Estimating the willingness to pay for a quality-adjusted life year in Thailand: does the context of health gain matter?

Montarat Thavorncharoensap1,2
Yot Teerawattananon2
Sirin Natanant2
Wantanee Kulpeng2
Jomkwan Yotheasamut2
Pitsaphun Werayingyong2

1Department of Pharmacy, Faculty of Pharmacy, Mahidol University, Thailand; 2Health Intervention and Technology Assessment Program (HITAP), Ministry of Public Health, Thailand

Background: This study aims to elicit the value of the willingness to pay (WTP) for a quality-adjusted life year (QALY) and to examine the factors associated with the WTP for a QALY (WTP/QALY) value under the Thai health care setting.

Methods: A community-based survey was conducted among 1191 randomly selected respondents. Each respondent was interviewed face-to-face to elicit his/her health state preference in each of three pairs of health conditions: (1) unilateral and bilateral blindness, (2) paraplegia and quadriplegia, and (3) mild and moderate allergies. A visual analog scale (VAS) and time trade off (TTO) were used as the eliciting methods. Subsequently, the respondents were asked about their WTP for the treatment and prevention of each pair of health conditions by using a bidding-game technique.

Results: With regards to treatment, the mean WTP for a QALY value (WTP/QALY_{treatment}) estimated by the TTO method ranged from 59,000 to 285,000 baht (16.49 baht = US$1 purchasing power parity [PPP]). In contrast, the mean WTP for a QALY value in terms of prevention (WTP/QALY_{prevention}) was significantly lower, ranging from 26,000 to 137,000 baht. Gender, household income, and hypothetical scenarios were also significant factors associated with the WTP/QALY values.

Conclusion: The WTP/QALY values elicited in this study were approximately 0.4 to 2 times Thailand’s 2008 GDP per capita. These values were in line with previous studies conducted in several different settings. This study’s findings clearly support the opinion that a single ceiling threshold should not be used for the resource allocation of all types of interventions.

Keywords: ceiling threshold, health resource allocation, time trade off, visual analog scale

Introduction

In a resource-constrained society, economic evaluations – particularly cost–utility analyses (CUA) – are increasingly used as important tools to support efficient health resource allocation. For the CUA to become a practical tool for systematic, transparent, and consistent policy decision making, a ceiling threshold is needed. Concerning such a threshold, health technology with an incremental cost-effectiveness ratio (ICER) below the accepted ceiling threshold is considered to be cost-effective. However, no scientific standard exists for setting that threshold. Sloan acknowledges an arbitrary threshold of US$50,000 per quality-adjusted life year (QALY),1 and the Commission on Macroeconomics and Health (2001) recommends an equally arbitrary threshold of 1–3 times the GDP per capita per disability-adjusted life year (DALY).2 Several arguments have frequently cited these numbers.3,5 However, rather than relying on such an arbitrary ceiling threshold, the value of a QALY estimated in terms of a society’s
willingness to pay (WTP) per QALY should be adopted as the ceiling threshold. This calculation simultaneously combines the WTP and the measured utility value.

Estimating the WTP/QALY value involves both theoretical and methodological challenges. Existing evidence reveals that the WTP/QALY is not a single value. It depends on several factors, including context and the type of health gain. A great deal of evidence indicates that a QALY gained by improving the quality of life or extending a life is worth less than a QALY gained by saving a life. A QALY gained from treatment is also theoretically different from a QALY gained from prevention. In addition, WTP/QALY values vary widely depending on socioeconomic status, duration of health gained, and the elicitation method used for estimating utility and WTP.

In spite of these challenges, a ceiling threshold is clearly needed. Determining the threshold should be a research priority. Johnson even insisted that, although an empirical study of this topic involves some methodological flaws, it is still preferable to using an arbitrary threshold value. Thus, scholars have proposed that a disease-specific and context-specific threshold should also be considered when assessing the value of a QALY.

Over the past few years, many attempts have been made to estimate an appropriate WTP/QALY value for the ceiling threshold. However, it is worth noting that almost all of these studies were confined to industrialized nations. This includes the recent European value of a quality-adjusted life year (EuroVaQ) study, which was conducted to determine the monetary value of a QALY across a number of European member states, including the Netherlands, UK, France, Spain, Sweden, Norway, Denmark, Poland, and Hungary. Although developing countries in Asia and other parts of the world have recently begun adopting reimbursement policies based on economic evaluations, only two studies exist that examine the WTP/QALY in these settings.

We consequently offer the first study aimed at estimating the WTP/QALY values and examining the factors associated with the WTP/QALY values in Thailand, where, since 2008, cost-effectiveness evidence has been required for medication reimbursement. In this study, we examine the value of a QALY derived from the treatment of the three selected health conditions—blindness, paralysis, and allergies. It should also be noted that this is the first study to compare the WTP/QALY values obtained from prevention with those obtained from treatment interventions.

Materials and methods

Samples

The sample size was calculated according to the following formula:

$$N = \left[ \frac{Z_{\alpha/2} V}{\Delta} \right]^2 D,$$

where $N$ is the required sample size, $Z_{\alpha/2}$ is the $(1 - \alpha)%$ confidence interval statistic, $V$ is the coefficient of variation, $D$ is the design effect, and $\Delta$ is the desired maximum fractional error in the mean. When $V$ is estimated at 2 and $\Delta$ is set at 0.15 and $D$ is equal to 1.5, the required sample size is 1025. In order to cover the incomplete interviews that might occur, the sample size was inflated to 1200. The samples were then allocated to each province based on the total population of each province.

A stratified multistage cluster random sampling technique was employed in this study. All provinces were classified into five regions based on their geographic locations (four regions and Bangkok). All provinces in each region were then stratified by income levels based on the average income per household obtained from the National Statistical Office of Thailand. The levels were high (average monthly household income per monthly household ≥ the 50th percentile) and low (average monthly household income < the 50th percentile), resulting in nine strata.

For each stratum, one province was purposely selected. In each selected province (Bangkok, Ang Thong, Chon Buri, Chaing Mai, Phayao, Khon Kaen, Surin, Trang, and Chumphon), members of the general Thai population who met the eligibility criteria were randomly chosen with the co-operation of the National Statistical Office. Inclusion criteria were as follows: (1) being between 15–65 years of age, (2) having a monthly household income > 5000 baht, and (3) being able to read and write Thai. Exclusion criteria were as follows: (1) the presence of a disability, (2) the inability to answer a series of complex theoretical questions, and (3) the refusal to participate in the study.

Study instrument

There were three questionnaire versions. The first version focused on a scenario of blindness, while the second and third versions were centered on scenarios of paralysis and allergies, respectively. For each version, two levels of severity were specified for each health condition: the first version, unilateral and bilateral blindness; the second version,
paraplegia and quadriplegia; and third version, mild and moderate allergies. The three scenarios were selected for the following reasons: (1) they represent a QALY gained from a wide range of severities, ie, mild, moderate, and severe health states; (2) the conditions can be easily imagined by most of the general population; and (3) the level of severity associated with the conditions can be clearly differentiated. As the questionnaires were complex, to reduce the respondent’s burden, each respondent answered only one version of the questionnaire.

The study was conducted in accordance with the Helsinki Declaration of the World Medical Association. The data collection process began after the conduction of three pilot tests on 145 respondents to ensure the clarity and understanding of the questionnaires and to determine the range of WTP values. Based on the pilot testing and actual treatment costs for the hypothetical health states, two different 12-starting-price ranges for the treatment situations were used (5000, 10,000, 20,000, 30,000, 50,000, 75,000, 100,000, 125,000, 150,000, 200,000, 250,000, and 300,000 baht for the blindness and paralysis scenarios, and 1000, 3000, 5000, 7500, 10,000, 12,000, 15,000, 20,000, 30,000, 50,000, 75,000, and 100,000 baht for the allergy scenario). Furthermore, for all health scenarios in the prevention situations, a different 12-starting-price range was used (500, 1000, 2000, 3000, 5000, 7500, 10,000, 12,000, 15,000, 20,000, 30,000, 50,000, 75,000, and 100,000 baht for the allergy scenario). For TTO, subjects were offered a choice between continuing to live in their current health state for 10 years or trading years of life in exchange for becoming in perfect health. Moreover, they were offered a choice between continuing to live in the given hypothetical health state for 10 years or trading years of life in exchange for returning to their current health state.

**WTP measure**

For each respondent, a WTP amount was determined by using a bidding-game technique. For the treatment situation, each respondent was asked to imagine being in the hypothetical scenario for 5 years then returning to his/her current health state without any treatment. Then he/she was told that there was a new treatment that could help him/her recover immediately and fully return to his/her current state of health. However, the treatment was not covered by any health insurance, so he/she had to pay out of pocket for the whole cost of the treatment. Then, he/she was asked to indicate his/her WTP for treatment of the given hypothetical scenario. In the prevention situation, the questions were used to arrive at a WTP amount associated with the percentile elimination of the risk (from 10% to 50%) of being in the hypothetical health states. In this study, each respondent was required to make a one-time out-of-pocket payment within the next 6 months. Each respondent was asked to think carefully before making his/her decision and to be sure that he/she could actually pay that amount of money. If he/she paid too little, he/she might not get the treatment, as it was not financially worthwhile for the provider. However, if he/she paid too much, it would impact his/her family’s finances.

To avoid a starting point bias, each respondent was randomly assigned a starting price. The yes/no answer to the first price offered to the respondent provided the criteria for the next price offered. If the answer was “yes,” the bid amount would increase to the next step until the respondent said “no.” If the initial answer was “no,” the bid amount would decrease one step, and this process would continue until the respondent said “yes.” The maximum WTP from this bidding method was determined as the midpoint between the

---

**Utility measure**

Each respondent was asked to imagine being in the two hypothetical health states that were presented in his/her version of the questionnaire. Descriptions of each hypothetical health state were also provided to the respondents during the interview. A visual analog scale (VAS) and time trade off (TTO) technique were employed to measure utilities associated with the current health state and the two hypothetical health states of each respondent.

In this study, VAS and TTO were used to measure utility for various reasons. First, both of them are commonly used methods. TTO is based on a solid theory, while VAS is a short, simple, and appropriate method to use as a warm-up exercise. Second, in our pre-test, the respondents seemed to understand these two methods better than the standard gamble (SG) method. Lastly, although the indirect methods of multi-attribute utility measure (MAUM), such as EQ-5D, might seem appropriate, no value set for Thailand’s population existed during our study period.

For VAS, the respondents were asked to look at the 20 cm, 0–100 unit thermometer scale, where 0 was labeled “The worst health state or dead” and 100 was labeled “The best health state or perfect health.” Then, they were asked to rate their current health state and the two corresponding hypothetical health states on the scale. For TTO, subjects were offered a choice between continuing to live in their current health state for 10 years or trading years of life in exchange for becoming in perfect health. Moreover, they were offered a choice between continuing to live in the given hypothetical health state for 10 years or trading years of life in exchange for returning to their current health state.

The study was conducted in accordance with the Helsinki Declaration of the World Medical Association. The data collection process began after the conduction of three pilot tests on 145 respondents to ensure the clarity and understanding of the questionnaires and to determine the range of WTP values.
lower acceptance bid and the higher rejection bid. However, if the respondents were willing to pay less than the minimum offered bid or higher than the maximum offered bid, their WTP amount was determined using open-ended questions. If the respondent indicated that he/she would not pay for treatment, his/her reason for not doing so was asked.

Data analysis

All statistical methods were performed using SPSS version 13.0 for Windows (SPSS Inc, Chicago, IL, USA) and STATA 8.0 (StataCorp LP, College Station, TX, USA). The WTP/QALY value was calculated using the following formula:

\[
\text{WTP/QALY} = \frac{\text{WTP amount}}{\sum_{i=1}^{5} (\text{Utility of current state} - \text{Utility of hypothetical state})/(1+r)^{t}}
\]

where the time horizon \(t\) is 1–5 years and the discounting rate \(r\) is 0.03.

For each scenario, the WTP/QALY values obtained from the two levels of severity were compared using a Wilcoxon signed-rank test. A Kruskal–Wallis test was also performed to examine the differences across the three scenarios. A generalized linear mixed model was used to evaluate the effects of the following factors on the WTP/QALY value: gender, household income, education level, and health scenario. These factors were chosen because they are significantly associated with WTP/QALY in bivariate analysis. Finally, since there are theoretical arguments associated with the validity of VAS, the WTP/QALY values estimated from the TTO method were emphasized in this study.

Results

Table 1 shows the characteristics of the respondents. Of the 1191 respondents, 399 were randomly selected for the blindness scenario, while 399 and 393 were randomly selected for the paralysis and allergen scenarios, respectively. Except for gender, there was no significant difference in the respondents’ characteristics across the three health scenarios. The mean age of the respondents was approximately 40 years. The respondents also had an average household income of 15,500 baht/month.

Table 2 displays the number of discounted QALYs gained for each health condition scenario. By using the TTO method, the mean discounted QALY gained ranged from 0.96 (for the mild allergen scenario) to 4.07 (for the quadriplegia scenario). Similarly, the mean QALY gained using the VAS method ranged from 1.08 (for the mild allergen scenario) to 3.00 (for the quadriplegia scenario). For each scenario,
the number of discounted QALY’s gained from more severe health states was higher than the number gained from mild and moderate health states.

Table 3 shows the WTP value for each situation. For all health scenarios, the WTP values for treatment (WTP_{treatment}) were significantly higher than those for prevention (WTP_{prevention}). In general, the WTP value was highest in the paralysis scenario, followed sequentially by blindness and allergies. The mean WTP_{treatment} ranged from 31,000 baht for mild allergies to 165,600 baht for quadriplegia. For each scenario, the mean WTP_{treatment} for the severe health states was higher than that for the mild and moderate health states. A similar pattern was found for prevention. The mean WTP_{prevention} ranged from 12,000 baht for mild allergies to 31,000 baht for bilateral blindness. Furthermore, the percentage of unwillingness to pay ranged from 0.8% to 2.3% in the treatment situation, while it ranged from 1.5% to 5.6% in the prevention situation.

Table 4 displays the WTP/QALY values estimated from different health state scenarios and utility elicitation methods.

Discussion

For the treatment situation, the mean WTP/QALY value ranged from approximately 59,000 to 285,000 baht (3578–17,283 US$ PPP). With a GDP per capita (in 2008) of approximately 138,000 baht, our values ranged between 0.4–2 times the GDP per capita, which – albeit slightly lower – is still in line with the range of 1–3 times the GDP per capita that the Commission on Macroeconomics and Health recommends. When
comparing our estimate in terms of the GDP per capita with those of previous studies, our results are generally consistent with the other studies. In addition, the findings from this study are consistent with the past allocation decisions of the National List of Essential Medicines (NLEM) committee in Thailand, which set a ceiling threshold of 1 GDP per capita/QALY.

Similar to a previous study, we found that male respondents with a high income and high level of education typically offered high WTP/QALY values. In line with prior knowledge, we also found that the value of a QALY varied depending on the specific health condition under consideration. In our study, the WTP/QALY values were significantly higher in the blindness scenario than in the allergic and paralysis scenarios. This may be due to the fact that respondents had different perceptions and attitudes toward each health condition scenario. Our findings clearly confirmed that QALY values are context-specific and may not be comparable across health conditions.

In fact, the source and nature of health conditions (such as self-inflicted conditions vs those that are unavoidably imposed upon the individual) can influence a QALY value. Nevertheless, it is arguable that the use of a condition’s specific threshold would discriminate against marginalized or stigmatized conditions. At present, we know little about the impact of incorporating conditional health preferences on societal decisions and how to implement their incorporation. This issue clearly deserves more attention, and future research is essential.

We found that, in blindness and paralysis scenarios, WTP/QALY values obtained from the more severe health states were lower than those obtained from the less severe states (ie, bilateral blindness < unilateral blindness; quadriplegia < paraplegia). The “ceiling effect” can explain these findings, as described in a previous study. The ceiling effect occurs when respondents are faced with a large QALY gain; however, because of the respondents’ limited budget, the WTP amount increased less than proportionately with respect to the increase in QALYs. In determining the WTP values in this study, the method of payment offered to the respondents was a lump sum payable within 6 months, which may also have contributed to this limitation.

One unique aspect of our study was our ability to compare the WTP/QALY estimates from both prevention and treatment situations. According to our findings, the value of a QALY varies between treatment and prevention situations. However, in line with a previous study, we revealed that the WTP amount for prevention is less than the WTP amount for treatment. As a result, WTP/QALY values were lower than WTP/QALY values. The fact that prevention and treatment situations involve different contexts (in which people have varying reference points and expectations) can explain this result. According to Kahneman and Tversky’s prospect theory, the preference of an individual is related to a reference point. For patients, the preference point for his/her health is a key factor in determining the preference of his/her treatment option. A recent study clearly found that a

Table 5 Factors affecting WTP per QALY (baht, rounded up/down in thousands)

<table>
<thead>
<tr>
<th>DV: Ln (WTP/QALY)</th>
<th>Treatment</th>
<th></th>
<th></th>
<th>Prevention</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>SE (β)</td>
<td>P-value</td>
<td>β</td>
<td>SE (β)</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>10.85</td>
<td>0.10</td>
<td>&lt;0.001</td>
<td>10.08</td>
<td>0.10</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Monthly income (baht)</td>
<td>0.00002</td>
<td>2.6e-06</td>
<td>&lt;0.001</td>
<td>0.00002</td>
<td>2.8e-06</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gender</td>
<td>Male (ref)</td>
<td>-0.32</td>
<td>0.08</td>
<td>&lt;0.001</td>
<td>-0.23</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Scenario</td>
<td>Blindness (ref)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paralysis</td>
<td>-0.89</td>
<td>0.09</td>
<td>&lt;0.001</td>
<td>-0.03</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Allergy</td>
<td>-0.82</td>
<td>0.09</td>
<td>&lt;0.001</td>
<td>-0.98</td>
<td>0.11</td>
</tr>
<tr>
<td>Education</td>
<td>Primary school or lower (ref)</td>
<td>0.29</td>
<td>0.08</td>
<td>0.001</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Secondary school</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Bachelor or higher</td>
<td>0.55</td>
<td>0.16</td>
<td>0.001</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

AIC = 7868.8
AIC = 7984.9

Note: Based on time trade off (TTO) method, as residuals were not normally distributed, so natural log transformations of WTP/QALY values were applied.

Abbreviations: AIC, Akaike’s information criterion; NA, not applicable; QALY, quality-adjusted life year; Ref, reference; SE, standard error; WTP, willingness to pay.
patient may become a risk seeker as his/her health declines.\textsuperscript{34} In the treatment scenario of our study, the QALY value was measured in terms of gaining health value while being in a declined health state. However, for prevention, we measured the QALY value in terms of avoiding the loss of health value while still being in a good health state. Consequently, it should be noted that the range of WTP offered for prevention was lower than the range of WTP offered for treatment. Given a situation in which societal values for WTP/QALY\textsubscript{prevention} are lower than WTP/QALY\textsubscript{treatment}, it is necessary that policy makers and other stakeholders pay closer attention to knowledge provisions and the employment of an effective strategy for increasing prevention awareness among the population. Our findings support the opinion that one universal ceiling threshold should not be applied for the resource allocation of all types of intervention.\textsuperscript{7,10,14,21}

As in previous studies,\textsuperscript{10,12,14,19,22} 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. Nevertheless, based on our preliminary testing, this did not become a major problem in our study. During the development of the questionnaire for each hypothetical health scenario, we included two levels of severity to permit a preliminary test of each respondent’s comprehension. Furthermore, according to our findings, the responses from the utility measure and the WTP measure were consistent with prior expectations. 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 treatment/prevention of the more severe health states was higher than that of the less severe health states.

While societal WTP/QALYs are needed to make informed resource allocation decisions, it should be noted that the methods used in this and other studies\textsuperscript{12,14,17,19,21} to derive the monetary value of a QALY often seek to measure an individual’s perspective. However, Smith (2005) indicates that an individual’s WTP may not be relevant for the societal decisions.\textsuperscript{35} It is likely that an individual’s personal WTP, which reflects the benefits that he/she receives from the goods/services paid for, is different from his/her social WTP, which reflects how much an individual is willing to pay for others in a group that possibly includes him/herself.\textsuperscript{35} Future research focused on measuring the societal WTP/QALY is definitely needed. By asking how much an individual is willing to pay through taxation for the addition of a service to the national health insurance scheme might be another way to measure societal WTP value.\textsuperscript{35} However, this alternative method seems to be inappropriate in Thailand, where only 2.3 million people from the total population of 67 million actually paid tax in 2009. Alternatively, using the person trade-off method,\textsuperscript{36} which requires respondents to take a societal rather than an individual perspective, may also be appropriate to measure societal WTP/QALY values in future studies.

Finally, it should be noted that there are other identifiable limitations in this study. For one, the standard deviation around the WTP values and WTP/QALY values were considerable, indicating a large variation in preferences. Furthermore, the sample in our study only reflected the population from the nine selected provinces. Future studies could employ a more rigorous method of random sampling from the general population. In addition, the applicability of our results to other types of health conditions should be made with caution. The WTP/QALY values derived from this study were based on only three specific conditions.

\textbf{Conclusion}

As health technology assessments are widely used to inform coverage decision making in Thailand, an empirical study examining the value of the QALY is essential for policy decision makers. Our study represents the first step in attempting to quantify the so called “unquantifiable” value of a QALY in Thailand. Despite some limitations, we believe that our empirical study also adds another missing component to studies determining the value of a QALY, especially in prevention situations. According to our findings, the value of a QALY varies widely, depending on the context of health gain. Thus, there should be more than one ceiling threshold for all situations.\textsuperscript{7,8,10,14} The value of other types of QALYs, especially in situations of life extension during terminal illness, deserves further investigation.

\textbf{Acknowledgments}

We are grateful for financial support from the Thai Health Promotion Foundation. We also wish to acknowledge with sincere thanks the contributions of the National Statistical Office, the Provincial Health Offices of the nine selected provinces, Dr Kanlaya Teerawattananon, Khannika Thitiboonsuwon, and Prapag Neramitpitagkulas, as well as all of the respondents in this study.

\textbf{Disclosure}

The authors report no conflicts of interest in this work.
References