A Comparison of Six Approaches for Measuring Utility Values Among Patients With Locally Advanced Cervical Cancer: a Cross-sectional Study

Kanyarat Katanyoo
Faculty of Medicine Vajira Hospital

Montarat Thavorncharoensap (montarat.tha@mahidol.ac.th)
Mahidol University Faculty of Pharmacy https://orcid.org/0000-0002-8256-2167

Usa Chaikledkaew
Mahidol University Faculty of Pharmacy

Arthon Riewpaiboon
Mahidol University Faculty of Pharmacy

Research

Keywords: cervical cancer, utility, agreement, EQ-5D, TTO, VAS

Posted Date: December 30th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-136123/v1

Abstract

Background: A variety of instruments to measure health utilities are available. Nevertheless, there was a lack of agreement between utility scores generated by the different instruments. To avoid difference in estimates of cost-effectiveness arising from the use of different instruments, it is essential to know whether these widely used instruments can be used interchangeable or not. While Euro-Quality of Life Five-Dimension (EQ-5D) is recommended as a preferred instrument for measuring health utilities there is no clear consensus on the preferred version of EQ-5D. Besides EQ-5D, time-trade-off (TTO), and visual analog scale (VAS) were commonly used instruments. This study aims to investigate agreement and performance of the following 6 utilities measures; EQ-5D-3L, EQ-5D-5L (cTTO Model), EQ-5D-5L (DCE Model), EQ-5D-5L (Hybrid model), TTO, and VAS among patients with locally advanced cervical cancer in Thailand.

Methods: Face-to-face interview were conducted with 194 locally advanced cervical cancer patients. Comparison of median utility scores derived from the 6 approaches was made using Friedman test. The agreement of utility scores between each pairwise comparison was assessed by intraclass correlation coefficient (ICC) and Bland-Altman plot. Known-validity was determined by comparing two known groups with respect to status of treatment (i.e. treatment, disease free, and progression), and presences of symptoms (no symptoms, 1 symptom, and > 1 symptoms).

Results: The mean (SD) utility values derived from the 6 approaches were, from low to high, 0.755±0.248 (EQ-5D-3L), 0.801±280 (TTO), 0.806±0.156 (VAS), 0.871±0.184 (cTTO model), 0.875±0.168 (Hybrid model), and 0.900±0.142 (DCE model). Intraclass correlation coefficient showed high agreement between EQ-5D-5L and EQ-5D-3L and very high agreement between all 3 models of EQ-5D-5L. Bland-Altman plots showed wide limit of agreement except the pairwise comparison between each model of the EQ-5D-5L; cTTO model and DCE model, cTTO model and Hybrid model, and DCE model and Hybrid. All 6 approaches can significantly differentiate treatment status and presence of symptoms.

Conclusions: The commonly used utility measures, which are TTO, VAS, EQ-5D-3L and EQ-5D-5L cannot be used interchangeable among locally advanced cervical cancer patients. High agreement was found only between each pairwise comparison of the 3 value sets of EQ-5D-5L.

Background

The limited resources and the rising cost of medical treatment have led to an increasing demand for health economic evaluation. Cost-utility analysis (CUA) is a preferred method in health economic evaluation, in which health outcome is measured in terms of quality-adjusted life year (QALY) [1]. QALY is a product of the years living in the health state and utility of that health state. Utility represents the strength of an individual's preferences for health-related outcomes. Generally, utility ranges from 0 (death) to 1 (perfect health) [1].

There are two main methods of measuring utility; direct valuation method, and indirect valuation method [1, 2]. Examples of direct measures are visual analog scale (VAS), time-trade-off (TTO), and standard gamble (SG) [3–5]. On the other hand, indirect method involves using some standardized questionnaires such as Health Utilities Index Mark 3 (HUI3) [6] and the Euro-Quality of Life Five-Dimension (EQ-5D) questionnaire [7].

According to many health technology assessment (HTA) guidelines, EQ-5D is the preferred measure of utility in many countries such as England [8], the Netherlands [9], Australia [10], Canada [11], and Thailand [12, 13]. Nevertheless, to date, there is no clear consensus on the preferred version of EQ-5D for health technology assessment. For the countries, in which the value set for both EQ-5D-3L and EQ-5D-5L are available, there is a concern that the use of different versions of EQ-5D, might lead to different utility values thus affect the economic evaluation results [14, 15]. While EQ-5D-5L is recommended in the Netherlands [9], many countries including Australia [10] and Canada [11] have no clear recommendation on which version of EQ-5D is preferred. In England, the position statement 2019 [16] has raised concern over the quality of the EQ-5D-5L valuation study in the country, hence, recommended to use the utility values derived from the EQ-5D-3L instead of the EQ-5D-5L.

There was a lack of agreement between utility scores generated by the different instruments [17–25]. While it is well documented that each commonly used instrument in direct valuation method (i.e. SG, TTO, and VAS) yield different utility values [18, 22–24, 26] there is limited number of studies examining the difference between direct and indirect method [25]. To avoid difference in estimates of cost-effectiveness arising from the use of different instruments it is essential to know whether these widely used instruments can be used interchangeable or not.

In Thailand, a variety of instruments to measure health utilities are available. According to Thai HTA guideline [12], EQ-5D is a preferred instrument. When EQ-5D is not appropriate, other methods such as SG, TTO, VAS, HUI, or Short Form-6 dimension (SF-6D) can be employed but justification should be made. At present, both EQ-5D-3L [27] and EQ-5D-5L [28] value sets were developed for Thai population. The following 3 value sets of EQ-5D-5L are available: Composite time trade-off (cTTO) model, Discrete choice experiment (DCE) model, and Hybrid model. To date, EQ-5D, TTO, and VAS were commonly used in Thailand. However, limited studies have yet comprehensively assessed the agreement between these instruments in Thailand.

The purposes of this study are to investigate agreement and performance of the following 6 approaches; EQ-5D-3L, EQ-5D-5L (cTTO Model), EQ-5D-5L (DCE Model), EQ-5D-5L (Hybrid model), TTO, and VAS for measuring utility value among Thai locally advanced cervical cancer (LACC) patients. LACC was selected as it is the fourth newly diagnosis and mortality for female cancers worldwide [29] and the second most common cancer of Thai women [30]. While there is currently considerable interest in evaluating the effectiveness and cost-effectiveness of cervical cancer treatment and prevention there is limited data on impact of different instruments on health utilities among cervical cancer patients, especially locally advanced stage.

Methods

A cross-sectional study consisting of face-to-face interviews with 194 LACC patients undergoing medical treatment at Radiation Oncology Unit of Faculty of Medicine Vajira Hospital and the National Cancer Institute of Thailand was conducted between June to December 2019. Both settings are tertiary care hospitals located in Bangkok, capital city of Thailand. Inclusion criteria included: 1) diagnosed with LACC (stage IB3-IVA) according to the International Federation of Gynecology and Obstetrics (FIGO) version 2018, 2) aged > 18 years, and 3) consented to be interviewed. Patients with cognitive impairment, or the disable, or have verbal communication problem were excluded. Written informed consent was obtained from all participants before the interview. Each participant was asked to complete a 4-part questionnaire: (1) General information (i.e. age, education, occupation, health insurance scheme, income, and presence of symptoms); (2) TTO; (3) EQ-5D-5L; and (4) EQ-5D-3L. Clinical information such as clinical stage, treatment status and presences of symptoms was collected from medical record review. As this study was a part of a larger project aims at estimating the cost of treatment, sample size was determined based on the sample size required to estimate cost of treatment, which was 194.

TTO

Each patient was asked to choose between a remaining life expectancy of 10 years with her current health state and “X” years in a state of perfect health, in which X is less than 10 years. The utility of the current health state was then calculated as the number of ‘X” years that participant feel indifferent between the two options divided by 10.

EQ-5D-3L

The EQ-5D-3L is a widely used generic preference-based instrument. It comprised the following 5 dimensions: (1) mobility; (2) self-care; (3) usual activities; (4) pain and discomfort; and (5) anxiety and depression. Each dimension has 3 response categories: 1) no problems; 2) some problems; and 3) severe problems. As of 2018, EQ-5D-3L has been translated more than 180 languages including Thai [31]. The descriptive responses from the EQ-5D-3L were, then, transformed into the utility score using Thai value set [27], as shown in Table 1.

Table 1
Characteristics of the Thai EQ-5D scoring functions

	EQ-5D-3L [27]	EQ-5D-5L [28]		
		cTTO model	DCE model	Hybrid model
Intercept	0.202	-	-	-
Mobility				
Slight problem	-	0.0622	0.0686	0.0661
Some/ moderate problem	0.121	0.1254	0.0684	0.0866
Severe	-	0.2426	0.1827	0.2110
Confine to bed/ unable to	0.432	0.3228	0.3569	0.3712
Self-care				
Slight problem	-	0.0331	0.0540	0.0581
Some/ moderate problem	0.121	0.0988	0.0408	0.0706
Severe	-	0.2168	0.1700	0.1925
Unable to	0.242	0.2488	0.2138	0.2499
Usual activities				
Slight problem	-	0.0499	0.0415	0.0583
Some/ moderate problem	0.036	0.0786	0.0388	0.0712
Severe	-	0.1747	0.1278	0.1535
Unable to	0.094	0.2165	0.2238	0.2483
Pain / discomfort				
Slight problem	-	0.0415	0.0354	0.0564
Some/ moderate problem	0.072	0.0726	0.0482	0.0655
Severe	-	0.2281	0.1825	0.2069
Unable to	0.209	0.2733	0.2282	0.2564
Anxiety / depression				
Slight problem	-	0.0435	0.0452	0.0581
Some/ moderate problem	0.032	0.1067	0.0808	0.0958
Severe	-	0.2187	0.2118	0.2327
Unable to	0.11	0.2591	0.2978	0.2953
N3	0.139	-	-	-
Highest value (11111)	1	1	1	1
Lowest value (33333/55555)	-0.454	-0.3205	-0.3205	-0.4211

cTTO Composite time-trade-off, *DCE* Discrete choice experiment

EQ-5D-5L

EQ-5D-5L is the new version of the EQ-5D. It comprised the same 5 dimensions as EQ-5D-3L. However, each dimension of EQ-5D-5L has 5 response categories: 1) no problem; 2) slight problems; 3) moderate problems; 4) severe problems; and 5) extreme problems. Previous

studies found that EQ-5D-5L have superior measurement properties than EQ-5D-3L [32–35]. The value set of EQ-5D-5L was developed based on standardized valuation study protocol (EQ-VT) developed by the EuroQol Group [36]. According to the EQ-VT protocol, composite time trade-off (cTTO) and Discrete choice experiment (DCE) were introduced as the valuation method. Then, the value sets were either developed based on the data from cTTO (i.e. cTTO model), DCE (DCE model), or the combination of cTTO and DCE (Hybrid model). In our study, the descriptive responses from the EQ-5D-5L were transformed into the utility score using the existing 3 Thai value sets [28], as shown in Table 1.

VAS

In this study, we obtained the VAS utility from EQ-5D-5L VAS. For VAS, each subject was asked to rate her state of health on the day of interviews by placing a mark on a vertical line, anchored by “the best health you can imagine” (100) and “the worst health you can imagine” (0). To get the utility score, reported VAS score was divided by 100.

Data analysis

The mean, standard deviation (SD), median and range of utility values obtained from each approach were reported and compared using Friedman test and Wilcoxon signed rank test. Intraclass correlation coefficient (ICC) was used to evaluate the agreement between the 6 utility measuring approaches. A value of ICC less than 0.5, 0.5–0.74, 0.75–0.9, and higher than 0.9 indicated poor, moderate, good, and excellent agreement, respectively [37]. Further, Bland-Altman plot [38, 39] was used to assess pairwise agreements between the different approaches of measuring utility scores. For each pairwise, the difference of utility values was plotted against the mean measurement for those two approaches, along with the limits of agreement (LOA), which was the range of values that would be expected to include 95% of individual differences. According to the previous study [40], minimally important differences (MID) of EQ-5D index for cancer patients ranged between 0.06–0.09. Therefore, the LOA of less than ± 0.09 implied that such approaches can be used interchangeably. Performance of each approach was assessed in terms of ceiling effect, and known group validity. Known-group validity was examined by comparing the mean utility scores between groups with expected differences by Mann-Whitney U test and the Kruskal-Wallis test. The following two groups were examined: presences of symptoms (no symptom, 1 symptom, > 1 symptoms); and treatment status (i.e. treatment, disease free, and progression). Data were analyzed using SPSS statistical software, version 22 (SPSS Inc., Chicago, IL).

Results

Characteristics of 194 patients were reported in Table 2. The average age of patients was 53.4 ± 11.4 years old. Majority of them (78.4%) had monthly income of less than 5,000 baht (30 baht ≈ 1$). About 63% graduated with primary school or lower. With respect to treatment status, 50% were disease free while 30% were receiving treatment as concurrent chemo radiation therapy (CCRT), and 20% were in disease progression after completed CCRT, respectively. Among patients who were receiving treatment, there were 9 (15.2%), 26 (44.1%), 21 (35.6%) and 3 (5.1%) patients with stage IB3-IIA2, IIB, IIIB and IVA, respectively. The three most common symptoms among patients who were in treatment stage were abnormal vaginal bleeding (AVB) (78.0%), abnormal discharge (28.9%) and pelvic pain (27.1%), respectively. For patients in progression stage, back pain (39.5%), abdominal pain (23.7%) and AVB (15.8%) were the three most common symptoms. Back pain (23.7%), mild radiation proctitis (9.3%) and cystitis (3.1%) were the most common symptoms reported by patients with disease free.

Table 2
General and clinical characteristics of the patients (n = 194)

Characteristics	Number	(%)
Age, mean)SD(, range	53.4(11.4), 29–80	
Education level		
Primary school or below	122	62.9
Secondary school	45	23.2
Diploma	6	3.1
Bachelor or higher	21	10.8
Occupation		
Unemployed	87	44.9
Self-employed	81	41.8
Paid-employed	20	10.3
Government officer	3	1.5
Others	3	1.5
Average income per month)Baht(		
< 5,000	152	78.4
5,000–10,000	14	7.2
10,001–20,000	17	8.8
20,001–30,000	9	4.6
> 30,000	2	1.0
Health insurance coverage		
Universal Coverage Scheme	111	57.2
Social Security Scheme	57	29.4
Civil Servant Medical Benefit Scheme	13	6.7
Self-pay	10	5.2
Other	3	1.5
Treatment status		
Treatment	59	30.4
Disease free	97	50.0
Progression	38	19.6
Clinical stage (for patients who receiving treatment)		
IB3-IIA2	9	15.2
IIB	26	44.1
IIIB	21	35.6
IVA	3	5.1
Presence of symptom		

SD standard deviation

Characteristics	Number	(%)
No symptom	64	32.9
1 symptom	77	39.7
> 1 symptoms	53	27.4
SD standard deviation		

The mean, standard deviation, and median of utility values obtained from the 6 approaches were displayed in Fig. 1. The mean (SD) values of the 6 utilities were, from low to high, 0.755 ± 0.248 (EQ-5D-3L), 0.801 ± 280 (TTO), 0.806 ± 0.156 (VAS), 0.871 ± 0.184 (cTTO model), 0.875 ± 0.168 (Hybrid model), and 0.900 ± 0.142 (DCE model). Significant difference in terms of utility scores across the 6 approaches was identified in Table 3. According to the Wilcoxon signed rank test, we found that all pairs were statistical different except the following 2 pairs: TTO-VAS; and EQ-5D-5L (Hybrid model)-EQ-5D-5D (cTTO model). As shown in Fig. 1, the overall ceiling effect was 20.1% for VAS, 39.2% for EQ-5D-3L, and 32.5% for EQ-5D-5L. It should be noted that floor effect was not observed in our study.

Table 3
Comparison of utility scores across the 6 approaches

Utility: mean (SD)/ median (range)[a]						*P*-value*
EQ-5D-3L	VAS	TTO	EQ-5D-5L (cTTO model)	EQ-5D-5L (DCE model)	EQ-5D-5L (Hybrid model)	
0.755(0.248)	0.806(0.156)	0.801(0.280)	0.871(0.184)	0.900(0.142)	0.875(0.168)	< 0.001
/0.726	/0.800	/0.950	/0.938	/0.952	/0.934	
(-0.085-1)	(0.300-1)	(0.100–0.950)	(0.031–0.959)	(0.256–0.965)	(0.140–0.944)	

SD standard deviation, *EQ-5D-3L* Euro-Quality of life five-dimension-three level, *EQ-5D-5L* Euro-Quality of life five-dimension-five level, *VAS* visual analog scale, *TTO* time trade-off, *cTTO* composite time-trade-off, *DCE* discrete choice experiment

*Friedman test. Significant difference across the 6 approaches (P < 0.001)

[a]Wilcoxon signed rank test revealed significant difference across each pairwise comparison except the following 2 pairs; TTO – VAS, EQ-5D-5L (cTTO model) – EQ-5D-5L (Hybrid model)

Agreement

Table 4 show ICCs between the 6 different approaches. Poor agreement (ICC < 0.5) were observed among the following pairs; TTO- EQ-5D-5L (cTTO model), TTO-EQ-5D-5L (DCE model), TTO- EQ-5D-5L (Hybrid model). Moderate agreement (ICC: 0.5–0.74) was observed among EQ-5D-3L-VAS, EQ-5D-3L-TTO, VAS-TTO, VAS-cTTO model, VAS-DCE model, and VAS-Hybrid model. On the other hand, good agreement (ICC > 0.75) was observed between EQ-5D-3L-EQ-5D-5L, and among 3 EQ-5D-5L approaches (cTTO model-DCE model, cTTO model-Hybrid model, DCE model-Hybrid model)

Table 4
Agreement of utility values between the 6 approaches using intraclass correlation coefficient

Approach	Intraclass correlation coefficient (95% Confidence interval)*					
	EQ-5D-3L	VAS	TTO	EQ-5D-5L (cTTO model)	EQ-5D-5L (DCE model)	EQ-5D-5L (Hybrid model)
EQ-5D-3L	-	0.594 (0.459–0.695)	0.525 (0.372-0.642)	0.849 (0.354-0.939)	0.749 (0.088-0.896)	0.832 (0.306-0.932)
VAS		-	0.515 (0.356-0.635)	0.609 (0.451-0.718)	0.561 (0.262-0.721)	0.615 (0.442-0.728)
TTO			-	0.453 (0.276-0.587)	0.399 (0.196-0.550)	0.453 (0.275-0.588)
EQ-5D-5L (cTTO model)				-	0.969 (0.926-0.984)	0.995 (0.993-0.996)
EQ-5D-5L (DCE model)					-	0.984 (0.918–0.993)
EQ-5D-5L (Hybrid model)						-

EQ-5D-3L Euro-Quality of life five-dimension-three level, *EQ-5D-5L* Euro-Quality of life five-dimension-five level, *VAS* visual analog scale, *TTO* time trade-off, *cTTO* composite time-trade-off, *DCE* discrete choice experiment

**P* < 0.001 for all comparison pairs

As shown in Fig. 2, the Bland-Altman plots for each pair-wise comparison showed wide LOA except for the comparison between each model of the EQ-5D-5L; cTTO model and DCE model (LOA = 0.19), cTTO model and Hybrid model (LOA = 0.10), and DCE model and Hybrid model (LOA = 0.12).

Known group

The mean utility scores of the known-groups by treatment status and presences of symptoms were displayed in Table 5–6. All 6 approaches were able to significantly differentiate treatment status (i.e. treatment, disease free, and progression) as well as presence of symptoms (i.e. no symptoms, 1 symptom, and > 1 symptoms). For treatment status, utility of disease free stage was the highest ranging from 0.829 in EQ-5D-3L to 0.943 in TTO and EQ-5D-5L (DCE model). Utility of treatment stage ranged from 0.637 in TTO to 0.889–0.914 in EQ-5D-5L. On the other hand, utility of progression stage was the lowest ranging from 0.572 in EQ-5D-3L to 0.709–0.772 in EQ-5D-5L. In post hoc analysis, all approaches can differentiate between disease free and progression stage. On the other hand, TTO cannot identify significant difference between treatment and progression stage. All 3 models of EQ-5D-5L cannot differentiate between treatment status and disease free. With regard to presence of symptoms, all 6 approaches were able to significantly differentiate the presence of symptoms (i.e. no symptom, 1 symptom, and > 1 symptoms). In all approaches, patients with > 1 symptoms had significantly lower utility scores than those with no symptom. Also, patients with > 1 symptoms had significantly lower utility scores than those with 1 symptom. It should be noted that all approaches except VAS can differentiate those with no symptom and 1 symptom.

Table 5
Known group validity: Treatment status

Approach	Utility: mean (SD)/ median (range)			
	Treatment (n = 59)	Disease free (n = 97)	Progression (n = 38)	*P*-value*
EQ-5D-3L[a,b,c]	0.750 (0.219) /0.726 (0.178-1)	0.829 (0.190) /1.000 (0.116-1)	0.572 (0.314) /0.677 (-0.085-1)	< 0.001
VAS[a,b,c]	0.800 (0.150) /0.800 (0.500-1)	0.851 (0.128) /0.850 (0.500-1)	0.701 (0.173) /0.725 (0.300-1)	< 0.001
TTO[a,b]	0.637 (0.350) /0.800 (0.100–0.950)	0.943 (0.066) /0.950 (0.300–0.950)	0.691 (0.304) /0.925 (0.100–0.950)	< 0.001
EQ-5D-5L (cTTO model)[b,c]	0.889 (0.135) /0.927 (0.518-1)	0.924 (0.121) /0.959 (0.149-1)	0.709 (0.268) /0.793 (0.031–0.959)	< 0.001
EQ-5D-5L (DCE model)[b,c]	0.914 (0.098) /0.952 (0.665-1)	0.943 (0.088) /0.965 (0.385-1)	0.772 (0.214) /0.851 (0.256–0.965)	< 0.001
EQ-5D-5L (hybrid model)[b,c]	0.889 (0.123) 0.934 (0.577-1)	0.927 (0.107) /0.944 (0.262-1)	0.718 (0.245) /0.808 (0.140–0.944)	< 0.001

SD standard deviation, *EQ-5D-3L* Euro-Quality of life five-dimension-three level, *EQ-5D-5L* Euro-Quality of life five-dimension-five level, *VAS* visual analog scale, *TTO* time trade-off, *cTTO* composite time-trade-off, *DCE* discrete choice experiment

* *P*-value of the Kruskal-Wallis test

[a] Mann-Whitney U test: Significant difference between treatment and disease free was identified

[b] Mann-Whitney U test: Significant difference between disease free and progression was identified

[c] Mann-Whitney U test: Significant difference between treatment and progression was identified

Table 6
Known group validity: Presence of symptoms

Approach	Utility: mean (SD)/ median (range)			
	No symptom (n = 64)	1 symptom (n = 77)	> 1 symptoms (n = 53)	*P*-value*
EQ-5D-3L[a,b,c]	0.866 (0.192)/ 1 (0.116-1)	0.748 (0.192)/ 1 (0.116-1)	0.630 (0.266)/ 0.693 (-0.085-1)	< 0.001
VAS [a,b]	0.856 (0.124)/ 0.825 (0.500-1)	0.820 (0.157)/ 0.800 (0.300-1)	0.725 (0.160)/ 0.750 (0.400-1)	< 0.001
TTO [a,b,c]	0.940 (0.081)/ 0.950 (0.300–0.950)	0.774 (0.314)/ 0.950 (0.100–0.950)	0.671 (0.313)/ 0.800 (0.100–0.950)	< 0.001
EQ-5D-5L (cTTO model) [a,b,c]	0.943 (0.132)/ 1 (0.149-1)	0.871 (0.159)/ 0.927(0.261-1)	0.785 (0.231)/ 0.884 (0.031-1)	< 0.001
EQ-5D-5L (DCE model) [a,b,c]	0.956 (0.097)/ 1 (0.385-1)	0.898 (0.129)/ 0.952 (0.334-1)	0.838 (0.176)/ 0.907 (0.256-1)	< 0.001
EQ-5D-5L (hybrid model) [a,b,c]	0.945 (0.116)/ 1 (0.262-1)	0.870 (0.152)/ 0.934 (0.245-1)	0.796 (0.207)/ 0.875 (0.140-1)	< 0.001

SD standard deviation, *EQ-5D-3L* Euro-Quality of life five-dimension-three level, *EQ-5D-5L* Euro-Quality of life five-dimension-five level, *VAS* visual analog scale, *TTO* time trade-off, *cTTO* composite time-trade-off, *DCE* discrete choice experiment

* *P*-value of the Kruskal-Wallis test

[a] Mann-Whitney U test: Significant different between no symptom and > 1 symptoms

[b] Mann-Whitney U test: Significant different between 1 symptom and > 1 symptoms

[c] Mann-Whitney U test: Significant different between no symptom and 1 symptom

Discussion

Our study is the first study examining the impact of different approaches, which included both direct and indirect method on the utility scores among LACC patients. It should be also noted that the recently developed 3 value sets of the EQ-5D-5L were also included in our comparison.

We found that EQ-5D-3L yielded the least score as compared to other approaches. This is consistent with previous studies, which found that utility scores obtained from EQ-5D-3L are significantly lesser than those from TTO [25] and EQ-5D-5L [32, 35, 41]. On the other hand, utility scores derived from both direct method (i.e. VAS and TTO) were quite similar but were significantly lower than those obtained from the three EQ-5D-5L approaches (i.e. cTTO, DCE, and hybrid model). It should be noted that DCE model provided the highest utility scores and seemed to be statistically different from the other EQ-5D-5L value sets. The reason of this result might be due to the smaller coefficient values of DCE in the dimension of pain/discomfort and anxiety/depression as compared to those of cTTO and hybrid model [28]. It is worth noting that the common symptoms of cervical cancer patients were pain and AVB, abnormal discharge which could reflect on the dimension of pain/discomfort and anxiety/depression domain. Although the utility scores obtained from the following 2 pairs of EQ-5D-5L: EQ-5D-5L (DCE model)-EQ-5D-5L (Hybrid model); and EQ-5D-5L (cTTO model)-EQ-5D-5L (DCE model) were seemed to be statistically different the magnitude of the differences ranged only between 0.025 to 0.029. As the result, it could be concluded that utility scores derived from all three EQ-5D-5L are still comparable.

High agreement among EQ-5D-5L and EQ-5D-3L and very high agreement among all 3 EQ-5D-5L approaches was identified from ICC. Similarly, Bland-Altman plots indicated acceptable agreement among all 3 EQ-5D-5L approaches. The high agreement across all 3 EQ-5D-5L approaches could be possibly explained as that they were derived from the same population under the EQ-VT protocol. In fact,

Hybrid model derived from both cTTO and DCE responses. The high agreement between EQ-5D-5L and EQ-5D-3L was probably due to that they were derived from the similar 5 domains. Poor to moderate agreement between direct method and indirect method might be explained as that in direct approach respondents have to integrate all health dimensions into one value and that they might take into account other factors not reflected in the 5 dimensions of the EQ-5D.

Our study revealed that approaches for measuring utility cannot be used interchangeably. This finding is similar to other previous studies [20, 41, 42], which found that instruments used to measure utility value has substantial impact on health economic study. Specifically, EQ-5D-5L resulted in fewer QALY gain and higher incremental cost effectiveness ratio (ICER), as compared to EQ-5D-3L [14, 15]. In terms of performance, consistent with previous studies, we found that ceiling effect was diminished in EQ-5D-5L as compared to EQ-5L-3L [33–35, 43] and that EQ-5D-5L resulting in identifying more health problems yet less severe (i.e. less utility scores) [35, 44]. Study among cancer patients also revealed that EQ-5D-5Lhas better psychometric properties than EQ-5D-3L [41]. Several studies [32, 44] recommend the value set of EQ-5D-5L over that of EQ-5D-3L.

In terms of known group, our preliminary result revealed that all approaches can identify significant different in terms of utility score across treatment status, and presences of symptoms. In terms of utility score, previous studies in Taiwan [45] found that the utility scores among patients in treatment stage and progression stage, measured with EQ-5D-3L were 0.85 and 0.77, respectively. Our study found that the utility was quite similar for treatment stage (0.83) but was lower for progression stage (0.57). Previous study [45] found that utility scores derived from VAS was slightly lower than the EQ-5D-3L. However, we found that utility scores derived from EQ-5D-3L was lower than those derived from VAS. These differences could be explained by the different values attached for each domain of EQ-5D-3L in each country as well as the differences in terms of patient characteristics and health care system.

To date there are few data on validity of EQ-5D-5L among cervical cancer. It is still unclear on the validity of EQ-5D-3L among cervical cancer [46], while some mentioned that it is a valid instrument [45], the other [46] mentioned that EQ-5D-3L appeared insufficiently sensitive for cervical cancer as it might not reflect some specific aspects of quality of life related to cervical cancer such as fertility and sexual relationship. In addition, common symptoms identified by most cervical cancer patients such as abdominal vaginal discharge, AVB, mild proctitis, mild cystitis might not be directly reflected in the generic instrument such as EQ-5D.

Although our study confirmed that different approaches to measure utility cannot be used interchangeably except the 3 models of EQ-5D-5L value set. It is still unclear which models of the EQ-5D-5L value set is the most appropriate. The recent study in Japan [47] recommended Hybrid model which combine the finding from cTTO and DCE model over the other 2 models. Nevertheless, further studies were warranted in other countries. Although, our preliminary result found that all approaches have acceptable psychometric properties further studies on the validity of each utility measure among LACC patients are strongly warranted. Construct validity with the valid disease specific instrument such as EORTC QLQ-C30 (European Organization for Research and Treatment of Cancer Quality of Life Questionnaire) [48] should be further investigated. Other limitation of our study was that we only conducted in LACC patients, wider groups of patients with larger sample size is warranted. The impact of using different instruments on actual economic evaluation findings are need to be explored.

Conclusions

Our studies revealed that the commonly use utility measures which are TTO, VAS, EQ-5D-3L and EQ-5D-5L cannot be used interchangeable among LACC patients. Performance of the instruments for measuring utility among LACC is needed to be further investigated.

Abbreviations

AVB: Abnormal vaginal bleeding; CCRT: Concurrent chemoradiation therapy; cTTO: Composite time trade-off model; CUA: Cost-utility analysis; DCE: Discrete choice experiment; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire; EQ-5D: Euro-Quality of Life Five-Dimension; FIGO: the International Federation of Gynecology and Obstetrics; HTA: Health technology assessment; HUI3: Health Utilities Index Mark 3; ICC: Intraclass correlation coefficient; LACC: Locally advanced cervical cancer; LOA: Limits of agreement; MID: minimally important differences; QALY: Quality-adjusted life year; SD: Standard deviation; SF-6D: Short Form-6 dimension; SG: Standard gamble; TTO: Time-trade-off; VAS: Visual analog scale

Declarations

Ethical approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the responsible committee on human experimentation and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standard. The study was approved by the Mahidol University Institutional Review Board (MU-IRB 2019/030.1705), the Institutional Review Board from Faculty of Medicine Vajira Hospital (IRB COA 039/62 and Research Committee of National Cancer Institute of Thailand (EC COA 044/2019).

Consent for publication

Participants' identity were concealed in the publication. Data present in the publication are anonymous.

Availability of data and material

The datasets generated during and/or analyzed during the current study are available upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors did not receive support from any organizations for the submitted work.

Authors' contributions

All authors contributed to the study conception and design. Data collection was conducted by KK. Data analysis was conducted by KK and MT. The first draft of the manuscript was written by KK. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgement

This study was a part of PhD thesis of Kanyarat Katanyooat Social, Economic and Administrative Pharmacy (SEAP) Graduate Program, Faculty of Pharmacy, Mahidol University.We would like to express our thanks to the EuroQoL group for allowing us to use the EQ-5D questionnaires. We thank all members of the Radiation Oncology at Faculty of Medicine Vajira Hospital and the National Cancer Institute of Thailand for their contribution to the parts of the medical record service. Special thanks also to all patients who participated in the data collection process.

References

1. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010; 96: 5-21.
2. Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. J Health Serv Res Policy. 1999; 4(3): 174-84.
3. Torrance GW. Toward a utility theory foundation for health status index models. Health Serv Res. 1976; 11(4): 349-69.
4. Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ. 1986; 5(1):1-30.
5. Torrance GW, Feeny D, Furlong W. Visual analog scales: do they have a role in the measurement of preferences for health states? Med Decis Making. 2001; 21(4): 329-34.
6. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002; 40(2): 113-28.
7. EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. Health Policy. 1990; 16(3): 199-208.
8. Guide to the Methods of Technology Appraisal 2013. National Institute for Health and Care Excellence (NICE). 2013. https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Accessed 1 October 2020.
9. Guideline for economic evaluations in healthcare. 2016. https://tools.ispor.org/PEguidelines/source/Netherlands_Guideline_for_economic_evaluations_in_healthcare.pdf. Accessed 1

October 2020.

10. Department of Health AG. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee Version 5. 2016. https://pbac.pbs.gov.au/. Accessed 1 October 2020.
11. Guidelines for the economic evaluation of health technologies: Canada. 4th ed. Ottawa: CADTH. 2017. https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf. Accessed 1 October 2020.
12. Chaikledkaew U, Teerawattananon Y. Thai national health technology assessment guidelines. Guidelines for Health Technology Assessment in Thailand. 2nd ed. Nonthaburi: Wacharin PP; 2013.
13. Thavorncharoensap M. Measurement of utility. J Med Assoc Thai. 2014;97 Suppl 5:S43-S49.
14. 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.
15. 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): 45-9.
16. National Institute for Health and Care Excellence. Position statement on use of the EQ-5D-5L valuation set for England (updated November 2018). 2018. http://rees-france.com/wp-content/uploads/2018/10/Position-statement-on-use-of-the-EQ-5D-5L.pdf. Accessed 1 October 2020.
17. Turner N, Campbell J, Peters TJ, Wiles N, Hollinghurst S. A comparison of four different approaches to measuring health utility in depressed patients. Health Qual Life Outcomes. 2013; 11: 81.
18. Robinson A, Dolan P, Williams A. Valuing health status using VAS and TTO: what lies behind the numbers? Soc Sci Med. 1997; 45(8): 1289-97.
19. Puhan MA, Schünemann HJ, Wong E, Griffith L, Guyatt GH. The standard gamble showed better construct validity than the time trade-off. J Clin Epidemiol. 2007; 60(10): 1029-33.
20. Pan CW, Zhang RY, Luo N, He JY, Liu RJ, Ying XH, et al. How the EQ-5D utilities are derived matters in Chinese diabetes patients: a comparison based on different EQ-5D scoring functions for China. Qual Life Res. 2020; 29(11): 3087-94.
21. Martin AJ, Glasziou PP, Simes RJ, Lumley T. A comparison of standard gamble, time trade-off, and adjusted time trade-off scores. Int J Technol Assess Health Care. 2000; 16(1): 137-47.
22. Lamers LM, Stalmeier PF, Krabbe PF, Busschbach JJ. Inconsistencies in TTO and VAS values for EQ-5D health states. Med Decis Making. 2006; 26(2): 173-81.
23. Bleichrodt H, Johannesson M. Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs. J Health Econ. 1997; 16(2): 155-75.
24. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ. 2002; 11(5): 447-56.
25. Arnold D, Girling A, Stevens A, Lilford R. Comparison of direct and indirect methods of estimating health state utilities for resource allocation: review and empirical analysis. BMJ. 2009; 339: b2688.
26. Badia X, Monserrat S, Roset M, Herdman M. Feasibility, validity and test-retest reliability of scaling methods for health states: the visual analogue scale and the time trade-off. Qual Life Res. 1999; 8(4): 303-10.
27. Tongsiri S, Cairns J. Estimating population-based values for EQ-5D health states in Thailand. Value Health. 2011; 14(8): 1142-5.
28. Pattanaphesaj J, Thavorncharoensap M, Ramos-Goñi JM, Tongsiri S, Ingsrisawang L, Teerawattananon Y. The EQ-5D-5L Valuation study in Thailand. Expert Rev Pharmacoecon Outcomes Res. 2018; 18(5): 551-8.
29. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018; 68(6): 394-424.
30. International Agency for Research on Cancer. 2018. https://gco.iarc.fr/today/home. Accessed 1 October 2020.
31. EuroQol Research Foundation. EQ-5D-3L User Guide, 2018. https://euroqol.org/publications/user-guides/. Accessed 1 October 2020.
32. 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): 675-97.
33. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013; 22(7): 1717-27.

34. 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. .
35. Thompson AJ, Turner AJ. A Comparison of the EQ-5D-3L and EQ-5D-5L. Pharmacoeconomics. 2020; 38(6): 575-91.
36. 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.
37. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016; 15(2): 155-63.
38. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999; 8(2): 135-60.
39. Francq BG, Govaerts B. How to regress and predict in a Bland-Altman plot? Review and contribution based on tolerance intervals and correlated-errors-in-variables models. Stat Med. 2016; 35(14): 2328-58.
40. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007; 5: 70.
41. Zhu J, Yan XX, Liu CC, Wang H, Wang L, Cao SM, et al. Comparing EQ-5D-3L and EQ-5D-5L performance in common cancers: suggestions for instrument choosing. Qual Life Res. 2020; 10.1007/s11136-020-02636-w.
42. Craig BM, Rand K. Choice Defines QALYs: A US Valuation of the EQ-5D-5L. Med Care. 2018; 56(6): 529-36.
43. Jia YX, Cui FQ, Li L, Zhang DL, Zhang GM, Wang FZ, et al. Comparison between the EQ-5D-5L and the EQ-5D-3L in patients with hepatitis B. Qual Life Res. 2014; 23(8): 2355-63.
44. Craig BM, Pickard AS, Lubetkin EI. Health problems are more common, but less severe when measured using newer EQ-5D versions. J Clin Epidemiol. 2014; 67(1): 93-9.
45. Lang HC, Chuang L, Shun SC, Hsieh CL, Lan CF. Validation of EQ-5D in patients with cervical cancer in Taiwan. Support Care Cancer. 2010; 18(10): 1279-86.
46. Ó Céilleachair A, O'Mahony JF, O'Connor M, O'Leary J, Normand C, Martin C, et al. Health-related quality of life as measured by the EQ-5D in the prevention, screening and management of cervical disease: A systematic review. Qual Life Res. 2017; 26(11): 2885-97.
47. Shiroiwa 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.
48. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993; 85(5): 365-76.

Figures

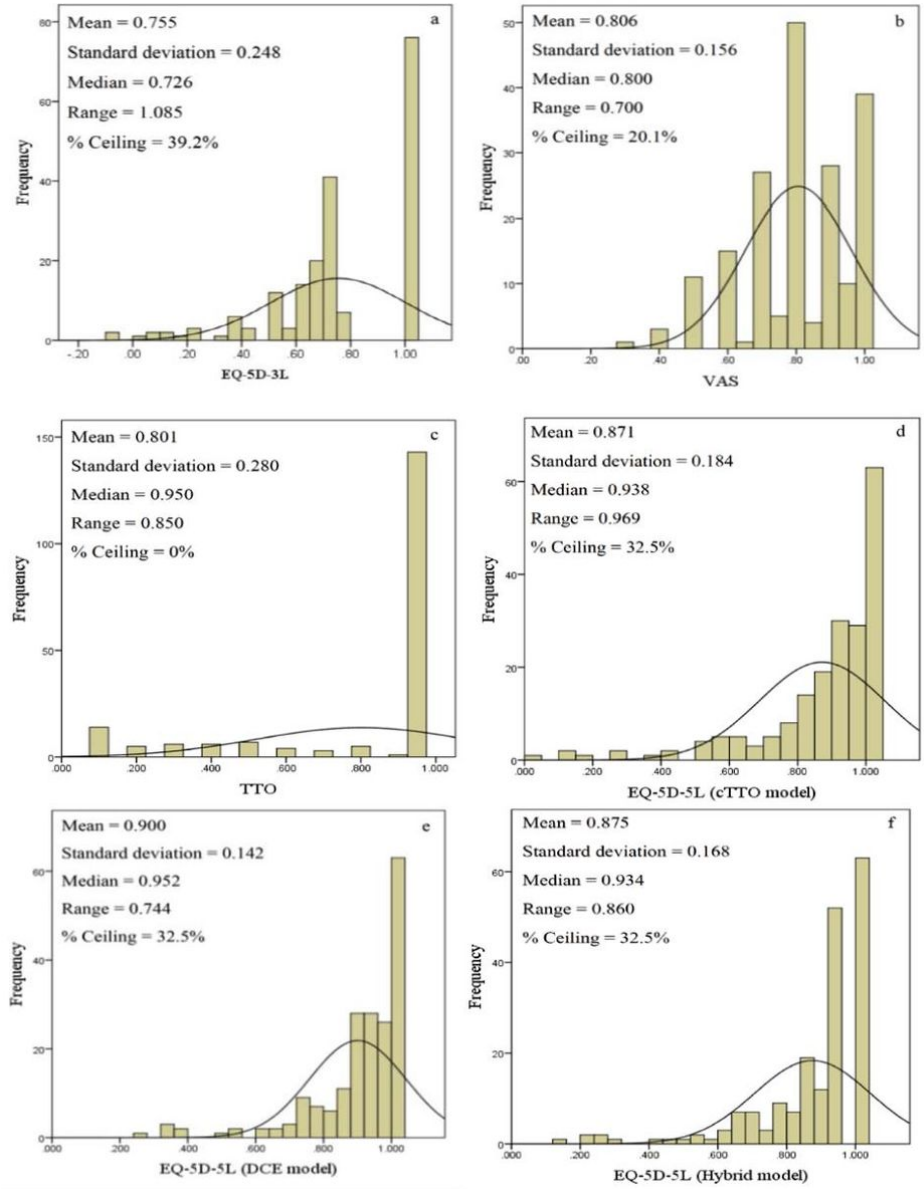

Figure 1

Histograms of utility scores from the 6 different approaches among locally advanced cervical cancer patients. a) EQ-5D-3L, b) VAS, c) TTO, d) EQ-5D-5L (cTTO model), e) EQ-5D-5L (DCE model), f) EQ-5D-5L (Hybrid model)

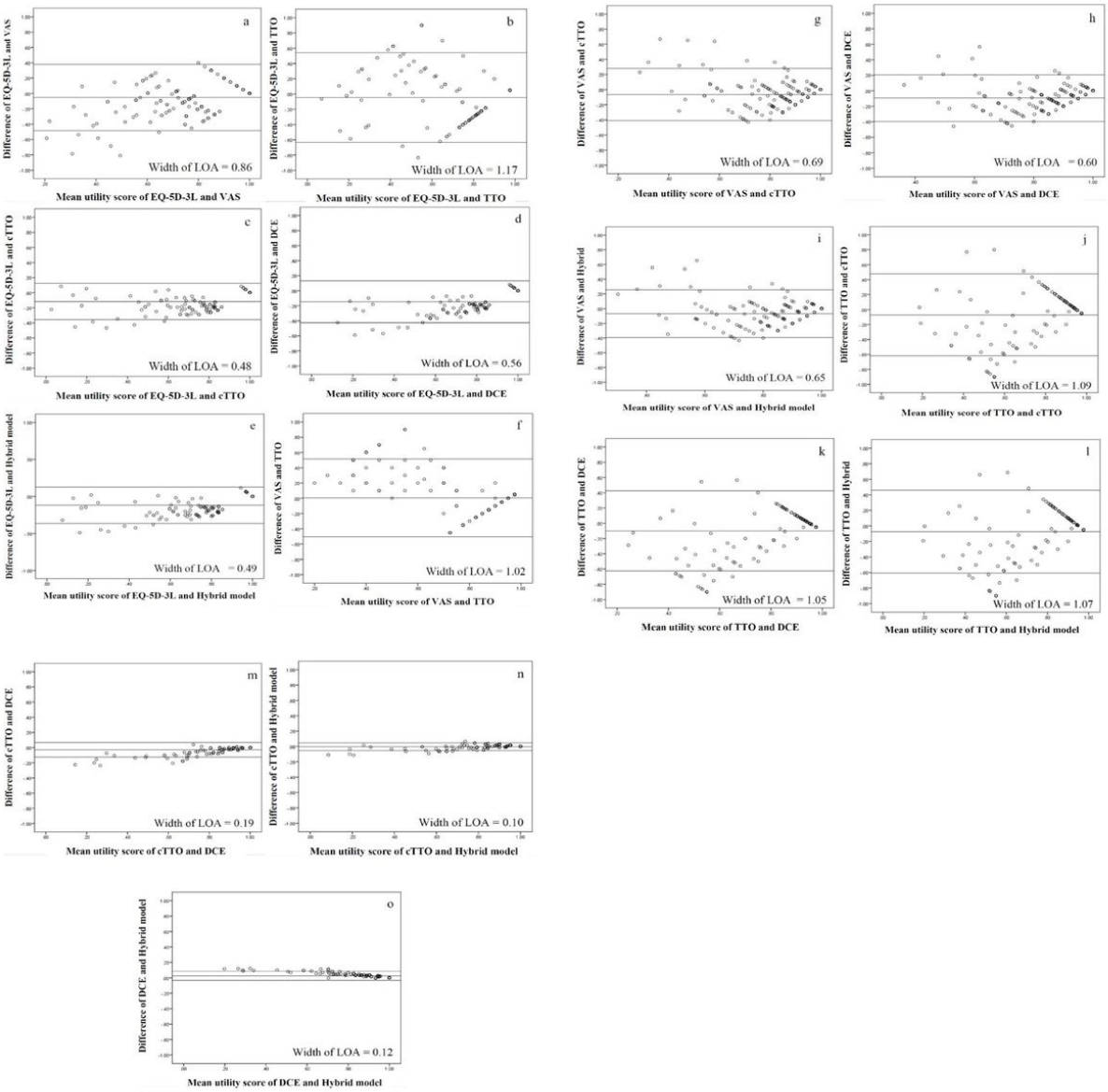

Figure 2

Bland-Altman plots between each pair of instrument a) EQ-5D-3L and VAS, b) EQ-5D-3L and TTO, c) EQ-5D-3L and EQ-5D-5L(cTTO model), d) EQ-5D-3L and EQ-5D-5L (DCE model), e) EQ-5D-3L and EQ-5D-5L(Hybrid model), f) VAS and TTO, g) VAS and EQ-5D-5L (cTTO model), h) VAS and EQ-5D-5L (DCE model), i) VAS and EQ-5D-5L (Hybrid model), j)TTO and EQ-5D-5L(cTTO model), k) TTO and EQ-5D-5L (DCE model), l) TTO and EQ-5D-5L (Hybrid model), m) EQ-5D-5L (cTTO model) and EQ-5D-5L (DCE model), n) EQ-5D-5L (cTTO model) and EQ-5D-5L (Hybrid model), o) EQ-5D-5L (DCE model) and EQ-5D-5L (Hybrid model)