International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 9, Issue 12, 2017

Original Article

RELIABILITY AND VALIDITY OF HYPOTHETICAL SCENARIOS FOR ESTIMATING WILLINGNESS TO PAY (WTP) PER QUALITY ADJUSTED LIFE YEAR (QALY) IN INDONESIA SETTING

SUSI ARI KRISTINA[1]*, DWI ENDARTI[1], TRI MURTI ANDAYANI[2], M. RIFQI ROKHMAN[1]

[1]Department of Pharmaceutics, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, Indonesia, [2]Department of Pharmacology and Clinical Pharmacy, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, Indonesia
Email: susiari_k@ugm.ac.id

Received: 19 Sep 2017 Revised and Accepted: 02 Nov 2017

ABSTRACT

Objective: To estimate validity and reliability of WTP questionnaire which WTP value can be taken as an indication of the monetary value of health gains, which may carry information regarding the appropriate height of the cost-effectiveness threshold.

Methods: Three hundred respondents, in Yogyakarta province, Indonesia, were interviewed during June 2017. We examine a value of WTP associated with the following scenarios: 1) improving moderate condition; 2) extending life during terminal illness, and 3) lifesaving. The interview ascertained maximum hypothetical WTP for one QALY using a dichotomous bidding format with an open-ended final question, along with questions about the socio-economic factors. Concerning validity, the WTP of the 3 versions of hypothetical scenarios were compared as known-group validity and analysis of the sensitivity and specificity was performed. Test-retest reliability and alpha Cronbach were employed to measure internal consistency.

Results: Analysis generally confirmed the validity and reliability of the WTP hypothetical scenarios. In terms of known group validity, there was significant difference across two scenarios (treatment v. s lifesaving), but no significant difference between mean WTP for treatment and terminal illness was found. Mean WTP for terminal illness (38 Million IDR) and lifesaving scenario (16 Million IDR) was significantly higher than that of treatment scenario (14 Million IDR). The WTP instrument showed good convergent validity (r=0.784), when comparing correlation between WTP value and utility score. Estimation of scenario's sensitivity and specificity in deriving expected WTP were 70.33 % and 38.98 %, respectively. The positive and negative predictive values were 64 % and 46 %. The test-retest reliability of WTP values indices excellent stability and reliability of the instrument with Spearman's rank correlation coefficient of 0.816 (p<0.001)

Conclusion: This study demonstrated that the WTP instrument is feasible and relatively reliable for measuring the WTP values in Indonesia. For wider application of the instrument, its validity should be investigated further. Meanwhile, adoption of WTP as an empirical evidence of societal values is encouraged.

Keywords: Reliability, Validity, Hypothetical scenario, Willingness-to-pay

DOI: http://dx.doi.org/10.22159/ijpps.2017v9i12.22657

INTRODUCTION

As health technology assessment is widely used in informing coverage decision-making in many countries, an empirical study examining the value of the QALY has recently become an important tool in policy decision making. Nowadays, more countries in Asia have recently begun to adopt economic evaluation-based reimbursement policy [1-3]. The relatively new of universal health coverage in Indonesia [4] has highlighted the need to undertake economic evaluations, especially cost-utility analysis (CUA). Interest in using willingness-to-pay when undertaking economic evaluations of health care has increased [5-6]. Willingness-to-pay refers to a method of valuing the benefits of health services with surveys using hypothetical scenarios [7]. Typically, when using willingness-to-pay, the benefits of health care services are estimated in monetary terms. Willingness-to-pay attempts to determine how much individuals are prepared to pay for defined health gain [8].

Estimating the WTP as s threshold in CUA is theoritical and methodological challenge [9-10]. A value of WTP depends on several factors including type of health gain as well as patients characteristics. A value of WTP may also be placed higher in some types of patients with higher illness severity. Furthermore, a value of WTP also varied widely on socioeconomic status, duration of health gain, and elicitation method used for interviewing [11-13]. As the result, applying WTP as a threshold for decision making tool, irrespective to context of health gain may create a potential problem. Individuals may have preferences for health intervention irrespective to the magnitude of the gain. Moreover, it is difficult for people to make these informed choices and that surveys asking people to state their preference for A versus B are hypothetical [14]. However, there is a clear need for a ceiling threshold, and more importantly, the threshold should be introduced on the basis of empirical evidence on societal values. Many studies have been made to estimate the WTP value as a ceiling threshold [2].

As with all outcome measures, it is important that reliability and validity is proven empirically to ensure that non-random error is minimized and that an instrument measures what it purports to [15]. One of the key concerns with using WTP is the hypothetical nature of such studies and the lack of evidence about whether hypothetical WTP would match the actual purchase decision [16]. Judgment of hypothetical health states is common to all valuation approaches in the health field but to date none of the non-monetary techniques have been amenable to a direct examination of validity and reliability.

The objective of this study was to assess validity and reliability of WTP questionnaire which can be taken as an indication of the monetary value of health gains, which may carry information regarding the appropriate height of the cost-effectiveness threshold.

MATERIALS AND METHODS

Questionnaire of WTP

The questionnaire consisted of three main components, namely general information, utility measure, and willingness to pay measure with three hypothetical scenario.

Utility measure

First, each respondent was assigned to imagine being in 1 hypothetical health state based on versions of questionnaire. Descriptive of each hypothetical health state was also provided. In this study, Visual analogue scale (VAS) as well as EQ-5D was employed to measure utility associated with the assigned hypothetical health state. For VAS, the respondents will be asked to look at the 20 cm, 0-100 thermometer scale where 100 is labeled "The best health state or perfect health", and 0 is labeled "the worst health state or dead". Then, they were asked to rate their current health state and the corresponding hypothetical health state on the scale.

Hypothetical scenarios

We developed 3 versions of scenarios namely treatment, terminal illness and lifesaving. The details of each version are as follow:

Version 1: treatment with health state 22333

Now, please imagine that you are staying in Health State 22333. (Point out picture of health state 22333 and read the description). After 8 mo, you can fully recover and return to perfect health without any treatment. (Point out picture of perfect health and read the description). Then, please assume that there is a new treatment that can make you immediately and fully recover to perfect health (point out the picture of perfect health). However, you have to pay-out-of pocket for the whole cost of treatment since it does not covered by any health insurance. Would you like to pay for the treatment?

You are required to pay full amount in one time within 6 mo.

Please ignore any income losses from this ill health, focus only on your quality of life over the coming 8 mo.

Version 2: terminal illness with health state 33443

Now, please imagine that you are staying in Health State 33443. (Point out picture of health state A and read the description). Without any treatment, you will die after 1 mo. Please assume that there is a new treatment that can extend your life so you will live in health state A5 for 5,5 y and 1 mo and then die. However, you have to pay-out-of pocket for the whole cost of treatment since it does not covered by any health insurance. Would you like to pay for the treatment?

You are required to pay full amount in one time within 6 mo.

Please ignore any income losses from this ill health, focus only on your quantity of life over the coming 5,5 y.

Version 3: Heart failure with health state 55555

Now, please imagine that you have severe disease (Health state 55555). Without any treatment, you will die today. Please assume that there is a new treatment that can make you fully recover to perfect health (see picture) but you will die afte 5 mo. However, you have to pay-out-of pocket for the whole cost of treatment since it does not covered by any health insurance. Would you like to pay for the treatment?

You are required to pay full amount in one time within 6 mo.

Please ignore any income losses from this ill health, focus only on your quantity and quality of life over the coming 6 mo.

WTP elicitation

Dichotomous bidding technique followed by open-ended questions was employed to examine respondents' WTP for each scenario. For treatment scenario, a specified period of time being in that hypothetical health state followed by complete recovery was assumed. Then, respondents were asked to indicate WTP value for the treatment that can make immediately recover to perfect health (EQ-5D state: 11111). For terminal illness, 1 mo lead time was introduced. In this situation, it was assumed that respondents in terminal illness (as described as EQ-5D state 33443) would die after one month from today. Respondents were then asked to indicate his/her WTP for a specific treatment that could have extended his/her life for a certain period. Regarding the lifesaving situation, it was assumed that respondents had a serious disease and he/she would die immediately. Then, those respondents were asked to pay for a treatment that can bring him/her back to full health (EQ5D state: 11111) but for only certain period of time.

To avoid starting point bias, each respondent was randomly assigned on a certain starting price. The yes/no answer to the first price offered to the respondent determine the next price offered. Dichotomous bidding as percentage of GDP per capita was used. If the answer is "yes", the bid amount increased in the second bid. If the initial answer was "no", the bid amount would be reduced. The open-ended question was asked after the second bidding to examine the maximum WTP amount. If the respondent indicated that he/she did not want to pay at all for the treatment, his/her reason(s) for not paying is asked. Each respondent was asked to determine his/her WTP for the treatment, for which he/she has to pay out-of-pocket one time within the next 6 mo. The respondent was asked to consider carefully before making his/her decision regarding the amount of money he/she is WTP for and the feasibility that he/she can pay that amount of money (table 1)

Table 1: Algoritm of dichotomous bidding

Starting point	Time of GDP per capita	First bidding value	Answer	Time of GDP per capita	First bidding value
1	0.02	900,000	No	0.0125	562,500
			Yes	0,05	2,250,000
2	0.05	2,250,000	No	0.025	1,125,000
			Yes	0.01	4,500,000
3	0.1	4,500,000	No	0.05	2,250,000
			Yes	0.2	9,000,000
4	0.2	9,000,000	No	0.1	4,500,000
			Yes	0.4	18,000,000
5	0.4	18,000,000	No	0.2	9,000,000
			Yes	0.8	36,000,000
6	0.8	36,000,000	No	0.4	18,000,000
			Yes	1.2	54,000,000
7	1.2	54,000,000	No	0.8	36,000,000
			Yes	1.5	67,500,000

Data collection

Before data collection begins, the proposal was submitted to Ethical Review Committee for Research on Human Subjects for approval. Data were collected via face-to face interview. Pilot testing was performed to ensure appropriateness, and clarity of developed questionnaire and scenario. All interviewers were trained and interviewer guideline was developed in order to ensure the consistency of data collection. During the field work, supervision plan was implemented.

Data analysis

Descriptive statistics were used to explain demographic and patients characteristics as well as the WTP values, using percentage and frequencies for the categorical variables and means, SD for the continuous variables. The chi-square test was employed for categorical variables and Kruskal-Wallis test was used to evaluate the differences between versions of scenario.

Internal consistency was assessed using Cronbach's alpha, and Spearman's rank correlation was used to assess test-retest reliability. Convergent validity was assessed using Spearman rank correlation between WTP values and the scores on utility. Known group validity was assessed through the association of three levels of scenario and WTP values using Chi square tests. Known group validity was assessed by using Kruskal Wallis test, assuming that low level of severity case also report lower WTP value. All analysis was performed using SPSS version 15.0. The significance level was set at p less than 0.05.

RESULTS AND DISCUSSION

Characteristics of respondents

Sociodemographic information of 300 respondents (response rate 75%) classified by each version was displayed in table 2. It was found that 30.33% of the respondents were male. Average age of the respondents was 40.64 y old with the standard deviation of 11.37 y. Most of the respondents (71.33%) graduated with primary school education or lower. Approximately 75% of the respondents had monthly household income lower than 2 million IDR. No significant differences across questionnaire versions were found.

Table 2: Sociodemographic characteristics of respondents

Characteristics	Total	Treatment (N=100) %	Terminal illness (N=100) %	Lifesaving (N=100) %	p-value
Gender					0.501
Male	91 (30.33)	33	35	23	
Female	29 (69.64)	67	65	77	
Age	40.64 (11.37)	40 (11.17)	39.88 (11.45)	42.04 (12.41)	
Years (Mean, SD)	40.64 (11.37)	40 (11.17)	39.88 (11.45)	42.04 (12.41)	0.437
Education					
Primary school or lower	214 (71.33)	74	75	65	0.086
Secondary school or higher	84 (28.67)	26	25	35	
Number of dependents (Mean, SD)	3.43 (2.34)	3.6 (2.00)	3.5 (2.19)	3.2 (2.56)	0.576
Income (000 IDR)					0.060
2,000 IDR or lower	225 (75)	62	84	79	
Higher than 2,000 IDR	75 (25)	38	16	21	

Proportions of respondents who were unwilling to pay in each version were displayed table 3. It was also found that there were significant differences across situations. As shown in table 3, only 3% of the respondents indicated that they were unwilling to pay for treatment situation while about 9 %, and 29% indicated that they were unwilling to pay for life extension during terminal illness, and lifesaving, respectively.

Table 3: Percentage of unwilling to pay for each version

Hypothetical scenario	N (%)	p-value
Treatment (100)	3 (3)	<0.001*
Terminal illness (100)	19 (9)	
Lifesaving (100)	29 (29)	
Total (300)	51 (17)	

*significant at p<0.05

Table 4 shows the reasons for unwillingness to pay for each situation. Concerning treatment scenario (version 1), it was found that the most frequent reasons for unwilling to pay for treatment was that the given health state was not too bad so they could live with it. On the other hands, the most frequent reason for unwilling to pay for terminal illness was that they do value the treatment but they could not afford to pay. Regarding lifesaving situation, "I would rather die right away" was the most frequent reasons for unwilling to pay.

Table 4: Reasons for unwilling to pay for three hypothetical scenario

Reasons for unwilling to pay	N (%)	p-value
The given health state is not too bad. I could live with it	4 (7.8)	<0.001*
I would get better anyway, so it is not worth paying for the treatment	7 (13.7)	
I do value the treatment but I cannot afford to pay anything for it	15 (29.4)	<0.001**
I'd rather die right away	25 (49.01)	
Total	51 (100)	

Validity and reliability

Known group validity

As shown in table 5, among those indicated that they were willing to pay, the respondents would pay IDR 24 million more for lifesaving as compared to those for treatment.

In addition, amount of willingness to pay for terminal illness was IDR 2 million higher than those for treatment.

When comparing treatment scenario with terminal illness and lifesaving scenario, it was found that there were significant differences across these 2 scenarios. Based on our analysis, mean WTP for terminal illness and lifesaving scenario was significantly higher than that of treatment scenario. However, no significant difference between mean WTP for treatment and terminal illness was found. When looking at treatment scenario, it was found that mean WTP values for treatment (version 1) were significantly lower than those for terminal illnesses (version 2) and lifesaving (version 3) condition.

Table 5: Known group validity of WTP values across three scenarios

Hypothetical scenario	Mean of WTP	Minimum (IDR)	Maximum (IDR)	P value
Treatment	14,097,000	1,000,000	100,000,000	0.065**
Terminal illness	16.715.000	500,000	100,000,000	
Lifesaving	38.890.000	1,000,000	200,000,000	<0.001*

Convergent validity

Table 6 showed the correlation between WTP value and utility score (r=0.784). Respondents were consistent to pay more for the lower utility score status. It means that utility score, and context of health gains are significant predictors of whether the respondents would pay higher or lower.

Table 6: Convergent validity of WTP value and utility score

WTP value	Mean VAS score	r
14,097,000[1]	40.72	0.784
16.715.000[2]	38.70	
38.890.000[3]	19.6	

Note: [1]scenario for treatment, [2]scenario for terminal illness, [3]scenario for lifesaving

Sensitivity and specificity

To determine how good the hypothetical scenario perform to identify responses among respondents willing to pay in three situations, sensitivity and specificity were evaluated. For the sensitivity and specificity analysis, only two groups of scenarios were used, treatment and terminal illness together as one group, and lifesaving as the second group. Estimation of scenario sensitivity and specificity were 70.33 % and 38.98 %, respectively. The positive and negative predictive values were 64 % and 46 %.

Reliability

Regarding the internal consistency reliability, Cronbach's alpha test was 0.721 for the items of WTP questionnaire, which is higher than acceptable value 0.7, indicates excellent reliability. Item 1 to 5 to total correlation coefficient ranged from 0.428 to 0.624 (table 7). The test-retest reliability of WTP values indices excellent stability and reliability of the instrument with Spearman's rank correlation coefficient of 0.816 (p<0.001) (table 8).

Table 7: Reliability test of questions of WTP scenarios

Question no	*mean±SD*	*Corrected item-total correlation*	*Cronbach's alpha* if item deleted
1	0.596±0.491	0.458	0.752
2	0.880±0.325	0.437	0.734
3	0.820±0.384	0.595	0.727
4	0.836±0.371	0.624	0.716
5	0.940±0.237	0.428	0.728

Table 8: Test retest reliability

Hypothetical scenario	Mean of WTP-1	Mean of WTP-2	r
Treatment	14,097,000	15,450,000	0.796
Terminal illness	16.715.000	17,289,000	0.832
Lifesaving	38.890.000	36,976,000	0.693

The reliability and validity of willingness-to-pay instruments have not been adequately addressed in the literature [12, 17-19]. This is a first attempt to predict validity and reliability of WTP scenario using dichotomous bidding combined with open ended question through the application of societal value perspective in Indonesian context. The results of this study appeal for taking WTP as a valid methodology. In this study significant difference was found between mean willingness-to-pay values in terminal illness and lifesaving. This suggests that the elicited willingness-to-pay is likely to reflect the real value.

However, we found a little gap of WTP value between treatment and terminal illness. In this case we need to pay attention more about participants' response in hypothetical situation, where participants might not be able to imagine the related situation [20]. The questionnaires used in this study were rather complex, involving hypothetical decision-making scenarios. As a result, we were concerned about the respondents' comprehension of the scenarios, which would affect the validity of their responses.

According to our findings, the responses from the utility score and the WTP values were consistent with prior expectations [18, 21]. For each scenario, the utility associated with the more severe health states was lower than that of the less severe health states. Similarly, the WTP associated with the more severe health states was higher than that of the less severe health states.

While societal WTP are needed to make decisions making in resource allocations, it should be noted that the methods used is often seek to measure an individual's perspective [8].

Many studies suggested that hypothetical willingness-to-pay typically overestimates real willingness-to-pay [16, 22]. Participants may exaggerate willingness-to-pay values and in real life do not necessarily behave the same way as stated in their responses [23].

It has been recommended that the mean willingness-to-pay value should be deflated by an ad hoc 50 percent to account for potential bias [18]. Despite the advantages of the WTP method compared with other approach, uncertainties about reliability and validity have thus limited its application in health care decision-making [14]. WTP study must be designed carefully and accounted for culturally specific issues.

Several limitations are need attentions to the study. First, the study was unable to differentiate the WTP values of treatment and terminal illness, despite the apparently gap between them. Experiences with health care services had an impact on patient's preferences and the WTP value. Second, only small sample sizes for each scenario, covered only one province and response rate is less than 90% which limited the statistical power.

CONCLUSION

This study demonstrated that the willingness-to-pay instrument is feasible and relatively reliable for measuring the WTP values in Indonesia. However, further investigations should be undertaken to confirm the validity of the approach, especially with respect to treatment services which cover wide range of services. Application of the method in other cultural or social contexts would require an appropriate adaptation, taking into account characteristics of the health care system and sociocultural features. Decision-makers should be encouraged to apply the willingness-to-pay approach in the economic evaluation threshold.

ACKNOWLEDGEMENT

We thank all participants who participated in this survey, and the master students Melviani, Fitria Dhirisma, Fitri Andriyani Fatimah, who helped in the collection of data.

AUTHORS' CONTRIBUTION

Data collection process was performed by Dwi Endarti and team, while Trimurti Andayani and M Rifqi Rokhman conducting review on questionnaire and data analysis. Susi Ari Kristina contributed on design of questionnaire, supervising data analysis process and writing manuscript.

COMPETING INTERESTS

The author(s) declare that they have no competing interests.

REFERENCES

1. Thavorncharoensap M, Teerawattananon Y, Natanant S, Kulpeng W, Yothasamut J, Werayingyong P. Estimating the willingness to pay for a quality-adjusted life year in Thailand: does the context of health gain matter? Clinicoecon Outcomes Res 2013;5:29-36.
2. Nimdet K, Chaiyakunapruk N, Vichansavakul K, Ngorsuraches S. A systematic review of studies eliciting willingness-to-pay per quality-adjusted life year: does it justify CE threshold? PLoS One 2015;10:e0122760.
3. Shillcutt SD, Walker DG, Goodman CA, Mills AJ. Cost effectiveness in low-and middle-income countries: a review of the debates surrounding decision rules. Pharmacoeconomics 2009;27:903-17.
4. Kristina SA, Endarti D, Prabandari YS, Ahsan A, Thavorncharoensap M. Burden of cancers related to smoking among the indonesian population: premature mortality costs and years of potential life lost. Asian Pac J Cancer Prev 2015;16:6903-8.
5. Pinto-Prades JL, Loomes G, Brey R. Trying to estimate a monetary value for the ALY. J Health Econ 2009;28:553-62.
6. Mason H, Baker R, Donaldson C. Willingness to pay for a QALY: past, present and future. Expert Rev Pharmacoecon Outcomes Res 2008;8:575-682.
7. Olsen JA, Smith RD. Theory versus practice: a review of 'willingness-to-pay' in health and health care. Health Econ 2001;10:39-52.
8. Bobinac A, Van Exel NJ, Rutten FF, Brouwer WB. Willingness to pay for a quality-adjusted lifeyear: the individual perspective. Value Health 2010;13:1046-55.
9. Zhao FL, Yue M, Yang H, Wang T, Wu JH, Li SC. Willingness to pay per quality adjusted lifeyear: Is one threshold applicable for all decision-making? Value Health 2010 13:A538.
10. Shiroiwa T, Sung YK, Fukuda T, Lang HC, Bae SC, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ 2010;19:422-37.
11. Lieu TA, Ray GT, Ortega-Sanchez IR ea. Willingness to pay for a QALY based on community member and patient preferences for temporary health states associated with herpes zoster. Pharmacol Economics 2009;27:1005-16.
12. Wagner TH, Wu T, Duenas GV, Pasick RJ. Willingness to pay for mammography: item development and testing among five ethnic groups. Health Policy 2000;53:105-21.
13. Weinstein MC. How much are Americans willing to pay for a quality-adjusted life year? Med Care 2008;46:343-5.
14. Bhatia MR, Fox-Rushby JA. Validity of willingness to pay: hypothetical versus actual payment. Appl Economics Lett 2003;10:737-40.
15. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med 2006;119:166, e7- e16.
16. Smith RD, Richardson J. Can we estimate the "social" value of a QALY? Four core issues to resolve. Health Policy 2005;74:77-84.
17. Gyrd-Hansen D, Kjaer T. Disentangling WTP per QALY data: different analytical approaches, different answers. Health Econ 2012;21:222-37.
18. Shiroiwa T, Igarashi A, Fukuda T, Ikeda S. WTP for a QALY and health states: More money for severer health states? Cost Effectiveness Resource Allocation 2013;11:22.
19. Krabe PFM, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five health-state valuation methods. Soc Sci Med 1997;45:1641-52.
20. Mahmudah RL, Ikawati Z, Wahyono D. A qualitative study of perspective, expectations, and needs of education in chronic obstructive pulmonary disease (COPD). Int J Chem Pharm Res 2017;9:32-5.
21. Park H. Association between rheumatoid arthritis and health-related quality of life in Korean women aged 50 y and over. Asian J Pharm Clin Res 2017;10:372-5.
22. Pinto-Prades JL, Sanchez-Martinez FI, Corbacho B, Baker R. Valuing QALYs at the end of life. Soc Sci Med 2014;113:5-14.
23. Martin-Fernandez J, Polentinos-castro E, del Cura-Gonzalez MI, Ariza-Cardiel G, Abraira V G-PS. Willingness to pay for a quality-adjusted life year: an evaluation of attitudes towards risk and preferences. BMC Health Serv Res 2014;14:28.