ORIGINAL RESEARCH ARTICLE

Willingness to Pay for One Additional Quality Adjusted Life Year: A Population Based Survey from China

Ziping Ye[1,2] · Raela Abduhilil[1] · Jiaxin Huang[1] · Lihua Sun[1]

Abstract

Objective The aim of this study was to estimate the population's willingness to pay (WTP) for an additional quality-adjusted life-year (QALY) in China.

Methods The WTP for an additional QALY (WTP/Q) was estimated using a contingent valuation survey with quota sampling and snowball sampling, using a pre-designed questionnaire with 18 hypothetical scenarios. The change in health state was depicted by the EQ-5D-5L. The questionnaires were completed by telephone and face-to-face interviews. Two-part regression models were used to test validity and how different factors affect WTP/Q.

Results A total of 2008 people participated in this survey and provided 3265 WTP responses for further analysis. The average WTP/Q for the entire sample is 113,120 Renminbi (RMB) (USD 16,884), which is 1.75 times the gross domestic product (GDP) per capita. For the quality-of-life improvement scenarios, the mean WTP/Q is RMB 78,907 (USD 11,777, 1.22 times GDP per capita), which is significantly lower than the life extension scenarios (RMB 177,761, USD 26,531, 2.76 times GDP per capita). Age was found to be negatively related to positive WTP. Educational level was positively related to the probability of reporting positive WTP and the level of WTP/Q. Although the EQ-5D-5L utility scores of respondents did not prove to be statistically significant determinants of WTP/Q, the two dimensions of EQ-5D-5L, pain/discomfort and anxiety/depression, had an impact on WTP/Q. In addition, WTP/Q was higher when the health outcome had a 50% probability of occurring than when the health outcome was 100% certain. WTP/Q was higher when a lower health gain was presented to the respondent.

Conclusion This study provides empirical evidence of the monetary value of an additional QALY from a sample of the Chinese population. In addition, a higher threshold for end-of-life therapies should be considered.

1 Introduction

In China, a series of health policy reforms in recent years have promoted economic evaluation. Since 2019, evidence of economic evaluation has been required by the National Health Insurance Administration for price negotiations. Guidelines for conducting economic evaluations in China were first published in 2011 and are updated as methods evolve [1]. The rules for general economic evaluation decisions suggest that a treatment is considered cost effective

✉ Lihua Sun
sunlh@syphu.edu.cn

Ziping Ye
ziping_ye@126.com

[1] College of Business Administration, Shenyang Pharmaceutical University, Shenyang 110016, China

[2] School of Public Administration, Hainan University, Hainan, China

Key Points for Decision Makers

The estimated willingness to pay for an additional quality-adjusted life-year (WTP/Q) lies in the range of the threshold proposed by the WHO for developing countries (1.75 times GDP per capita versus 1–3 times GDP per capita).

The WTP/Q was found to be determined by age, income, pain/discomfort, and anxiety/depression dimensions of EQ-5D-5L, QALY size and certainty of health outcomes.

Different thresholds for quality-of-life improvement and end-of-life therapies should be considered.

if the incremental cost-effectiveness ratio (ICER) of the cost effectiveness analysis (CEA) falls below a certain cost-effectiveness threshold. Currently, the threshold in China is 1–3 times the gross domestic product (GDP) per capita per disability-adjusted life year (DALY) averted [1], which was once recommended by the World Health Organization (WHO) but no longer applies [2]. A related question is how much health departments in China should pay to improve health, which can be measured in quality-adjusted life-years (QALYs) [3].

The methods for determining this threshold are controversial. In general, two ways of estimating the threshold are proposed, namely the demand-side perspective and the supply-side perspective [4]. The supply-side threshold reflects the health opportunity cost, meaning that whether an intervention is worth its fee depends on whether the quantity of health outcomes it produces is greater than the health outcomes that would have been achieved if a different intervention had been funded [5]. However, the demand-side strategies are consistent with the approach taken in other public sectors in the context of conventional welfare economics, where the monetary value of an additional QALY is measured as the willingness to pay (WTP) per QALY (WTP/Q) of people from society. It was advocated that WTP/Q can help in resource allocation within the health sector as well as between sectors [6].

There is an ongoing debate about appropriate ways of estimating WTP/Q. A major problem raised against these methods is the theoretically unachievable task of linear translation of QALYs to WTP [4]. Two conditions that must be met are scope insensitivity and independence of severity. The former refers to the sensitivity of the WTP to the scope of the good, the latter tests the extent to which the severity of the initial health condition affects the valuation [7]. A whole series of empirical studies have shown that WTP/Q varies according to severity and scope [8–11]. For example, Sund and Svensson pointed out that the standard assumption that the more QALYs the better can only partially apply to the largest improvements [11]. It was also found that larger QALY changes lead to lower WTP/Q estimates. Recently, Kouakou and Poder [12] and Ye et al. [13] published similar reviews with meta-regression analysis on this topic, implying age, income and mode of administration (online or face-to-face interview) have significant effects on WTP/Q. All these published results raise the question of how we determine the WTP/Q that can represent the monetary value of health from the population's point of view. We believe that the essential but probably insufficient way to determine WTP/Q requires a variety of scenarios and a representative sample.

Some empirical research on cost-effective thresholds has been conducted in China. Ochalek and colleagues [14]

calculated the cost per DALY averted as 27,923–52,247 (2017 Renminbi [RMB]) (central estimate RMB 37,446) from the supply-side perspective. Nevertheless, the less common measure of health-related outcome, DALY, was used in this study. Therefore, this threshold could only be used to a limited extent for economic evaluations. Using the published literature, Cai et al. [15] first estimated the value of statistical life and then calculated the cost-effective threshold to be 1.5 times the GDP per capita. Furthermore, Zhao et al [16] proved the feasibility of quantifying the monetary value per QALY with a WTP approach. However, due to the small sample size, they did not succeed in determining a reliable threshold.

The aim of this study was to investigate the monetary value for an additional QALY by eliciting the WTP of citizens for a given health scenario in China. First, we conducted a population-based survey with quota sampling and snowball sampling. In addition, two-part regression models were used to test validity and how different factors affect WTP/Q.

2 Materials and Methods

In general, there are two approaches to determining the WTP/Q, namely the direct and the indirect approach. In the direct approach, a sample of the population is surveyed to indicate their WTP preferences for health gain. A set of hypothetical scenarios is used to assign a monetary value to the improvement in quality of life or life extension. The indirect approach is also known as the Value of Statistical Life (VSL) method, where WTP/Q is derived by estimating the WTP for mortality risk reduction.

Since the indirect method leads to a very high WTP/Q due to the income constraint not being taken into account, we focus on the direct approach in this study. The design of this study is based on previous publications [8–11, 17], including questionnaires for different types of QALY (quality-of-life improvement and life extension), different QALY sizes and different certainty (50%, 75% and 100%) of health outcome.

2.1 Study Design and Sample

The monetary value of a QALY was estimated using a contingent valuation survey (1 October 2020 to 25 January 2021) by telephone (a mobile app 'WeChat') and face to face[1]. As reported in previous research, the response rate of probability sampling was low [18]. Therefore, quota sampling and snowball sampling with quotas based on gender, age, income, place of residence, and education were used in the final survey. The quotas were calculated according to

---

[1] We found that it was difficult to interview the elderly via mobile phones, so we also had to consider face-to-face interviews.

the China Statistical Yearbook 2020 [19]. We first recruited samples[2] that matched our quotas. Then we asked each respondent to recommend two people for this survey and those who matched our quotas were further interviewed.

The interviewers in this survey are undergraduate or master's students. All interviewers were required to attend two training sessions and pass the final test. Respondents were personally recruited by team interviewers. We then interviewed those who met the quotas. All subjects were asked for their full consent to participate in the study and no financial incentives were offered.

2.2 Questionnaire Design

The structure of the questionnaire can be found in Fig. 1. Each questionnaire contains five sections, including the introduction, hypothetical scenarios, the demographic section, and the ending. An example of the questionnaire translated from Chinese can be found in Online Resource 1 in the electronic supplementary material (see ESM_1). First, the introduction contained a brief explanation of this survey. Then, the EQ-5D-5L[3] was used to measure the respondents' health status. The third part consists of two hypothetical health conditions in which we request participants to state the maximum amount they would be willing to pay for the treatment of each hypothetical condition. The demographic section included five demographic questions on age, gender, education, place of residence, and annual family income per person.

The overview of all 18 hypothetical scenarios can be found in Table 1[4]. In total, there are 9 versions of questionnaires, each containing 2 scenarios. This survey combined two ways of demonstrating health improvement, namely life extension (terminal illness) and quality-of-life improvement (mild, moderate and severe disease states).

When asked about mild, moderate and severe WTP, a hypothetical scenario with description of EQ-5D-5L (the health conditions listed in Table 1) was explained to participants. Without any treatment, they would live XX months with the described health state. After XX months, they would recover completely/with 75% probability/with 50% probability. The terminal illness scenario reflects the assumption that
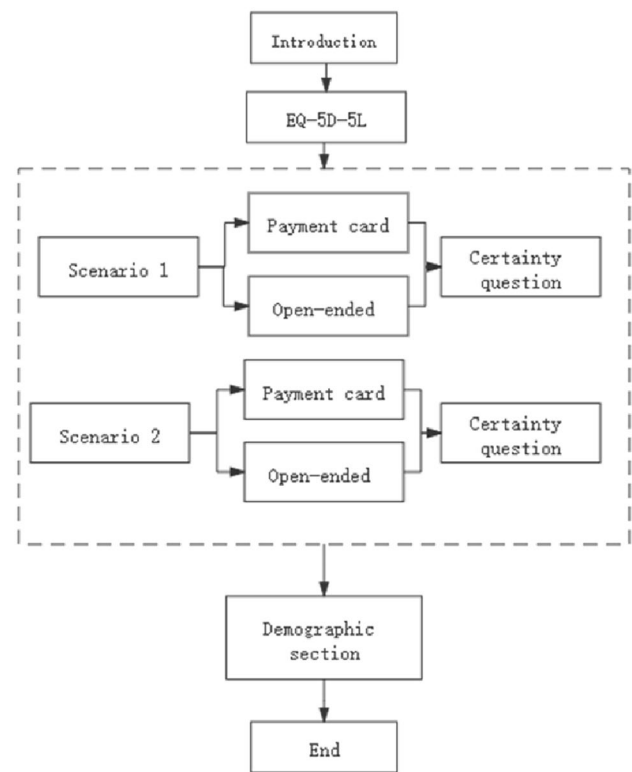
---

[2] The initial stage of the survey is more convenience sampling. We included people who were easy to reach and met our quotas.

[3] The simplified Chinese version of EQ-5D-5L in our research is approved by the EuroQol Group.

[4] To better reflect reality, QALY gains of 0.1 and 0.2 QALY with 100% certainty of health outcomes were set for mild scenarios. For moderate scenarios, QALY gains were set at 0.2 (more frequent) and 0.4 QALY, with 100%, 75% and 50% certainty of health outcomes. For severe scenarios and terminal illness, QALY gains were set at 0.2, 0.4 and 0.8 (less frequent) QALY with 100%, 75% and 50% certainty of health outcomes. In total, we created 18 scenarios.

Fig. 1 Questionnaire structure

participants have a terminal illness and have life length of 3 months (EQ-5D-5L description: 44332). A newly developed treatment could have a 100%/75%/50% chance of extending life expectancy in this severe health state by XX months. For each hypothetical health state, WTP was measured by respondents' willingness to purchase the treatment.

The concept of probability was explained directly, followed by a simple analogy. For example: "...you have a 75% chance of getting perfectly well immediately if you receive the new treatment, which means that in a similar situation, out of four people taking the medicine, three will get perfectly well immediately..."

The WTP payment was described as the amount of out-of-pocket costs for the purchase of an adopted therapy. Participants were asked if they would pay for the therapy. Those who said "no" were then asked to give their reasons. If the answer was "yes', the participant was asked to indicate the maximum amount he or she was willing to pay out of pocket using a payment card. In this study, respondents were offered a payment card with total amounts ordered from low to high. The payment card had the following categories: RMB 3200 (5% of Chinese GDP per capita, USD 478), RMB 6450 (10% of Chinese GDP per capita, USD 963), RMB 12,900 (20% of Chinese GDP per capita, USD 1925), RMB 25,800 (40% of Chinese GDP per capita, USD 3851), RMB 51,600 (80% of Chinese

Table 1 Description of different types of scenarios

Code	QALY gain[§]	Question type	Change of health (EQ-5D-5L)	Period (months)	Certainty of health outcome (%)	Version
1	0.104	Mild	12122 → 11111	8	100	A
2	0.195	Mild	12122 → 11111	15	100	B
3	0.198	Moderate	23332 → 11111	5	100	C
4	0.397	Moderate	23332 → 11111	10	100	D
5	0.174	Moderate	23332 → 11111	6.5	75	E
6	0.178	Moderate	23332 → 11111	10	50	F
7	0.198	Severe	44332 → 11111	3	100	G
8	0.396	Severe	44332 → 11111	6	100	H
9	0.792	Severe	44332 → 11111	12	100	I
10	0.164	Severe	44332 → 11111	4	75	A
11	0.178	Severe	44332 → 11111	6	50	B
12	0.208	Terminal illness	44332, 3–15 mo	15	100	C
13	0.416	Terminal illness	44332, 3–27 mo	27	100	D
14	0.832	Terminal illness	44332, 3–51 mo	51	100	E
15	0.187	Terminal illness	44332, 3–27 mo	27	50	F
16	0.173	Terminal illness	44332, 3–19 mo	19	75	G
17	0.374	Terminal illness	44332, 3–51 mo	51	50	H
18	0.345	Terminal illness	44332, 3–35 mo	35	75	I

[§]We first decided that 0.1, 0.2, 0.4 and 0.8 quality-adjusted life-year (QALY) gains should be the usual health-related outcomes of treatment in the real world. Accordingly, we calculated the time period of the health outcome. Since the period should be a whole number, we then calculated the actual QALY gains

GDP per capita, USD 7701), RMB 77,400 (120% of Chinese GDP per capita, USD 11,064), RMB 103,200 (160% of Chinese GDP per capita, USD 15,403), RMB 206,400 (320% of Chinese GDP per capita, USD 30,806) and RMB 309,600 (480% of Chinese GDP per capita, USD 46,209). Those who could not provide their payment card WTP were asked an open-ended question. To increase the chance of getting reliable estimates, a follow-up question was asked to clarify whether they were really sure about their WTP, with the options: very sure; sure; neutral; not sure; not sure at all. We also specified the following conditions to each respondent to clarify the assumed situation: therapy is not covered by statutory health insurance; loss of income due to the disease does not have to be taken into account.

2.3 Data Analysis

2.3.1 Zero Willingness to Pay (WTP), Protest Values and Irrational Values

A zero value means that the respondent is not willing to pay anything to purchase a hypothetical intervention. If the respondent refuses to state his or her maximum WTP, the result of the WTP survey (often the zero value) is called a protest value. Irrational values are those where respondents were willing to pay a higher amount for smaller health

outcomes than for larger health outcomes. Since protest values and irrational values cannot reflect the true WTP of respondents, they are not included in the main results but in the sensitivity analysis. The definition of protest zero value in this study has three components: first, when respondents refused to state any reason for zero WTP; second, when the respondents refused to pay anything for both WTP questions; third, when the respondents clearly indicated that the intervention should be reimbursed by health insurance, according to the literature [20]

2.3.2 Calculation of WTP per Quality-Adjusted Life-Year (QALY)

In this study, three different probabilities (100%/75%/50%) for the health outcome were considered. According to prospect theory, people tend to assign a large value to small chances and a small value to large chances. If this weighting of probabilities is not corrected, this can lead to incorrect estimates of WTP/Q [10]. Therefore, it is necessary to adjust the probability. In this study, Kahneman and Tversky's [21] estimating equation was used to adjust the probability when the probability of the health outcome occurring is < 100%. The parameter equation is as follows, where $\gamma = 0.674$ [22]:
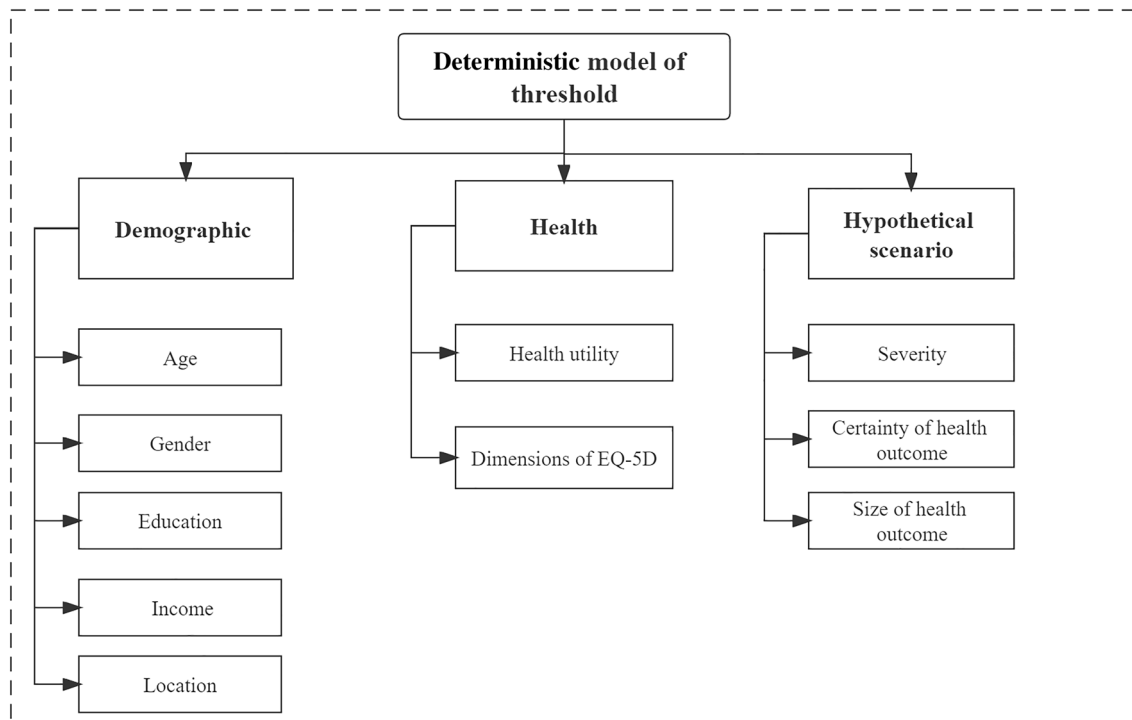
Fig. 2 Determinants of threshold

$$\omega(p) = \frac{p^\gamma}{\left[p^\gamma + (1-p)^\gamma\right]^{1/\gamma}}.$$

The disaggregated method was used in this study because it takes into account the heterogeneity of preferences and the marginal rate of substitution of individual WTP/Q between health and money [23, 24]. This means that first the individual WTP/Q was calculated and subsequently the mean value of WTP/Q was estimated. The descriptive statistics (mean, SD, median, interquartile range, minimum, maximum) of the WTP/Q values were calculated for the whole sample and subgroups. We also conducted a sensitivity analysis as follows: first, we included all protest values and irrational values; second, to reduce the effect on means of extreme values and zero values, we excluded all 0 values and the top 1% of maximum WTP/Q values [25]; furthermore, we removed the WTP/Q responses if the answer to the follow-up certainty question was "not sure" or "not at all sure."

2.3.3 Determinants of WTP per QALY Gain

In this study, an attempt was made to build a multiple determinant model of thresholds, which included three parts: socio-demographic characteristics, health status of respondents, and aspects related to hypothetical scenarios, as shown in Fig. 2. Considering the zero WTP values, a two-part regression model was conducted. In the first part, the probit method was used to estimate the deterministic of the probability of a yes response to the WTP questions; in the second part, the generalized linear model (GLM) with gamma distribution and log function was adopted to estimate the individual's WTP/Q.[5] All samples contained two responses from each person, hence we allowed for clustering at the individual level. We first performed univariate regression for all variables in Online Resource 2 (see ESM_2). Statistically significant variables in univariate regressions and factors mentioned in previous research were included in the multivariate regression [12, 13]. We also conducted subgroup regression as well as sensitivity analyses for the whole sample and the improving quality of life and extending life length subgroups.

---

[5] In order to choose an appropriate variance function for the GLMs, we performed a modified Park test, which indicated gamma distribution. As for the log link, this has the advantage of focusing on differences between groups of participants with respect to arithmetic rather than geometric means.

Table 2 Demographic characteristics of survey respondents

Demographic characteristics	Survey ($N = 2008$)	Quotas (%)
Gender, n (%)		
Male	961 (48)	50
Female	1047 (52)	50
Age, n (%)		
18–34 y	815 (40.6)	35.3
35–54 y	851 (42.4)	47.05
55–65 y	342 (17)	17.65
Education, n (%)		
Primary school or less	445 (22)	25
Secondary school	478 (23)	40
High school	680 (25)	17
College or more	413 (21)	18
Location, n (%)		
Urban	1309 (65)	60
Rural	699 (35)	40
Annual household income (RMB) per person, n (%)		
≤ 12,000	300 (15)	20
12,001–24,000	463 (23)	20
24,001–36,000	536 (27)	20
36,001–60,000	428 (21)	20
> 60,000	281 (14)	20

3 Results

3.1 Demographic Characteristics of Respondents

Table 2 shows the demographic characteristics of the respondents. A total of 2008 individuals were included in this survey, of which 87.5% were interviewed by telephone, and 12.5% were interviewed face to face. In this survey, 52% of the respondents are female, which is consistent with the general Chinese population from the China Statistical Yearbook 2020 [19]. The respondents are between 18 and 65 years old. In the sample, 18- to 34-year-olds are slightly overrepresented (40.6% compared with 35.3%), and 35- to 54-year-olds are slightly underrepresented (42.4% compared with 47.05%). In terms of educational background, 445 (22%) of the respondents reported that they had attended primary school or less, which is in line with the general Chinese population. However, those with only a junior school degree are underrepresented while those with a diploma of high school or higher are overrepresented. Sixty-five percent of respondents live in urban areas, which is similar to the general population (60%). Middle-income groups are slightly overrepresented, whereas annual income groups of less than RMB 12,000 (USD 1791) and more than RMB 60,000 (USD 8955) are underrepresented.

## 3.2 Descriptive Statistics of WTP and WTP per QALY (WTP/Q)

Of the 4016 WTP responses in this survey, there are 662 (16.5%) zero responses and 270 (6.7%) responses are considered protest zeros according to the aforementioned definition. In addition, 481 (12%) WTP responses are classified as irrational responses. A total of 3265 WTP responses were used for further analysis, excluding protest responses and irrational responses. Specifically, 2135 (65%) WTP responses are related to the quality-of-life improvement scenarios (mild 495 [15%], moderate 737 [23%] and severe 903 [28%] disease states). The rest of the responses are regarding life extension scenarios (1130, 35%).

Figure 3 is the distribution of payment card selections. As can be seen from Fig. 3, more respondents chose the lower values while fewer chose higher WTP values. Considering the relatively even distribution of WTP responses to RMB 3200 (USD 448), RMB 6450 (USD 963), and RMB 12,900 (USD 1925), it is reasonable to assume that the bias of starting point bias and range in this research had little impact on the WTP survey.

Table 3 demonstrates the descriptive statistics of WTP/Q of the whole sample and subgroups. The average WTP/Q for the whole sample is RMB 113,120 (USD 16,884) (range from RMB 108,802 to RMB 129,788 in the sensitivity analysis), which equals to 1.75 times GDP per capita, while the median is RMB 36,236 (USD 5408), which equals to 0.65 times GDP per capita. For quality-of-life improvement scenarios, the mean WTP/Q is RMB 78,907 (USD 11,777, 1.22 times GDP per capita). However, the mean WTP/Q of life extension scenarios is significantly higher than that of quality-of-life improvement scenarios, at RMB 177,761 (USD 26,531, 2.76 times GDP per capita).

3.3 Determinants of WTP/Q

According to the results of the univariate regressions as well as the literature theory, the following variables were included in the final two-part regression models: gender, age, education, place of residence, income, EQ-5D-5L utility, the two dimensions of EQ-5D-5L (pain/discomfort and anxiety/depression), the size of the QALY, aspects related to hypothetical scenarios (the quality-of-life improvement or life extension scenarios) and certainty of health outcomes. However, insignificant variables were omitted from Table 3 (whole sample) and Table 4 (subgroups). The results of the sensitivity analysis can be found in Online Resource 3 (see ESM_3).
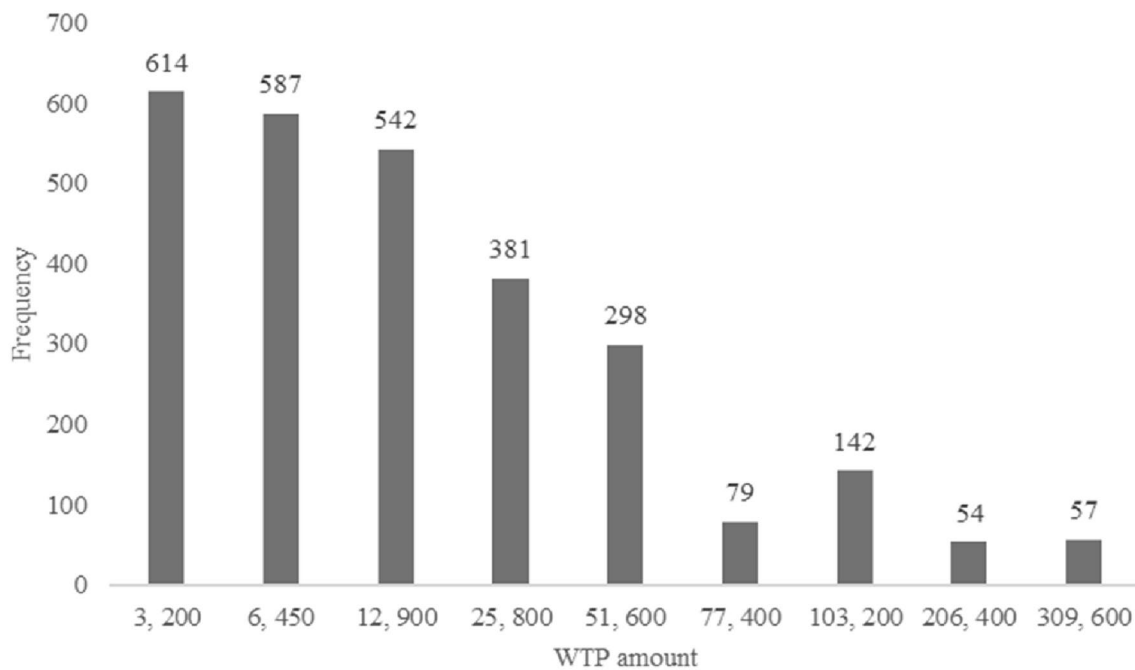
Fig 3 Distribution of payment card selections

3.3.1 Two-Part Regression for the Whole Sample

Table 3 displays results from the two-part regressions for the whole sample. The first column represents the coefficients from the probit model (the first part of the model) that estimates the probability of reporting any positive WTP amount. The third column represents the coefficients from the second part, the GLM model (second part of the model) estimating the relationship among those who reported a positive WTP amount. The model predicted a mean WTP/Q of RMB 113,209 (range from 108,819 to 128,915 in the sensitivity analysis) which was similar to the descriptive mean displayed in Table 5.

In the first part, the results show that higher education and higher income are associated with a higher probability of positive WTP. Compared with individuals without depression, respondents with moderate depression have a lower probability of positive WTP. A larger QALY value increases the probability of a positive WTP. WTP questions related to life extension were also less likely to have a positive WTP. In addition, the probability of a positive WTP was higher when health outcome had a probability of 50% when compared with a probability of 75%.

The second part of the regression shows the factors influencing the magnitude of WTP/Q. Age was associated with a lower WTP/Q. In addition, WTP/Q was higher when respondents had higher education and household income. In addition, respondents with slight or severe pain/discomfort had a lower WTP/Q than respondents who did not have pain/discomfort. QALY size was positively related to WTP/Q. Compared with the scenarios of improving quality of life, the life extension scenarios had a much higher WTP/Q, which is inconsistent with descriptive statistics. Furthermore, WTP/Q was higher when the health outcome had a 50% probability of occurring than when the health outcome was 100% certain.

3.3.2 Two-Part Regression for Subgroups

Subgroup modelling demonstrated the different determinants of WTP/Q in terms of quality-of-life improvement and life extension, which is displayed in Table 4. The mean estimates from the two-part regression models were RMB 75,116 (72,853–82,371) for quality-of-life improvement and RMB 189,374 (175,621–227,388) for life extension.

In the first part, the results show that income is associated with a higher probability of positive WTP for quality-of-life improvement, but not for life extension. In contrast to the results for the whole sample, the probability of positive WTP was lower for 100% of the possibility of a health outcome regarding life extension than for 50% of the possibility. The second part of the regression showed that age was negatively associated with WTP/Q for quality of life improvement but not for life extension. Furthermore, certainty of health outcome is negatively related to WTP/Q only for the life extension scenarios.

Table 3 Results of two-part regression models for the whole sample

	Probit		GLM	
	Coef	SE	Coef	SE
Age			− 0.007*	0.003
Education (vs primary school or less)				
Secondary school	**0.17***	0.072	**0.182***	0.082
High school	0.192	0.102	**0.372****	0.112
College or more	**1.025****	0.151	**0.543****	0.149
Annual household income (RMB) per person (vs < 12,000)				
12,001–24,000	0.023	0.089	0.191	0.139
24,001–36,000	0.064	0.088	**0.355****	0.125
36,001–60,000	**0.261****	0.098	**0.602****	0.122
>60,000	**0.261***	0.121	**1.017***	0.142
Pain/discomfort (vs no problems)				
Slight	− 0.088	0.076	**− 0.307***	0.084
Moderate	− 0.342	0.192	− 0.364	0.309
Severe	− 0.362	0.439	− 0.512	0.629
Very severe	− 0.043	0.623	**− 1.455****	0.467
Anxiety/depression (vs no depression)				
Slight depression	− 0.113	0.069	0.118	0.088
Moderate	**− 0.386****	0.138	0.248	0.234
Severe	− 0.141	0.459	− 0.05	0.467
Very severe	− 0.978	0.659	0.237	0.666
QALY	**0.984****	0.211	**− 1.375****	0.166
Extending life length (vs improving quality of life)				
Extending life length	**− 0.445****	0.07	**1.052****	0.051
Certainty (vs 50%)				
75%	**0.225***	0.095	0.044	0.075
100%	− 0.117	0.076	**− 0.2****	0.068
N	3265		2876	
Log pseudo-likelihood	− 1085		− 35,990	
WTP/Q estimate			113,209	
Pseudo R^2	0.0904			
AIC			25	

*Indicates that there are statistically significant differences at the 5% level; **indicates that there are statistically significant differences at the 1% level; insignificant variables were omitted in this table.

AIC Akaike information criterion, *GLM* generalized linear model, *QALY* quality-adjusted life-year, *RMB* Renminbi, *SE* standard error, *WTP/Q* willingness to pay per QALY

Table 4 Results of two-part regression models for subgroups

	Improving quality of life				Extending life length			
	Probit		GLM		Probit		GLM	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Age	−0.005	0.004	−0.01*	0.004	−0.006	0.005	−0.001	0.004
Education (vs primary school or less)								
Secondary school	0.166	0.097	0.182	0.103	0.204	0.129	0.19	0.098
High school	0.334*	0.142	0.366**	0.135	0.009	0.172	0.355*	0.14
College or more	1.085**	0.197	0.552**	0.188	1.005**	0.242	0.535**	0.142
Annual household income (RMB) per person (vs < 12,000)								
12,001–24,000	0.072	0.118	0.079	0.159	−0.021	0.165	0.408**	0.146
24,001–36,000	0.109	0.115	0.267	0.148	0.037	0.162	0.577**	0.136
36,001–60,000	0.47**	0.137	0.5**	0.142	0.039	0.171	0.865**	0.142
> 60,000	0.204	0.151	0.96**	0.169	0.386	0.223	1.19**	0.149
Pain/discomfort (vs no problems)								
Slight	−0.157	0.103	−0.329**	0.103	−0.033	0.129	−0.255*	0.101
Moderate	−0.499	0.273	−0.446	0.383	−0.191	0.333	−0.249	0.38
Severe	−0.769	0.655	−0.44	0.708	−0.019	0.767	−0.318	0.761
Very severe			−1.815**	0.495				
Anxiety/depression (vs no depression)								
Slight	−0.234**	0.09	0.054	0.114	0.086	0.121	0.258**	0.091
Moderate	−0.241	0.211	0.211	0.291	−0.583*	0.246	0.226	0.194
Severe			−0.411	0.618	−0.659	0.595	0.455	0.574
Very severe	−0.821	0.802	0.346	0.669			−2.022**	0.248
QALY	1.409**	0.436	−1.402**	0.29	0.952**	0.251	−1.428**	0.161
Certainty (vs 50%)								
75%	0.269*	0.129	0.176	0.094	0.242	0.15	−0.207*	0.102
100%	0.012	0.108	−0.161	0.087	−0.265*	0.115	−0.21*	0.09
N	2123		1927		1128		949	
Log pseudo-likelihood			−23548		−449		−12422	
WTP/Q estimate			75,116				189,374	
Pseudo R²	0.095				0.093			
AIC			25				26	

*Indicates that there are statistically significant differences at the 5% level; ** indicates that there are statistically significant differences at the 1% level; insignificant variables were omitted in this table.

AIC Akaike information criterion, GLM generalized linear model, QALY quality-adjusted life year, RMB Renminbi, SE standard error, WTP/Q willingness to pay per QALY

Table 5 Descriptive statistics of willingness to pay for one additional quality-adjusted life-year

	Whole sample		Quality-of-life improvement		Life extension	
	Base case	Sensitivity analysis	Base case	Sensitivity analysis	Base case	Sensitivity analysis
N	3265	(2738, 4016)	2135	(1842, 2666)	1130	(895, 1335)
Mean	113,120	(108,802, 129,788)	78,907	(76,821, 88,523)	177,761	(166,677, 214,038)
Standard deviation	223,362		182,231		273,993	
Median	36,236	(33,077, 62,019)	21,446	(32,576, 33,077)	74,566	(68,984, 124,038)
Minimum	0		0		0	
Maximum	2,976,923		2,976,923		2,673,797	
25th quantile	16,288	(16,162, 29,750)	16,288	(16,161, 17,978)	17,213	(15,385, 62,019)
75th quantile	124,038	(124,038, 132,308)	72,472	(72,472, 74,138)	248,077	(248,076, 275,936)

4 Discussion

In this paper, we report the results of a survey to estimate WTP for a QALY gain among the Chinese population. A total of 2008 individuals were included in this survey. The mean WTP/Q for the entire sample was RMB 113,120 (USD 16,884, 1.75 times GDP per capita). For the quality-of-life improvement scenarios, the mean WTP/Q is 78,907 RMB (USD 11,777, 1.22 times GDP per capita), for the life extension scenario, the average WTP/Q is 177,761 RMB (USD 26,531, 2.76 times GDP per capita). Two-part regression models indicated that the determinants of WTP/Q were age, income, pain/discomfort, and anxiety/depression dimensions of EQ-5D-5L, QALY size and certainty of health outcomes.

It is crucial to ensure the validity and practical usefulness of our empirical results. Testing theoretical validity is the most popular validity test for WTP surveys because it is comparatively easy to perform. Theoretically, a positive effect of income on WTP/Q would argue for the theoretical validity of WTP research [8, 9, 13]. Moreover, it was assumed that worse health state scenario [26–28] and smaller QALY gain [4, 26, 30, 31] should lead to a higher WTP/Q. Our regressions confirmed previous results and proved the theoretical validity of our estimates: WTP/Q was higher when respondents had higher incomes, when analyzing both the whole sample and subgroups; the mean WTP/Q value for life extension is almost 2.25 times that for quality-of-life improvement; negative effect of QALY gain on WTP/Q was observed. Furthermore, a comparison of our results with those of a very comparable study would provide further evidence of the validity of this research. Cai et al. [15] estimated the monetary value of a QALY based on the value of statistical life to be 1.5 times the GDP per capita in China, equivalent to almost RMB 100,000 (USD 14,952). Song and Lee [29] found that the average WTP/Q in South Korea was about RMB 145,000 (USD 21,642). Both studies agreed with our estimates.

It is important for policy makers in health economics to know the public's WTP value for a QALY. Our study provides empirical evidence on WTP/Q that questions the appropriateness of the WHO's recommendation. According to the systematic review published by Butt and colleagues [30], the most commonly used threshold was three times GDP per capita. As we have shown, estimates of WTP/Q can vary considerably depending on the specific subgroups and hypothetical scenarios. The large discrepancy between the quality-of-life improvement and life extension scenarios leads to the question of whether we should adopt a higher threshold for end-of-life treatment. For example, the National Institute for Health and Care Excellence (NICE) standard threshold is £20,000 to £30,000, but for end-of-life

interventions the threshold is £50,000 [31]. Whether and how this heterogeneity should be reflected in health care decision making in low- and middle-income countries is a question beyond the scope of this paper, but clearly a subject for future research.

Respondents who refused to pay were assigned a WTP of zero. Zero responses were quite common in all WTP surveys. For instance, in the European Value of a QALY (EuroQol) study [25], about 30% of the total responses were zero WTP. Similarly, in the study by Shafie and colleagues [26], almost half of the responses were zero. In this research, 16.5% of responses were zero, of which almost a third are defined as protest values. Twelve percent of WTP responses are classified as irrational responses because a lower WTP value was given for a higher benefit. In general, protesters and irrational respondents are not considered to have accepted or understood the scenario and they should therefore be excluded for the main analysis. However, we included all protesters and irrational respondents in the sensitivity analysis, which showed limited impact on the final estimates.

There are a number of limitations of this study that need to be mentioned. First, the final sample was broadly representative of the 18–65 years age group, but with a significant overrepresentation of middle-income groups and an underrepresentation of low-education groups. Both income and education have been shown to be positively related to WTP/Q, which is consistent with previous findings. Therefore, our study may overestimate the true WTP/Q of the population. Furthermore, a snowballing method was used to recruit the samples, which was a very effective method given the very limited funds. Nevertheless, there is a high probability that unobserved variables influence the representativeness of the sample, leading to problems with external validity. Third, the mode of this survey might influence our results. According to recent publications [12, 13], surveys conducted over the internet or telephone are more likely to have a significantly higher value for WTP/Q than face-to-face surveys. This is what we call mode bias in surveys. Almost 90% of our survey was conducted over the phone, the rest were face-to-face interviews. However, it is still not clear which mode better represents the true value. Fourth, this is a cross-sectional study limited to one survey at one point in time, conducted under the influence of the coronavirus pandemic; longitudinal time-series data would be required for robust WTP/Q estimation. Finally, an important limitation is the hypothetical nature of all WTP surveys [32]. It can be difficult for individuals to imagine an assumed scenario that differs from conditions they have experienced.

5 Conclusion

Our results suggest that the mean WTP/Q for the entire sample was RMB 113,120 (USD 16,884, 1.75 times GDP per capita). For the quality-of-life improvement scenarios, the mean WTP/Q was RMB 78,907 (USD 11,777, 1.22 times GDP per capita), for the life extension scenario, the average WTP/Q was RMB 177,761 (USD 26,531, 2.76 times GDP per capita). WTP/Q was found to be determined by age, income, pain/discomfort, and anxiety/depression dimensions of EQ-5D-5L, QALY size and certainty of health outcomes. We recommend that future researchers seek to explore the WTP/Q of major diseases from multiple perspectives. Furthermore, different thresholds for quality-of-life improvement and end-of-life therapies should be considered.

Declarations

References

1. China Center for Health Economic Research. China guidelines for pharmacoeconomic evaluations, 2017.

2. Bertram MY, et al. Cost–effectiveness thresholds: pros and cons. Bull World Health Organ. 2016;94(12):925.

3. Wilson HS. Methods for the economic evaluation of health care programmes. J Epidemiol Community Health. 1987;41(4):355–6.

4. Baker R, Chilton S, Donaldson C, et al. Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE. Health Econ Policy Law. 2011;6(04):435–47. https://doi.org/10.1017/S1744133111000181.

5. Woods B, Revill P, Sculpher M, et al. Country-level cost-effectiveness thresholds: initial estimates and the need for further research. Value Health. 2016;19(8):929–35. https://doi.org/10.1016/j.jval.2016.02.017.

6. Robinson A, Gyrd-Hansen D, Bacon P, Baker R, Pennington M, Donaldson C, Team E. Estimating a WTP-based value of a QALY: the 'chained' approach. Soc Sci Med. 2013. https://doi.org/10.1016/j.socscimed.2013.05.013.

7. Nielsen JS, Gyrd-Hansen D, Kjær T. Sample restrictions and the elicitation of a constant willingness to pay per quality adjusted life year. Health Econ. 2021;30(5):923–31.

8. Bobinac A, van Exel J, Rutten FFH, et al. The value of a QALY: individual willingness to pay for health gains under risk. Pharmacoeconomics. 2014;32(1):75–86.

9. Bobinac A, van Exel NJA, Rutten FFH, et al. GET MORE, PAY MORE? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. J Health Econ. 2012;31(1):158–68.

10. Bobinac A, Van Exel NJA, Rutten FFH, et al. Valuing QALY gains by applying a societal perspective. Health Econ. 2013;22(10):1272–81.

11. Sund B, Svensson M. Estimating a constant WTP for a QALY—a mission impossible? Eur J Health Econ. 2018;19(6):871–80.

12. Kouakou CRC, Poder TG. Willingness to pay for a quality-adjusted life year: a systematic review with meta-regression. Eur J Health Econ. 2022;23(2):277–99.

13. Ye Z, Ma J, Liu F, et al. A systematic review and meta-regression of studies eliciting willingness-to-pay per quality-adjusted life year in the general population. Expert Rev Pharmacoecon Outcomes Res. 2021. https://doi.org/10.1080/14737167.2021.1878881.

14. Ochalek J, Wang H, Gu Y, et al. Informing a cost-effectiveness threshold for health technology assessment in China: a marginal productivity approach. Pharmacoeconomics. 2020. https://doi.org/10.1007/s40273-020-00954-y.

15. Cai D, Shi S, Jiang S, et al. Estimation of the cost-effective threshold of a quality-adjusted life year in China based on the value of statistical life. Eur J Health Econ. 2021. https://doi.org/10.1007/s10198-021-01384-z.

16. Zhao FL, Yue M, Yang H, et al. Willingness to pay per quality-adjusted life year: is one threshold enough for decision-making? Med Care. 2011;49(3):267–72. https://doi.org/10.1097/MLR.0b013e31820192cd.

17. Gloria M, Thavorncharoensap M, Chaikledkaew U, et al. A systematic review of demand-side methods of estimating the societal monetary value of health gain. Value Health. 2021;24(10):1423–34.

18. Ye Z, Liu F, Ma J, et al. Comparing the monetary value of a quality-adjusted life year from the payment card and the open-ended format. Cost Effect Resour Alloc. 2021. https://doi.org/10.1186/s12962-021-00298-0.

19. National Bureau of Statistics of the People's Republic of China. China Statistical Yearbook. China Statistics Press. 2020.

20. Greg L. Market models, protest bids, and outliers in contingent valuation. J Water Resour Plan Manag. 1994;120(1):121–9.

21. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain. 1992;5(4):297–323.

22. Bleichrodt H, Pinto Prades JL. A parameter-free elicitation of the probability weighting function in medical decision analysis. Manag Sci. 2000;46:1485–96.

23. Rashidian A, Rasekh H, Foroughi M, Olyaeemanesh A. Monetary value of quality-adjusted life years (QALY) among patients with cardiovascular disease: a willingness to pay study. Iran J Pharm Res. 2017;16(2):820–30.

24. Gyrd-Hansen D, Kjær T. Disentangling WTP per QALY data: different analytical approaches, different answers. Health Econ. 2012;21(3):222–37. https://doi.org/10.1002/hec.1709.

25. Pennington M, Baker R, Brouwer W, et al. Comparing WTP values of different types of QALY gain elicited from the general public. Health Econ. 2015;24(3):280–93. https://doi.org/10.1002/hec.3018.

26. Shafie AA, Lim YW, Chua GN, et al. Exploring the willingness to pay for a quality-adjusted life-year in the state of Penang, Malaysia. ClinicoEconomics Outcomes Res. 2014. https://doi.org/10.2147/CEOR.S67375.

27. Gyrd-Hansen D. Willingness to pay for a QALY. Pharmacoeconomics. 2005;23(5):423–32. https://doi.org/10.1002/hec.79.

28. Ryen L, Svensson M. The willingness to pay for a quality adjusted life year: a review of the empirical literature. Health Econ. 2015;24(10):1289–301. https://doi.org/10.1002/hec.3085.

29. Song HJ, Lee EK. Evaluation of willingness to pay per quality-adjusted life year for a cure: a contingent valuation method using a scenario-based survey. Medicine. 2018. https://doi.org/10.1097/MD.0000000000012453.

30. Butt T, Liu GG, Kim DD, Neumann PJ. Taking stock of cost-effectiveness analysis of healthcare in China. BMJ Glob Health. 2019;4(3): e001418.

31. Bovenberg J, Penton H, Buyukkaramikli N. 10 years of end-of-life criteria in the United Kingdom. Value Health. 2021;24(5):691–8.

32. Hanemann WM. The economic theory of WTP and WTA. Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries. 1999:42–96.