Editorial note
Manuscript received: Sep 17, 2022;
Notified: Dec 11, 2022;
Revised: Dec 14, 2022;
Accepted in final form: Jan 23, 2023;
Published online: September 30, 2023.

Abstract

Objective: To investigate the performance and agreement of utility scores elicited from various elicitation methods of EQ-5D-3L, the three value sets of the EQ-5D-5L (cTTO model, DCE model, and Hybrid model), and VAS among undergraduate students. **Methods:** A cross-sectional survey study was conducted with 393 undergraduate students between March and April 2022. Intraclass correlation (ICC) was used to determine the agreement of utility values derived from five approaches. Multiple regression was used to compare the utility values with differences in gender, age, smoking status, alcohol consumption and medical conditions. **Results:** The mean (SD) utility values derived from five approaches were as follows: 0.79 ± 0.13 (VAS), 0.84 ± 0.18 (EQ-5D-3L), 0.92 ± 0.11 (Hybrid model) 0.93 ± 0.12 (cTTO model), and 0.94 ± 0.10 (DCE model). The ICC showed excellent agreement among these following pairs: cTTO model-Hybrid model, cTTO model-DCE model, DCE model-Hybrid model. However, the agreement of utility values from VAS and EQ-5D-3L and EQ-5D-5L was poor. Females and smokers reported lower the utility values from EQ-5D-3L, cTTO model, DCE model, and the Hybrid model than their counterparts (P-value < 0.05). **Conclusion:** The EQ-5D-5L (DCE model) is the best elicitation method among undergraduate students because it can discriminate utility scores between predefined subgroups.

Keywords: Health-related quality of life, Utility scores, EQ-5D-3L, EQ-5D-5L, agreement

Journal website: http://ejournals.swu.ac.th/index.php/pharm/index

Introduction

Health related quality of life (HRQoL) is known as a humanistic outcome that can be directly measured from an individual's perspective to determine the specific effects of health, illness and medical treatment on an individual's quality of life.[1,2] Furthermore, it can be considered an essential indicator for decision-making to determine whether a novel intervention is worthwhile for resource allocation compared to an existing or standard intervention.[3,4]

Cost-utility analysis (CUA) is an economic analysis approach to compare between cost and outcome of health care interventions recommended by several health technology assessment (HTA) guidelines including Thai guideline.[5] For CUA, the outcome is measured in terms of quality adjusted life years wherein it can be computed by multiplying the amount of life year expectancy and utility scores to indicate an individual's preference towards a given health state.[6,7]

The utility score generally ranges from 0 (the worst possible health state or death) to 1 (the perfect health).[8,9] Two major approaches used to elicit the utility score include the direct method such as standard gamble (SG), time-trade off (TTO), and visual analogue scale (VAS), and the indirect method where a multidimensional questionnaire is used to measure the utility score such as the Short Form-6-dimension (SF-6D), Health Utility Index (HUI), the 15D, the Quality of Wellbeing (QWB), EQ-5D, etc.[10]

Given its simplicity and self-completion with low burden to respondents, the EQ-5D, developed by the EuroQoL group in the 1980s, is designed to measure the utility score in both the general population and clinical area, and it is commonly used to elicit the utility score for economic analyses strongly recommended by several HTA guidelines.[11-14] When EQ-5D is not appropriate to elicit the utility score for some conditions, other elicitation methods, SG, TTO, VAS, or Short Form-6 dimension (SF-6D), can be employed with justification.

Currently, the EQ-5D has two versions including 1) EQ-5D-3L and 2) EQ-5D-5L. However, there is no consensus guideline to suggest which EQ-5D version should be employed to elicit the utility score, especially for some countries that both versions of the EQ-5D are available because the use of different versions of the EQ-5D can yeild different utility scores which may affect the results of incremental cost-effectiveness ratio in ecnomic analyses.[15,16]

In Thailand, the EQ-5D-5L is also recommended for economic analyses by the Thai HTA guideline.[5] Currently, there are value sets for eliciting utility score for both EQ-5D-3L[17], and EQ-5D-5L for general Thai population.[18] Furthermore, the Thai EQ-5D-5L has different value sets used to elicit the utility scores as follows: Composite time trade-off (cTTO) model, Discrete choice experiment (DCE) model, and Hybrid model. However, evidence is limited for the agreement between the utility scores derived from EQ-5D-3L and the three value sets of the EQ-5D-5L (cTTO model, DCE model, and Hybrid model).

To date, the study was conducted to estimate the agreement of utility score derived from the EQ-5D-3L, the three value sets of the EQ-5D-5L (cTTO model, DCE model, and Hybrid model), TTO, and VAS in cervical cancer patients[19], and the results showed that TTO, VAS, EQ-5D-3L and EQ-5D-5L could not be used interchangeably and the use of different instruments might affect the results of economic analyses. Although previous evidence supported the use of Hybrid model for eliciting the utility scores in the Thai population[18], evidence is still limited for the measurement properties of the utility values derived from the three value sets of the EQ-5D-5L and other elicitation methods including the EQ-5D-3L and VAS including the agreement and known-groups validity among the undergraduate students. Regarding the measurement properties, previous studies have shown that the utility scores could be associated with some demographic factors including gender, smoking, and alcohol consumption in that female, smokers and drinkers were more likely to have lower utility scores than their counterparts[20-22], so these variables should be selected to investigate the known-group analysis for the utility scores elicited from five approaches. Therefore, this study aimed to investigate the performance and agreement of utility scores elicited from various elicitation methods as follows: EQ-5D-3L, the three value sets of the EQ-5D-5L (cTTO model, DCE model, and Hybrid model), and VAS among the undergraduate students.

Methods

Sample and settings

A cross-sectional survey study with a face-to-face interview was conducted with 393 undergraduate students from 19 faculties at Burapha University, Bangsean campus (Chonburi Thailand). Participants were eligible if they were 1) undergraduate students in the academic year of 2021, 2) willing to participate into this study, and 3) able to understand and read Thai language. Those who could not complete the questionnaire and those were aged less than 18 years old were excluded. To ensure fair representation of the study population, our participants were recruited using a quota-sampling method in the proportion to the study population of each faculty, and sample size was computed using Jacob Cohen[23] because it was conducted to compare the utility values from the five approaches. Considering 0.05 significant level (α), 0.80 statistical power (1-β), and 0.20 effect size, the sample size according to the Jacob Cohen table was 61 participants per group. As a result, a sample size of 05 participants was required. However, the sample size was adjusted for unintentional error/missing responses using the following formula, $n_1 = n/(1-d)$[24], where n = 305 and d = 20%. Therefore, a sample of 381.25 particpants was required. However, we collected the data with 393 complete responses (participants) in our study.

Data collection procedure

All interviews were conducted between March and April 2022. Prior to the data collection process, a participant information sheet explaining the study objectives and overall data collection process in plain language was given to an individual undergraduate student. Then, a written inform consent was obtained from the participants in case they consented to be interviewed. Each sample was asked to complete a three-part questionnaire as follows: 1) general information (gender, age, year of study, faculty, average family income, underlying diseases, smoking and alcohol history), 2) EQ-5D-3L, 3) EQ-5D-5L, and 4) EQ-VAS. Furthermore, both EQ-5D-3L and EQ-5D-5L were permitted to use for conducting this study by the EuroQoL group (Registration ID: 43532), and the study also obtained ethical approval from the Burapha University Institutional Review Board (IRB1-027/2565).

Research instruments

EQ-5D-3L

The EQ-5D-3L is considered a generic health preference-based instrument widely used to elicit the utility scores for the general Thai population. It consists of five items, one for each of the following dimensions: mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD). Each item has three response options including no problems, some/moderate problems, and extreme problems of inability or confined to a bed. The descriptive system is used to compute the utility scores using the value sets developed for the Thai population.[17]

EQ-5D-5L

The EQ-5D-5L is a newer version of the EQ-5D which consists of five items for each health dimension, and previous evidence has suggested that it has better measurement properties than those of the EQ-5D-3L in the Thai population.[22,25] However, each dimension has five response options namely no problems, slight problems, moderate problems, severe problem, extreme problems/unable to perform. Similar to the EQ-5D-3L, the descriptive system is also employed to calculate the utility score using the value sets developed for the Thai population.[18] In this study, the utility scores were derived from three following value sets: Composite time trade-off (cTTO) model, the discrete choice experiment (DCE) model, and the hybrid model based on a standardized valuation study protocol (EQ-VT) developed by the EuroQol Group.[26]

EQ-VAS

The VAS is considered a direct elicitation method, where the sample is asked to rate their current health status on the day the interview was performed by placing a cross-mark on a 20-centimeter vertical line, which is anchored at 0 labeled as “worst imaginable health state,” and 100 labeled as “best imaginable health state.” The utility scores are derived from the number marked on the VAS divided by 100.

Data analysis

Regarding the sample characteristics, descriptive statistics with frequency, proportions, mean and standard deviations (SD) were used as appropriate. Mean, SD, median, interquartile ranges (IQR), and range were also employed to report the utility scores from each approach and the utility scores obtained from various approaches were compared using the Friedman test and Wilcoxon signed rank test.

The agreement levels of utility scores derived from various approaches were computed and compared using the intraclass correlation coefficients (ICCs) with two-way mixed-effects model and absolute agreement and average measures, yielding the ICCs within a range from 0.00 to 1.00. According to the guideline for reporting ICCs, the computed ICCs were classified into four levels of agreements as poor agreement (ICCs $<$ 0.50), moderate agreement ($0.50 \leq$ ICCs $<$ 0.75), good agreement ($0.75 \leq$ ICCs $<$ 0.90), and excellent agreement (ICCs $\geq$ 0.90).[27]

The performance of each elicitation approach was evaluated using ceiling/floor effects and known-group validity. The ceiling and floor effects were reported and computed as proportions of respondents with the highest and lowest utility scores derived from each elicitation method. For the known-group validity, it was performed to determine whether the mean utility score changes appeared to be against participant sub-groups classified by participants’ characteristics. We expected that the utility scores would be lower among women, youngsters, self-reported smokers, self-reported drinkers, and participants without underlying disease than their counterparts.[22,28]

A multivariable Tobit regression model was employed to investigate the association between participant’s characteristics including gender, age, smoking status, alcohol

consumptions and self-reported comorbidities as the independent variables, and EQ-5D utility scores derived from the EQ-5D-3L, the three value sets of the EQ-5D-5L (cTTO model, DCE model, and Hybrid model), and VAS as the dependent variables. Given the fact that the utility scores usually show a severe ceiling effect in which most of the participants usually rate themselves as full health, with the utility scores of 1.00. Therefore, the data could be interpreted as being bounded or censored at 1.0 which could contribute to biased coefficient estimations using ordinary least square regression.[29] The multivariable Tobit regression model was recommended as a more appropriate alternative method to deal with the censored and skewed data in econometrics research.[29] Two goodness of fit indices, Alkaline Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used for determination of the model. The Tobit regression model contributing lowest AIC and BIC values was the best regression model to predict the EQ-5D utility and the VAS scores.

All statistical analyses were performed using IBM SPSS version 23 except the multivariable Tobit regression model was performed using STATA version 17.0. Statistical significance was set a type I error of 5%.

Results

Of the 393 participants, most were female (72%) and first year students (32.8%). The average age of the participants was 20.07±1.19 years old. The majority of the participants were non-smokers (95.2%) and drinkers (55.5%). Furthermore, most of them reported themselves as healthy (90.3%) (Table 1).

Descriptive statistics of utility scores

It was found that the mean (SD) utility scores from the five approaches ranging from low to high were 0.79±0.13 (VAS), 0.84±0.18 (EQ-5D-3L), 0.92±0.11 (EQ-5D-5L Hybrid model), 0.93±0.12 (EQ-5D-5L cTTO model), and 0.94±0.10 (EQ-5D-5L DCE model) (Table 2). It should be noted that there were statistically significant for the utility scores derived from five approaches (P-value < 0.01). Furthermore, the Wilcoxon signed rank test revealed that all pairwise comparisons of utility scores were statistically significant (P-value < 0.01). Notably, the ceiling effects of the VAS and EQ-5D-3L were 4.8% and 49.9%, while those of the EQ-5D-5L for cTTO, DCE and Hybrid models were equal, at 46.3%. Conversely, no floor effects were observed across the five approaches (Table 2).

Table 1 Demographic characteristics (N = 393).

Characteristics	N (%)
Sex	
Male	110 (28.0)
Female	640 (72.0)
Age	
Mean ± SD	20.07 ± 1.19
Field of study	
Health sciences	85 (21.6)
Pure sciences and technology	131 (33.3)
Social sciences and humanities	177 (45.0)
Year of study	
First year	129 (32.8)
Second year	100 (25.4)
Third year	125 (31.8)
Fourth year	26 (6.6)
Fifth year	13 (3.3)
Self-reported smoking	
Non-smokers	374 (95.2)
Smokers	19 (4.8)
Self-reported alcohol consumption	
Non-smokers	175 (44.5)
Smokers	218 (55.5)
Self-reported underlying diseases	
Yes	38 (9.7)
No	355 (90.3)

Table 2 Descriptive statistics of utility scores from 5 methods (N = 393).

Elicitation methods	Mean	SD	Median	Interquartile range	Range	% Floor	% Ceiling
EQ-VAS	0.79	0.13	0.80	0.70 - 0.90	0.25 - 1.00	0	5.1
EQ-5D-3L	0.84	0.18	0.77	0.71 - 1.00	0.04 - 1.00	0	49.9
EQ-5D-5L (cTTO model)	0.93	0.12	0.96	0.89 - 1.00	0.21 - 1.00	0	46.3
EQ-5D-5L (DCE model)	0.94	0.10	0.96	0.92 - 1.00	0.37 - 1.00	0	46.3
EQ-5D-5L (Hybrid model)	0.92	0.11	0.94	0.89 - 1.00	0.27 - 1.00	0	46.3

Agreement of the utility scores

The ICCs were used to determine the agreement level of utility scores across five approaches (Table 3). It showed that poor agreement (ICCs < 0.05) was found in the pairs of the VAS-cTTO model, the VAS-DCE model, and the VAS-Hybrid model. Moderate agreement (0.50 ≤ ICCs < 0.75) was found in these three pairs, EQ-5D-3L-VAS, EQ-5D-3L-cTTO model, and EQ-5D-3L-DCE. Moreover, good agreement (0.75 ≤ ICCs < 0.90) was found only in one pair of EQ-5D-3L-Hybrid model, while an excellent agreement (ICCs ≥ 0.90) was observed among utility scores from three EQ-5D-5L models including the cTTO model-DCE model, the cTTO model-Hybrid model, and the DCE model-Hybrid model (Table 3).

Table 3 Agreement of utility scores from various elicitation approaches using intraclass correlation coefficients (ICCs) (N = 393).

Approaches	Intraclass correlation coefficient (95% Confidence interval)				
	EQ-5D-3L	VAS	EQ-5D-5L (cTTO model)	EQ-5D-5L (DCE model)	EQ-5D-5L (Hybrid model)
EQ-5D-3L	-	0.565 (0.458 - 0.650)	0.753 (0.369 - 0.875)	0.685 (0.208 - 0.841)	0.770 (0.398 - 0.884)
EQ-VAS	-	-	0.436 (-0.095 - 0.683)	0.387 (-0.161 - 0.661)	0.444 (-0.090 - 0.689)
EQ-5D-5L (cTTO model)	-	-	-	0.960 (0.936 - 0.973)	0.990 (0.987 - 0.992)
EQ-5D-5L (DCE model)	-	-	-	-	0.970 (0.890 - 0.987)
EQ-5D-5L (Hybrid model)	-	-	-	-	-

Known-group validity

It was found that the five elicitation approaches could discriminate utility scores in regard to the defined sample characteristics (Table 4). The assumptions for gender and smoking status were confirmed for the utility scores derived from EQ-5D-3L and three EQ-5D-5L models because female and smokers had lower utility scores than their counterparts (P-value < 0.05), while it showed that only non-drinkers had lower utility scores derived from the VAS than those of drinkers (P-value < 0.05). Moreover, we found that the EQ-5D-5L (DCE model) yielded the lowest AIC and BIC values (Table 4).

Table 4 Known-group validity of the utility scores from various elicitation approaches using multivariable Tobit regression model (N = 393).

Participant characteristics	Coefficients (95% Confidence interval)				
	EQ-5D-3L	EQ-5D-5L (cTTO model)	EQ-5D-5L (DCE model)	EQ-5D-5L (Hybrid model)	EQ-VAS
Sex (Ref: Male)					
Female	-0.095* (-0.178, -0.013)	-0.059* (-0.107, -0.011)	-0.045* (-0.084, -0.006)	-0.056* (-0.103, -0.009)	-1.427 (-4.569, 1.715)
Age	-0.016 (-0.045, 0.013)	-0.003 (-0.020, 0.014)	-0.002 (-0.016, 0.012)	-0.003 (-0.019, 0.014)	-0.525 (-1.660, 0.609)
Smoking (Ref: Non-smokers)					
Smokers	-0.130* (-0.238, -0.022)	-0.072* (-0.135, -0.008)	-0.057* (-0.109, -0.005)	-0.069* (-0.131, -0.008)	-1.749 (-6.062, 2.564)
Alcohol (Ref: Non-drinkers)					
Drinkers	0.001 (-0.074, 0.075)	0.017 (-0.026, 0.061)	0.013 (-0.023, 0.049)	0.013 (-0.029, 0.056)	-2.567 (-5.455, 0.322)
Reported comorbidities (Ref: None)					
Yes	0.060 (-0.058, 0.178)	0.017 (0.051, 0.084)	0.014 (-0.041, 0.070)	0.016 (-0.050, 0.081)	0.444 (-4.112, 5.000)
AIC	403.442	149.281	67.903	144.142	3059.397
BIC	431.258	177.098	95.719	171.959	3087.214

* P-value < 0.05.

AIC: Alkaike Information Criterion, BIC: Bayesian Information Criterion.

Discussions and Conclusion

Ours is the first study examining the utility scores derived from the five elicitation approaches covering both direct (VAS) and indirect methods (EQ-5D-3L and EQ-5D-5L). Notably, the three EQ-5D-5L value sets including the cTTO model, DCE model, and Hybrid model, were also employed into this comparison. Due to the limited budget, this study was nevertheless conducted to investigate those aspects among undergraduate students and these findings are expected to be the preliminary results for future studies investigating those aspects in the general Thai population.

We found that the EQ-5D-3L yeileded smaller utility scores than those from the EQ-5D-5L. This finding is consistent with the previous studies in which the utility scores from the EQ-5D-3L were significantly less than those from the EQ-5D-5L in a wide range of population including general population and clinical area.[19, 30-32] We explained that adding two more levels of severity for the EQ-5D-5L produces higher discriminatory power than those of the EQ-5D-3L which is consistent with the findings from the previous studies conducted in both general population and clinical area[22, 25, 33, 34]; therefore, the three EQ-5D-5L models can yeild higher utility scores than those of the EQ-5D-3L.

Like several previous studies[19, 22, 35], our study also showed that by adding two more levels of severity to the EQ-5D-3L could reduce the ceiling effect of the EQ-5D-5L since the ceiling effects of the EQ-5D-5L and the EQ-5D-3L were 46.3% and 49.9%, respectively. However, no floor effects were not detected because a majority of the recruited participants (90.3%) reported themseleves as healthy.

Our study also revealed that the utility scores derived from the VAS (direct method) were lower than those obtained from both EQ-5D-3L and EQ-5D-5L (indirect method). Similar to the previous study conducted with locally advanced cervical cancer patients[19], it showed that the utility scores obtained from the VAS and TTO were lower than those of the EQ-5D-5L questionnaire. Therefore, the utility score from the direct elicitation method may be lower than those from the indirect method in the Thai population. Similar to the previous study[19], our study also revealed that the EQ-5D-5L DCE model yeilded the highest utility scores compared to the other two EQ-5D-5L models because the EQ-5D-5L DCE model has the smallest coefficients compared to other two elicitation methods especially for the dimensions of PD and AD for all five severity

levels[18]. Our participants also reported themselves as having health problems especially for the severe and extreme problems for the dimensions of PD and AD, resulting in the highest utility scores obtained from the EQ-5D-5L DCE model.

As expected, excellent agreement (ICCs ≥ 0.90) was observed among the set of EQ-5D-5L approaches because the three EQ-5D-5L models were designed by the EQ-VT protocol for EQ-5D-5L valuation study among the general Thai population.[18] Moreover, the Hybrid model is generated by combining the EQ-5D-5L cTTO model and EQ-5D-5L DCE model, resuling in excellent agreement among the three EQ-5D-5L models. Similar to the previous study[19], poor to moderate agreements between the direct method (VAS) and indirect method (EQ-5D-5L and EQ-5D-3L) were detected. Again, it may be due to different valuation methods might account for this phenomenon in that the indirect method requires the respondents to rate their current health status through the predefined health dimension in the questionnarie, while the direct method requires the respondents to rate their health status into one value by considering all health aspects which may influence on their current health status.

Similar to previous studies [20, 22], both EQ-5D-3L and EQ-5D-5L showed that significant decreases in the uility scores were observed among females and smokers implying that both EQ-5D-3L and EQ-5D-5L could discriminate the utility scores for gender and smoking status among undergraduate students. Therefore, the interventions are urgently needed to tackle the respondents with some sociodemographic factors such as female and smokers, which independently or in combination associated with utillity loss among undergraduate students. Similar to known-group results from the previous study[22], the coefficients of EQ-5D-3L were slightly higher than those of the EQ-5D-5L for smoking status and gender variables, which may imply that the magnitude of decreases in utility scores derived from the EQ-5D-3L were higher than that of the EQ-5D-5L. We reasonsed that the EQ-5D-3L has three response options contributing to lower utility scores when the respondents rated themselves as having health impairments compared to the EQ-5D-5L.

Nonetheless, our study did not show the significant decreases in utility scores obtained from five elicitation approaches between respondents with/without health comorbidies. It may be due to the fact that most of the respondents reported themselves as healthy, and only a few participants (9.7%) indicated themselves with some health comorbidities. Notably, most of the respodents with health comorbidities reported themselves as having allergic rhinitis (54%) which is an acute disease and might not be active during the interviews. Moreover, known-group validity showed that age was not associated with the utility scores from all five approaches because the age range of undergraduate students was quite narrow (18 and 24 years old), and most reported themselves as healthy. Therefore, the differences in utility scores were not observed among this age range of the participants.

An unexpected association of the utillity scores derived from both EQ-5D-5L and EQ-5D-3L was observed for the drinkers/non-drinkers group because it showed that the utility scores derived from both EQ-5D-5L and EQ-5D-3L for non-drinker were lower than those of drinkers. It seems to be a specific characteristic of the general Thai population because a previous study also showed that the utility scores from both EQ-5D-3L and EQ-5D-5L for drinkers were higher than those of non-drinkers[22, 36]. Nevertheless, the EQ-VAS scores showed otherwise. It may possibly due to the fact that the VAS might only reflect the altered health perceptions for respondents rather than their real health status or health changes.[37, 38] However, the known-group analysis on alcohol consumptions and age of the participants deserves to be reinvestigated in the general Thai population.

The psychometric results showed that the utility scores elicited from the EQ-5D-5L (DCE model) was the best elicitation method because it could discriminate utility scores between predefined subgroups, and it has lower ceiling effect, AIC and BIC values than other elicitation methods Therefore, it implies that the EQ-5D-5L (DCE model) approach should be used to elicit the utillity scores among undergraduate students or general population who aged less than 25 years old. However, this study did not investigate some psychometric testing including test-retest reliability and responsiveness. To confirm this finding, a full psychometric testing for the utillity scores from these five approaches should be therefore investigated with the larger sample size , and more varied health conditions of general Thai population for the future studies.

There are some limitations to be addressed in this study. Firstly, other elicitation approaches used to elicit the utility scores including SG, TTO or other health-preferenced based instruments (Short-Form-6 Dimension) were not employed, so future studies should investigate the differences of utility

scores obtained from those approaches and impact of those differences on economic analyses. Secondly, this study was conducted with undergraduate students and most reported themselves as healthy, so future study investigating the differences of utility scores derived from various elicitation methods with a more varied population should be greatly encouranged. Thirdly, only known-group validity and celing effects were investigated, so a full psychometric testing for the utillity scores from these five approaches should be reinvestigated with the larger sample size among the more varied health conditions of general Thai population in the future studies.

These preliminary results suggest that the EQ-5D-5L yielded highest utility scores followed by the EQ-5D-3L and VAS, and EQ-5D-5L DCE model yielded the highest utility scores among three EQ-5D-5L models. Poor agreement of the utility scores from direct and indirect elicitation method was found which indicates that both direct and indirect approaches cannot be used to elicit the utility scores interchangeably. Furthermore, the known-group analysis showed that decreases in utility scores were found among female and smokers, which highlights the required health interventions for prevention or alleviation on undergraduate students' health. Based on the psychometric performance, the EQ-5D-5L (DCE model) is recommended to use for eliciting utility scores because it can discriminate utility scores between predefined subgroups, and it has lower ceiling effect, AIC and BIC values than other elicitation methods However, a full psychometric testing for the utillity scores from these five approaches should be reinvestigated with the larger sample size among the more varied health conditions of general Thai population for the future studies.

Acknowledgements

The authors would like to express their gratitude to all study participants for their valuable time completing the questionnaire and participating in this study. We also express our gratitude to the Faculty of Pharmaceutical Sciences, Burapha University for their grant support of this research.

References

1. Ferrans CE, Zerwic JJ, Wilbur JE, Larson JL. Conceptual model of health-related quality of life. *J Nurs Scholarsh*. 2005;37(4):336-42.
2. Kozma CM, Reeder CE, Schulz RM. Economic, clinical, and humanistic outcomes: a planning model for pharmacoeconomic research. *Clin Ther* 1993;15(6):1121-1132.
3. Goodacre S, McCabe C. An introduction to economic evaluation. *Emerg Med J* 2002;19(3):198-201.
4. Rudmik L, Drummond M. Health economic evaluation: important principles and methodology. *Laryngoscope* 2013;123(6):1341-1347.
5. Tavorncharoensap M, Sakthong P. Health utility. In: Pannarunothai S, Pilasant S, Saengsri W, Kingkaew P (eds). The Guideline of Health Technology Assessment in Thailand. Nonthaburi. Health Systems Research Institute, 2019. (in press)
6. Devlin NJ, Lorgelly PK. QALYs as a measure of value in cancer. *J Cancer Policy* 2017;11:19-25.
7. Rashidi AA, Anis AH, Marra CA. Do visual analogue scale (VAS) derived standard gamble (SG) utilities agree with Health Utilities Index utilities? A comparison of patient and community preferences for health status in rheumatoid arthritis patients. *Health Qual Life Outcomes*. 2006;4:25.
8. Sakthong P. Measurement of clinical-effect: utility. *J Med Assoc Thai*. 2008;91(Suppl 2):S43-52.
9. Wolowacz SE, Briggs A, Belozeroff V, et al. Estimating health-state utility for economic models in clinical studies: An ISPOR Good Research Practices Task Force Report. *Value Health* 2016;19(6):704-719.
10. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. *Pharmacoeconomics* 2000;17(1):13-35.
11. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33(5):337-343.
12. Gusi N, Olivares PR, Rajendram R. The EQ-5D Health-Related Quality of Life Questionnaire. In: Preedy VR, Watson RR (eds.). Handbook of disease burdens and quality of life measures. New York, NY. Springer New York, 2010: pp.87-99.
13. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37(1):53-72.
14. The National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal third edition. 2012 (Accessed on Sep. 17,2022, at http://www.nice.org.uk/media/CB1/43/GuideToMethodsOfTechnologyAppraisal2012.pdf.)
15. Hernandez Alava M, Wailoo A, Grimm S, et al. EQ-5D-5L versus EQ-5D-3L: The impact on cost effectiveness in the United Kingdom. *Value Health* 2018;21(1):49-56.
16. Yang F, Devlin N, Luo N. Cost-utility analysis using EQ-5D-5L Data: Does how the utilities are derived matter? *Value Health* 2019;22(1):45-49.
17. Tongsiri S, Cairns J. Estimating population-based values for EQ-5D health states in Thailand. *Value Health* 2011;14(8):1142-1145.
18. 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):551-558.
19. Katanyoo K, Thavorncharoensap M, Chaikledkaew U, Riewpaiboon A. A comparison of six approaches for measuring utility values among patients with locally advanced cervical cancer. *Expert Rev Pharmacoecon Outcomes Res* 2022;22(1):107-117.

20. Maheswaran H, Petrou S, Rees K, Stranges S. Estimating EQ-5D utility values for major health behavioural risk factors in England. *J Epidemiol Community Health* 2013;67(2):172-180.
21. Kraemer KL, Roberts MS, Horton NJ, et al. Health utility ratings for a spectrum of alcohol-related health states. *Med Care* 2005;43(6):541-550.
22. Kangwanrattanakul K, Parmontree P. Psychometric properties comparison between EQ-5D-5L and EQ-5D-3L in the general Thai population. *Qual Life Res* 2020;29(12):3407-3417.
23. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. New York. Routledge, 1988.
24. Sakpal TV. Sample size estimation in clinical trial. *Perspect Clin Res* 2010;1(2):67-69.
25. 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. (doi: doi: 10.1186/s12955-014-0203-3)
26. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goñi JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics* 2016;34(10):993-1004.
27. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15(2):155-163.
28. Kangwanrattanakul K. A comparison of measurement properties between UK SF-6D and English EQ-5D-5L and Thai EQ-5D-5L value sets in general Thai population. *Expert Rev Pharmacoecon Outcomes Res* 2021;21(4):765-774.
29. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. *Qual Life Res* 2000;9(8):901-910.
30. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *Pharmacoeconomics* 2018;36(6):675-697.
31. Thompson AJ, Turner AJ. A comparison of the EQ-5D-3L and EQ-5D-5L. *PharmacoEconomics* 2020;38(6):575-591.
32. Zhu J, Yan XX, Liu CC, et al. Comparing EQ-5D-3L and EQ-5D-5L performance in common cancers: suggestions for instrument choosing. *Qual Life Res* 2021;30(3):841-854.
33. Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013;22(7):1717-1727.
34. Martí-Pastor M, Pont A, Ávila M, et al. Head-to-head comparison between the EQ-5D-5L and the EQ-5D-3L in general population health surveys. *Pop Health Metrics* 2018;16(1):14. (doi: https://doi.org/10.1186/s12963-018-0170-8)
35. Kim TH, Jo MW, Lee SI, Kim SH, Chung SM. Psychometric properties of the EQ-5D-5L in the general population of South Korea. *Qual Life Res* 2013;22(8):2245-2253.
36. Kangwanrattanakul K, Auamnoy T. Psychometric testing of the health-related quality of life measurement, SF-36v2, in the general population of Thailand. *Expert Rev Pharmacoecon Outcomes Res* 2019;19(3):313-320.
37. Voutilainen A, Pitkäaho T, Kvist T, Vehviläinen-Julkunen K. How to ask about patient satisfaction? The visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric Likert scale. *J Adv Nurs* 2016;72(4):946-957.
38. Tran BX, Nguyen LH, Ohinmaa A, Maher RM, Nong VM, Latkin CA. Longitudinal and cross sectional assessments of health utility in adults with HIV/AIDS: a systematic review and meta-analysis. *BMC Health Serv Res* 2015;15:7. (doi: https://doi.org/10.1186/s12913-014-0640-z)