DETERMINATION OF COST-EFFECTIVENESS THRESHOLD FOR MALAYSIA

by

LIM YEN WEI

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science (Pharmacoeconomics)

October 2015

ACKNOWLEDGEMENT

The successful completion of this study was made possible with the support from a number of persons. First and foremost, I would like to express my deepest gratitude to my main supervisor, Associate Professor Dr Asrul Akmal Shafie, co-supervisor, Professor Dr Mohamed Azmi Ahmad Hassali and co-researcher, Ms Chua Gin Nie for their guidance and instruction throughout this study. They have never given up the hope on me and their constructive feedback, continuous encouragement and patience have definitely made the completion of my study.

I would like to thank the Ministry of Health of Malaysia for approving my study leave which provided me with protected time to concentrate on my study. I would also like to express my appreciation to Universiti Sains Malaysia (USM) for funding me with a research grant (APEX Delivering Excellence 2012, 1002/PFARMASI/910330) throughout the study. This has definitely helped in reducing my financial burden in conducting this research.

I appreciate all the staff members in the School of Pharmaceutical Sciences, USM and my fellow postgraduate students in the Discipline of Social and Administrative Pharmacy, USM for their support throughout my study period. My sincere thanks go to all the members of the research team especially the final year undergraduate students who assisted in data collection during the study period. Their enthusiasm and active participation have definitely smoothed the data collection process. I would also like to dedicate my special thanks to all the international collaborators in this collaboration project for their assistance and involvements in making this project a success.

Last but not least, my deepest gratitude to my mother, my aunty who treats me like her own son and all my family members for their prayers, loves and endless support which rendered me a positive attitude in completing this thesis. Also, thanks to my fiancée for her support and encouragement throughout the whole study period.

TABLE OF CONTENTS

CHAPTER 3 – METHODS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CAD	Canadian Dollar
CBA	Cost-benefit Analysis
CE	Cost-effectiveness
CEA	Cost-effectiveness Analysis
CUA	Cost-utility Analysis
CV	Contingent Valuation
CVM	Contingent Valuation Method
DALY	Disability-adjusted Life-year
EQ-5D-3L	EQ-5D with Three Level Version
EQ-5D-5L	EQ-5D with Five Level Version
EUR	Euro
EuroQoL	European Quality of Life
GBP	Great Britain Pound
GDP	Gross Domestic Product
HRQoL	Health Related Quality of Life
HTA	Health Technology Assessment
HUI	Health Utilities Index
HUI2	HUI Mark II System
HUI3	HUI Mark III System
HYE	Healthy Year Equivalent
ICER	Incremental Cost-effectiveness Ratio
Interviewer-QWBS	Interviewer-administered QWBS
JPY	Japanese Yen
LMIC	Low- and Middle-income Countries
MYR	Malaysian Ringgit
NICE	National Institute for Health and Clinical Excellence
NZD	New Zealand Dollar
QALY	Quality-adjusted Life-year
QWBS	Quality of Well-being Scale

QWBS-SA	Self-administered Version of QWBS
SAVE	Safe-young-life Equivalent
SD	Standard Deviation
SEK	Swedish Krona
SEM	Standard Error of the Mean
SG	Standard Gamble
TTO	Time Trade-off
USD	United States Dollar
USM	Universiti Sains Malaysia
VAS	Visual Analogue Scale
WHO	World Health Organisation
WHO-CHOICE	World Health Organisation’s CHOosing Interventions that are Cost-Effective
WTA	Willingness-to-accept
WTP	Willingness-to-pay

LIST OF APPENDICES

LIST OF PUBLICATIONS

1. **Lim Y.W.**, Shafie A.A., Hassali M.A., Chua G.N. (2013). Exploring The Feasibility of The Instrument Used In Determination of Willingness-To-Pay Per Quality-Adjusted Life-Year In Malaysia. *Value In Health*, 16(7): A596 – A597.
2. **Lim Y.W.**, Shafie A.A., Chua G.N., Hassali M.A. (2014). Determination of Cost-effectiveness Threshold for Malaysia. *Value In Health*, 17(7): A438.
3. Shafie A.A., **Lim Y.W**., Chua G.N., Hassali M.A. (2014). Exploring the willingness to pay for a quality-adjusted life-year in the state of Penang, Malaysia. *ClinicoEconomics and Outcomes Research,* 6: 473 – 481.

LIST OF CONFERENCE PRESENTATIONS

1. 'Exploring Willingness-to-pay for A Quality-Adjusted Life-Year in Malaysia' at *1st Quality Use of Medicine (QUM) 2013 Conference*, Subang, Selangor, Malaysia (Poster Presentation).
2. 'Exploring the Feasibility of the Instrument Used In Determination of WTP/QALY In Malaysia' at *10th Annual Conference of Health Technology Assessment International (HTAi) 2013*, Seoul, South Korea (Podium Presentation).
3. 'Exploring the Feasibility of the Instrument Used In Determination of Willingness-to-Pay Per Quality-Adjusted Life-Year In Malaysia' at *16th International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual European Congress 2013*, Dublin, Ireland (Poster Presentation).
4. 'Exploring Willingness-to-pay per Quality-Adjusted Life-Year in Malaysia' at *5th Social Science Postgraduate National Seminar 2013*, Universiti Sains Malaysia, Penang, Malaysia (Podium Presentation).
5. 'Determination of Cost-Effectiveness Threshold for Malaysia' at *2nd Pharmacoeconomics and Outcome Research Conference 2014*, Kuala Lumpur, Malaysia (Poster Presentation).
6. 'Determination of Incremental Cost-Effectiveness Ratio Threshold for Malaysia' at *3rd HTAsiaLink Conference 2014*, Beijing, China (Podium Presentation).
7. 'Exploring the Feasibility of the Instrument Used In Determination of Cost-effectiveness Threshold In Malaysia' at *8th National Conference for Clinical Research (NCCR) 2014*, Kuching, Sarawak, Malaysia (Poster Presentation).

8. ‘Determination of the Cost-Effectiveness Threshold In Malaysia’ at *8th National Conference for Clinical Research (NCCR) 2014*, Kuching, Sarawak, Malaysia (Podium and Poster Presentation).
9. ‘Determination of Cost-Effectiveness Threshold for Malaysia’ at *17th International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual European Congress 2014*, Amsterdam, The Netherlands.

LIST OF AWARDS AND ACHIEVEMENTS

1. Successfully secured APEX Delivering Excellence 2012 Research Grant (1002/PFARMASI/910330) with the amount of MYR 120,000.
2. First Prize for Poster Presentation at *1st Quality Use of Medicine (QUM) 2013 Conference*, Subang, Selangor, Malaysia.
3. 2013 ISPOR International Student Travel Grant Award for *16th International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual European Congress 2013*, Dublin, Ireland with the amount of USD 1,300.
4. First Prize in Young Investigator Award for Podium Presentation at *8th National Conference for Clinical Research (NCCR) 2014*, Kuching, Sarawak, Malaysia.

PENENTUAN NILAI AMBANG KOS BERKESAN DI MALAYSIA

ABSTRAK

Dalam sistem penjagaan kesihatan, penentuan untuk kos berkesan bagi teknologi penjagaan kesihatan adalah sukar terutamanya dalam keadaan di mana terdapat alternatif-alternatif yang lebih mahal tetapi lebih berkesan. Dalam keadaan ini, satu kriteria luaran dalam bentuk nilai ambang kos berkesan ataupun WTP/QALY diperlukan untuk menentukan kos keberkesanan intervensi penjagaan kesihatan. Namun, amalan kini di Malaysia dalam penentuan kos keberkesanan untuk teknologi penjagaan kesihatan baru dilakukan tanpa satu nilai ambang kos berkesan yang jelas. Jadi, tujuan utama kajian ini adalah untuk menentukan nilai ambang kos berkesan bagi intervensi penjagaan kesihatan di Malaysia. Satu kajian *cross-sectional*, penilaian kontingen telah dijalankan dengan menggunakan teknik persampelan rawak kelompok berstrata di Pulau Pinang, Kedah, Selangor dan Wilayah Persekutuan Kuala Lumpur. Responden berumur antara 20 – 60 tahun yang memahami sama ada Bahasa Inggeris atau Bahasa Melayu ditemuduga secara bersemuka. Mereka ditanya tentang latar belakang sosio-ekonomik, mutu kehidupan dan kesanggupan untuk membayar satu senario hipotetikal (rawatan, lanjutan nyawa dalam penyakit terminal dan situasi penyelamatan nyawa dengan tiga tahap keterukan penyakit dan dua tahap QALY – 0.2 QALY dan 0.4 QALY). Nisbah purata nilai WTP bagi satu QALY diterokai dengan menggunakan kaedah bukan parametrik *Turnbull* dan model parametrik *interval regression*. Model parametrik *interval regression* juga digunakan untuk menganalisa faktor-faktor yang mempengaruhi nilai ambang kos berkesan. Seribu tiga belas responden telah ditemuduga semasa kaji

selidik dijalankan. Nilai ambang kos berkesan yang diterokai dari kaedah bukan parametrik *Turnbull* ialah antara Ringgit Malaysia 12,810 – 22,840 (~ Dolar Amerika Syarikat 4,000 – 7,000) manakala nilai ini dianggar antara Ringgit Malaysia 19,929 – 28,470 (~ Dolar Amerika Syarikat 6,200 – 8,900) dengan menggunakan model parametrik *interval regression*. Faktor-faktor utama yang mempengaruhi nilai ambang kos berkesan ialah tahap pendidikan, anggaran pendapatan bulanan isi rumah dan penerangan senario keadaan kesihatan. Nilai ambang kos berkesan yang didapati dalam kajian ini ialah antara Ringgit Malaysia 19,929 – 28,470. Keputusan yang diperoleh menyokong bahawa nilai QALY tidak boleh dirujuk sebagai satu nilai tunggal. Nilai ambang kos berkesan yang dianggar untuk Malaysia adalah didapati lebih rendah daripada nilai ambang kos berkesan dengan satu hingga tiga kali keluaran dalam negara kasar per kapita yang disyorkan oleh pertubuhan kesihatan sedunia.

DETERMINATION OF COST-EFFECTIVENESS THRESHOLD FOR MALAYSIA

ABSTRACT

In healthcare system, decisions on the cost-effectiveness (CE) of healthcare technologies are difficult especially when alternatives are more expensive but more effective. In this situation, an external criterion in the form of CE threshold or willingness-to-pay for a quality-adjusted life-year (WTP/QALY) is necessary to decide on the CE of healthcare interventions. Nevertheless, current practice in Malaysia on coverage decisions of new healthcare technologies is made without an explicit CE threshold. Thus, this study aimed to determine a CE threshold value for healthcare interventions in Malaysia. A cross-sectional, contingent valuation study was conducted using stratified multistage cluster random sampling technique in Penang, Kedah, Selangor and Kuala Lumpur Federal Territory. Respondents aged between 20 – 60 years old who can understand either English or Malay language were interviewed face-to-face. They were asked for the socioeconomic background, quality of life and their WTP for a hypothetical scenario (treatment, extended life in terminal illness and life saving situations with three severities and two QALY gained levels – 0.2 QALY and 0.4 QALY). The mean ratio of the amount of WTP for an additional QALY gained was explored by non-parametric Turnbull method and parametric interval regression model. Parametric interval regression model was also used to analyse the factors that affect the CE threshold. One thousand thirteen respondents were interviewed during the survey. The CE threshold explored from non-parametric Turnbull method ranged from MYR 12,810 – 22,840 (~

USD 4,000 – 7,000) whereas it was estimated to range between MYR 19,929 – 28,470 (~ USD 6,200 – 8,900) using parametric interval regression model. Key factors that affect the CE threshold were education level, estimated monthly household income and the description of health state scenarios.

The cost-effectiveness threshold found in this study was reported as MYR 19,929 – 28,470. The findings support that there is no single value of a QALY. The CE threshold estimated for Malaysia was found to be lower than the threshold value of one to three times the gross domestic product per capita recommended by the World Health Organisation.

CHAPTER 1

INTRODUCTION

1.1 The Importance of Economic Evaluation in Healthcare

Economic is defined as a study on how men and society end up choosing, with or without the use of money, to employ scarce productive resources that could have alternative uses, to produce various commodities and distribute them for consumption, now or in the future, among various people and groups in society. It analyses the costs and benefits of improving patterns of resource allocation (Sameulson and Nordhaus, 1998, Shiell *et al.*, 2002).

For a long time, it is noted that health expenditures have been seen rising far more rapidly than the national income generated in most countries. In response, many countries are having dilemma in getting sufficient funds to ensure universal access to all health interventions and services (Chisholm and Evans, 2007). Under the situation of limited budget and the pervasive scarcity of resources available to supply health demands, economic considerations are getting an increasingly prominent role in the evaluation of healthcare interventions.

Economic evaluation, is an economic tool that can be usefully employed to address the question of how to achieve the highest possible overall level of population health for the available resources (Chisholm and Evans, 2007). In this context, the basic task of economic evaluation is to identify, measure, value, and compare the costs and consequences of the alternatives being considered which in turn is crucial to provide information for policymakers in determining the best forgone choice between the competing alternatives (Berger *et al.*, 2003, Drummond *et al.*, 2005). Decision making in health is inherently value-laden, where policymakers need sound evidence of the likely costs and benefits of their decisions, articulated through a prism of societal values. A good decision making in healthcare setting should always consider the additional costs of an intervention as well as the health consequences of those additional costs that are reflected from the perspective of individual patient level to a societal level. Economic evaluation serves as a robust methodology for determining the costs and benefits of an intervention (NICE International, 2014).

1.2 Basic Concept of Economic Evaluation

Economic evaluation is defined as the comparative analysis of alternative courses of action in terms of both their costs and consequences (Drummond *et al.*, 2005). However, there are also some cases where the evaluations deal only with the analysis of either costs or consequences of the competing alternatives. Some even deal with the analysis of cost and consequence within a single programme without comparing to other alternatives. In these cases, they are classified as partial economic evaluation. In full economic evaluation, both the costs and consequences of the alternatives are examined. Three major analytic techniques are commonly used in healthcare. They are cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and cost-utility analysis (CUA). All these three types of analyses are similar in how they measure costs but differing in the way that the consequences of healthcare programmes are measured and valued (Drummond *et al.*, 2005).

CBA is a form of economic evaluation that enumerates and compares the net costs of a healthcare intervention with the benefits that arise as a consequence of applying that intervention. Both the net costs and the consequences of the health intervention are expressed in a commensurate unit, often money (Shiell *et al.*, 2002, Berger *et al.*, 2003).

CEA is also a form of full economic evaluation that compares both the costs and consequences of health care programmes, where the consequences of the programmes are measured in the same common units – natural units related to the clinical objective of the programmes, such as life-years gained, points of blood pressure reduction and so on (Berger *et al.*, 2003). The outcomes in CEA are single, programme specific, and unvalued. Typically the results are expressed as cost per effectiveness ratio (Drummond *et al.*, 2005).

The final analytical technique, CUA, is another form of economic evaluation which has a lot of similarities to CEA except for the outcomes that are measured in the units of utility or preference, often quality-adjusted life-year (QALY) gained, or possibly some variant, like disability-adjusted life-year (DALY) or healthy year equivalent (HYE) (Shiell *et al.*, 2002, Berger *et al.*, 2003, Drummond *et al.*, 2005). Unlike CEA, the outcomes in CUA may be single or multiple, are generic as opposed to programme specific, and incorporate the notion of value. In CUA, the results are normally presented in the terms of cost per QALY gained (Drummond *et al.*, 2005).

1.2.1 Moving Towards Cost-Utility Analysis

CEA is unequivocally the most widely used technique of economic evaluation in the field of health economics. Nevertheless, CUA has increasingly become more popular and its application has been extended to a wide variety of health interventions including pharmaceuticals, surgical procedures and diagnostic imaging (Neumann *et al.*, 2005, Fang *et al.*, 2011). The steady growth of CUA is due to its credibility in overcoming few problems faced when using CEA as a technique in economic evaluation.

First, the outcomes measured in CEA are programme-specific units such as millimetres mercury of blood pressure reduction, cases prevented, life-years gained and so on where different programmes may be designated with different outcome units. In such a condition, CEA can only be used to compare alternatives with similar outcome units but cannot be used to make comparisons across a broad set of interventions with different outcome units (Drummond *et al.*, 2005). In order to be able to compare the different options for the use of common resources, the quantification of health outcomes using a common measurement unit is necessary. CUA was therefore developed to address this shortage. It provides a method through which the various disparate outcomes can be combined into a single composite summary outcome (e.g. QALY gained), which in turn, enable broad comparisons across widely differing programmes (Drummond *et al.*, 2005, Dernovsek *et al.*, 2007).

Second, CEA cannot address the issue of opportunity cost of funding the new programme whereas CUA on the other hand is able to do so (Drummond *et al.*, 2005). The standard outcomes measurement of cost per QALY gained in CUA allows one to determine a level of 'acceptable' cost utility for health care choices: that is, a 'threshold level' of cost per QALY (Berger *et al.*, 2003). Third, in any one programme there is usually more than one outcome of interest. In CEA, the outcomes presented are single, programme-specific and cannot be valued. On the contrary, the outcomes in CUA may be single or multiple, are generic as opposed to programme specific, and there are values attached to the outcomes where the more important outcomes are weighted more heavily (Drummond *et al.*, 2005).

In addition, CUA has the advantage over other cost-effectiveness methodologies because it incorporates quality of life measure. Health related quality of life (HRQoL) represents a crucial measure of therapeutic effectiveness especially when two alternatives differ in their effect on quality of life as well as survival and it has become an important element in the economic evaluation. CUA is an adaptation of CEA which measures the effect of treatment on both the quantitative (length of life or mortality) and qualitative (quality of life or morbidity) aspect of health (Berger *et al.*, 2003, Dernovsek *et al.*, 2007). CUA is now becoming more common and can be considered as the "gold standard" methodology among other cost-effectiveness techniques (Berger *et al.*, 2003). Therefore, many health authorities worldwide and experts in the field of health economics and outcomes research have recommended the use of CUA in evaluating the cost-effectiveness of the health care programmes (Fang *et al.*, 2011).

1.2.2 The concept of Quality-adjusted Life-year

Many cost-effectiveness (or clinical) studies express health outcomes in term of programme-specific measures such as number of cases avoided or life year gained without taking into the consideration of quality adjustment. Although these measures are useful in comparing the effects of particular treatment but they do not permit the comparison across widely differing diseases or programmes. For this reason, the concept of QALY has been introduced as a convenient metric for such purposes (Neumann and Greenberg, 2009, Smith *et al.*, 2009).

QALY is a universal health outcome measure used in CUA which is able to capture simultaneously both the changes in quantity of life (mortality) and the changes in quality of life (morbidity), and integrate these into a single measure. It is applicable to all individuals and all diseases, thereby enabling broad comparisons across differing programmes or diseases (Berger *et al.*, 2003, Smith *et al.*, 2009).

QALYs are the life years adjusted by a preference-based quality weight (health-related quality of life). The quality weights for QALYs must be based on preferences, anchored on perfect health with a score of 1.0 and being dead has a score of 0.0, and measured on an interval scale. In order to qualify as preference-based instrument, the quality weights must represent the preferences of individuals for the relevant health states as measured with appropriate preference measurement instruments such as standard gambling, time trade-off, Health Utilities Index, EuroQoL EQ-5D, quality of well-being instrument or visual analogue scale (VAS). The more desirable (or more preferred) health states receive greater weight and will be favoured in the analysis compared to those that are worse or less desirable. The preference exist in QALY satisfies the von Neumann-Morgenstern axioms where an individual should be indifferent between two risky prospects (Berger *et al.*, 2003, Drummond *et al.*, 2005, Neumann and Greenberg, 2009).

The scale of QALY weights must contain two anchor points on the interval scale – perfect health and death. These two points can be given any two arbitrary values as long as death has a smaller value than perfect health. For instance, a value of zero may represent death while perfect health may be represented by a value of one. In fact, there is no well-defined upper end or bottom of the scale. Nevertheless, the pair of values zero and one is conventionally assigned to represent death and perfect health respectively in the scale of QALY weights (Drummond *et al.*, 2005).

The QALY value is calculated by the product of the survival time in a particular health state and its quality-adjustment weight (the utility value) (Drummond *et al.*, 2005).

As illustrated in Figure 1.1 (Torrance, 1996), without intervention, an individual's HRQoL would deteriorate according to the lower path and the person would die at time Death 1. With the intervention, the individual would live longer with the HRQoL deteriorate more slowly and would die at time Death 2. The area between the two curves is the QALY gained by the intervention. From the diagram shown, Part A is the amount of QALY gained due to quality improvement (the gain in health related quality of life) and Part B is the amount of QALY gained due to quantity improvement (the amount of life extension). Simply taking, QALY is calculated by the area under the curve, where the duration of the health state in years is multiplied by the quality weight for the health state. (Berger *et al.*, 2003, Drummond *et al.*, 2005).

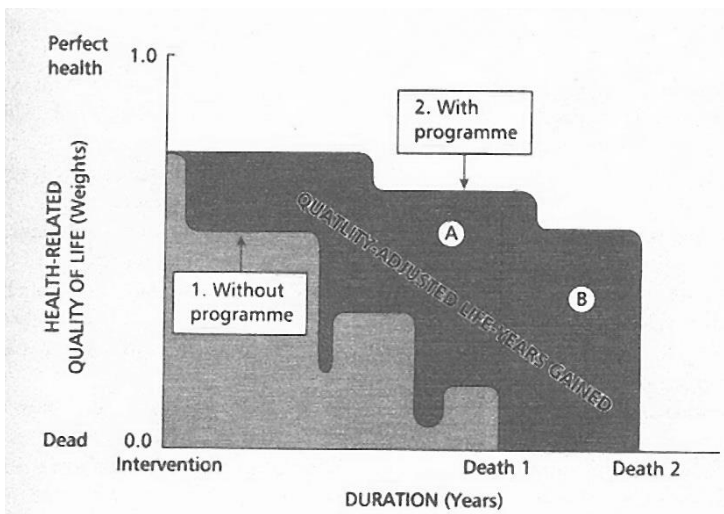

Figure 1.1: Quality-adjusted Life-years Gained from an Intervention
Source: Adapted from Torrance, 1996

The concept of QALY has been used universally but is not without controversy. Arguments ranged from those questioning that the QALY approach is needlessly complex and should be replaced by simpler disaggregated measures to those who think that the QALY concept is overly simplistic and should be replaced by more complex methods (Mehrez and Gafni, 1989, Mehrez and Gafni, 1992). Several alternatives to QALY have therefore been suggested. For example, the World Bank uses DALY and others suggest the use of HYE and saved-young-life equivalent (SAVE) (Berger *et al.*, 2003, Drummond *et al.*, 2005).

Conceptually, the DALY approach is similar to QALY approach but an important difference between them is that DALY uses an age-weighting function that values life years differently depending on the age of disease onset (Whitehead and Ali, 2010). The value choices built into the DALY (the age weight, the discount rate and the disability weights) have a major influence on the rankings of programmes, and yet these value choices are arbitrary and are far from transparency (Drummond *et al.*, 2005). For HYE, it truly reflects a person's utility function over the lifetime through the measurement of the preferences over the entire path of health states rather than a single health state in QALY approach. In such a condition, it is theoretical attractive but more difficult to be implemented in practice (Drummond *et al.*, 2005, Whitehead and Ali, 2010). Compared to QALY approach, the SAVE approach is less willing to take mortality risks to improve quality of life because it appears to give more emphasis to quantity of life, and less to quality of life (Drummond *et al.*, 2005).

As an overall comparison, the lack of a simple better measure as an alternative makes the QALY an indispensable tool (Smith *et al.*, 2009). The QALY approach provides an imperfect but nonetheless useful proxy as a measure of value to inform reimbursement decisions in healthcare and most cost-utility analyses (CUAs) are conducted using cost per QALY as the unit of measurement (Berger *et al.*, 2003, Neumann and Greenberg, 2009).

1.2.3 Incremental Cost-effectiveness Ratio

In both CEA and CUA, the results are usually expressed in the terms of incremental cost-effectiveness ratio (ICER). ICER is a ratio calculated by dividing the difference in cost (incremental cost) between two alternatives to the difference in effects (incremental effects) of these two alternatives (Berger *et al.*, 2003). The formula for the calculation of ICER is shown below:

$$\text{ICER} = \frac{\text{Cost of Programme A} - \text{Cost of Programme B}}{\text{Effect of Programme A} - \text{Effect of Programme B}}$$

In ICER, the numerator (cost) is expressed as monetary unit and the denominator (effect) is expressed in appropriate health units which must be the same units between two interventions. For instance, life-years gained in CEA whereas QALY in CUA. Therefore, ICER may be expressed as cost per life-years gained or cost per QALY. ICER reflects the additional cost per unit of health benefits obtained in switching from one medical intervention to another. It compares medical interventions and provides information for resource allocation. When used in proper context, ICER serves as one of the vital tools in guiding decision making on allocating scarce resources across competing medical interventions (Bambha and Kim, 2004).

1.3 Decision Making Using Cost-effectiveness Plane

Cost-effectiveness (CE) plane is a two-dimensional graphical device used to illustrate the cost-effectiveness of two comparing alternatives (Black, 1990). CE plane is often used to show how decisions can be related to both costs and effects (Briggs and Fenn, 1998). It is now a generally accepted method of presenting the results of CEA and CUA (Sendi and Briggs, 2001). The CE plane is shown in Figure 1.2.

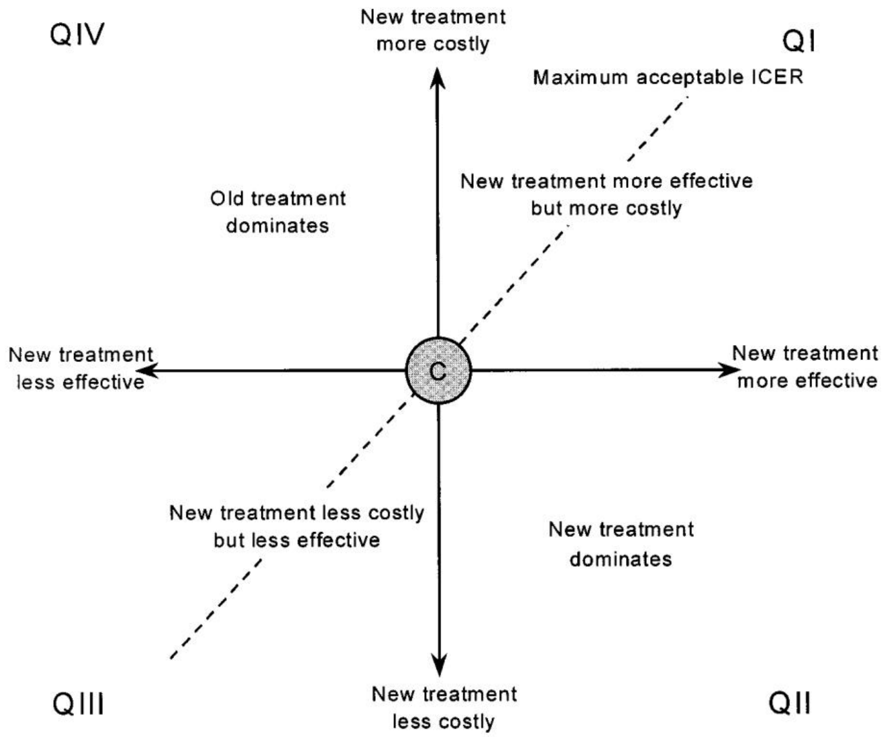

Figure 1.2: The Cost-effectiveness Plane
Source: Adapted from Briggs, 1998

In Figure 1.2, the horizontal axis represents the differences in effectiveness while the vertical axis represents the differences in cost. The CE plane is divided into four quadrants indicating four possible situations in relation to the additional costs and additional health outcome effects of a new intervention compared to the controlled intervention (the origin) (Briggs and Fenn, 1998). In CE plane, any intervention characterised by a certain cost and effectiveness may be represented by a point. A point will be allocated to the right side on the CE plane if an intervention is more effective and vice versa. Similarly, the more costly the intervention is, the higher the point will be seen on CE plane (Bambha and Kim, 2004).

If a new intervention is simultaneously cheaper and more effective than the other alternative, it will be plotted in quadrant II of CE plane. In this situation, new intervention is clearly the treatment of choice because it is the most cost-effective and is said to dominate or sometimes it is called 'dominant'. In contrast, the least cost-effective strategies fall on quadrant IV in which the intervention has an increased cost and decreased effectiveness compared to the alternative. In such a condition, the new intervention is dominated by the old intervention and should not be adopted. However, if a new intervention is both more costly and more effective (quadrant I), or less costly and less effective (quadrant III), then the decision is no longer clear. In this context, a judgement must be made concerning whether the difference in costs resulting from a switch in therapy is justified by the difference in effectiveness that such a switch would bring about (Briggs and Fenn, 1998, Bambha and Kim, 2004).

1.3.1 What is Cost-effectiveness Threshold?

ICER alone does not allow policymakers to judge whether a healthcare intervention represents good value for money. To draw conclusions on the CE of healthcare interventions, the derived ICER (from CEA and CUA) needs to be compared to a reference value, which is CE threshold value or sometimes it is also known as ICER threshold or ceiling threshold (Cleemput *et al.*, 2008, Thavorncharoensap *et al.*, 2013). The ceiling threshold can be represented by the dashed line on the CE plane in Figure 1.2. Any intervention with calculated ICER appears above the CE threshold line would be deemed not cost-effective while it would be considered cost-effective if it falls below the line, indicating some monetary value of a QALY. Apparently, this CE threshold value serves as the maximum acceptable value for ICER (Briggs and Fenn, 1998). In other words, CE threshold value is a boundary for the last intervention that would still be financed from a given fixed budget (Cleemput *et al.*, 2008). CE threshold also reflects the opportunity cost, where it represents the additional cost that needs to be imposed to forgo one QALY worth of health through displacement (Claxton *et al.*, 2013).

1.4 Application of Cost-effectiveness Threshold in Economic Evaluation

Economic evaluations (particularly CEA and CUA) have increasingly been used as an important tool to support efficient resource allocation especially in the resource-constraint society in healthcare sector. Inevitably, application of thresholds for CE ratios are appeared as a tool for the reimbursement of pharmaceuticals or formulary listing (Eichler *et al.*, 2004).

As mentioned before, CE threshold would serve as a marginal benchmark (the last chosen point) for any new medical intervention to be deemed cost-effective. With this, it means that if a new medical intervention has a lower cost per QALY value (or ICER) than this marginal benchmark, it is likely to be accepted by healthcare policy makers (Weinstein, 2008). Taking an example of a study done by Freeman et al, they found that the ICER of high dose dabigatran compared with warfarin for stroke prevention in atrial fibrillation was USD 45,372 and concluded that high dose dabigatran was deemed to be more cost-effective compared to adjusted-dose warfarin, based on the CE threshold of USD 50,000 per QALY in the United States (Freeman *et al.*, 2011). Apparently, as seen from this study, CE threshold expressed in the term of cost per QALY would serve as a robust tool in the cost-effectiveness decision making.

1.5 Problem Statement

In real world, outcomes from economic evaluations alone are not sufficient to inform policymakers in decision making. It is unclear to what degree CE is used to guide coverage decision. CEA cannot be considered as a proper decision making tool because it would lack a systematic and universally recognisable decision criterion (Bobinac *et al.*, 2010).

Decision making is a complex process, where effectiveness and CE are not the only two of many considerations in making policy choices. No matter how explicit and openly the decisions are taken, it is somehow desirable to have a preference-based value that citizens place on the gains in health and life expectancy that can be achieved with the new medical interventions. In such circumstances, estimating a country specific CE threshold is necessary (Eichler *et al.*, 2004, Ahlert *et al.*, 2013).

A number of countries have explicitly stated their own threshold values such as the United Kingdom, Ireland and so on. In Malaysia, however, there is lack of an explicit Malaysian threshold value of QALY. Current practice on covering decisions of new healthcare technologies is made without a transparent decision criterion. In this situation, it will contribute to more room for arbitrariness and 'ad hoc' consideration in decision making process. This will affect the identification of true opportunity cost of a new medical intervention, which, in turn, will impose inefficiency and inconsistency in decision making, and ruin the sustainability of healthcare funding system (Eichler *et al.*, 2004, Donaldson *et al.*, 2010). Therefore, establishing a Malaysian CE threshold expressed in the term of cost per QALY is vital as it evolves as a solid criterion in decision making process.

1.6 Study Objectives

The general objective of this study was to determine a CE threshold value for healthcare interventions in Malaysia. Specifically, it was aimed to find out the amount of willingness-to-pay (WTP) for a QALY in Malaysian society and the factors that affect the amount of WTP for a QALY.

CHAPTER 2

LITERATURE REVIEW

2.1 Cost-effectiveness Threshold Used Globally

For many years, clinical evidence is often the only evidence required for deciding funding of healthcare interventions or drugs reimbursement. However, many countries are now considering the cost of drug as part of the important criteria in decision making due to the finite resources available in healthcare sector (Barnieh *et al.*, 2014). In conjunction with this, economic evaluations are adopted in the decision making process, where they assist in determining the relative value for money of the interventions (Donaldson *et al.*, 2010).

Results of economic evaluations (especially CEA and CUA) are usually summarised in ICER. ICER represents the incremental cost per incremental outcome of one intervention compared to another. In this context, a commonly used measure of health outcome is QALY. By default, ICER is expressed as incremental cost per QALY gained (Bobinac *et al.*, 2010).

To decide on the CE of medical interventions, external criterion in the form of CE threshold need to be applied to the ICER. Interpreting whether a derived ICER is acceptable requires the use of CE threshold (Claxton *et al.*, 2013). Hence, the CE threshold, also presented as WTP/QALY gained value, is vital for decision making using economic evaluation. Johannesson and Meltzer, 1998, had argued the importance of explicating a threshold value. They claimed that without CE threshold, CEA cannot be considered a proper decision making tool because it would lack a systematic and universally recognisable decision criterion (Johannesson and Meltzer, 1998).

In conjunction with this, three countries in Europe, namely the United Kingdom (England and Wales), Ireland and the Slovak Republic have explicitly stated a CE threshold each for funding or informing decision about listing a drug on formulary. (Barnieh *et al.*, 2014). The National Institute for Health and Clinical Excellence (NICE) in the United Kingdom has set a threshold value of Great Britain Pound (GBP) 20,000 – 30,000 per QALY gained for England and Wales, whereas Ireland's threshold is Euro (EUR) 20,000 per QALY gained (Appleby *et al.*, 2007, Shiroiwa *et al.*, 2010, Barnieh *et al.*, 2014). In the Slovak Republic, they use a threshold ranging from EUR 18,000 – 26,500 per QALY for drug reimbursements (Barnieh *et al.*, 2014).

The Netherlands, Sweden and France also apply threshold values in the practice, although these values are generally not explicitly acknowledged by the policymakers in these countries (Donaldson *et al.*, 2010). In the United States, a threshold of United States Dollar (USD) 50,000 – 100,000 per QALY is widely used as a benchmark for assessing the CE of an intervention (Grosse, 2008, Shiroiwa *et al.*, 2010). In addition, a threshold of Canadian Dollar (CAD) 20,000 – 100,000 was recommended in Canada since two decades ago (Laupacis *et al.*, 1992).

Besides Western countries, Asia-Pacific countries like Australia, New Zealand and Japan have also mentioned CE threshold for the use in their countries, despite the fact that the values are arbitrary stated (Simoens, 2009, Shiroiwa *et al.*, 2013). Nevertheless, as economic evidence is increasingly required in healthcare decision making, CE threshold plays an important role in the process.

In the recently done European value of a quality-adjusted life-year (EuroVaQ) study in Europe, it is clearly noted that CE threshold has become an indispensable tool in economic evaluation to have a more consistent and transparent decision making process. Along with this, HTAsiaLink, a network of health technology assessment (HTA) organisations in Asia has embarked on first collaborative research on determining the CE threshold across 4 countries in Asia Pacific region namely Korea, Japan, Malaysia, and Thailand. Table 2.1 shows the summary of the CE threshold used in the public domain in some countries, although some are generally not explicitly acknowledged by the healthcare decision making body within each country.

Table 2.1: Summary of the Currently Used Cost-effectiveness Thresholds

Country	CE Threshold in Local Currency	Reference
Australia	AUD 42,000 – 76,000 per life year	Simoens, 2009
Canada	CAD 20,000 – 100,000 per QALY	Laupacis *et al*., 1992
France	EUR 50,000 per QALY	Donaldson *et al*., 2010
Ireland	EUR 20,000 per QALY	Barnieh *et al*., 2014
Japan	JPY 5 – 6 million per QALY	Shiroiwa *et al*., 2013
Netherlands	EUR 20,000 per QALY	Donaldson *et al*., 2010
New Zealand	NZD 3,000 – 15,000 per QALY	PHARMAC, 2007
Slovak Republic	EUR 18,000 – 26,500 per QALY	Barnieh *et al*., 2014
Sweden	SEK 400,000 – 655,000 per QALY	Donaldson *et al*., 2010
United Kingdom (England and Wales)	GBP 20,000 – 30,000 per QALY	Appleby *et al*., 2007
United States	USD 50,000 – 100,000 per QALY	Grosse, 2008 & Shiroiwa *et al*., 2010

Abbreviations: AUD, Australian Dollar; CAD, Canadian Dollar; EUR, Euro; JPY, Japanese Yen; NZD, New Zealand Dollar; SEK, Swedish Krona; GBP, Great Britain Pound; USD, United States Dollar; QALY, Quality-Adjusted Life-Year.

2.2 Conceptual Tools in the Determination of Cost-effectiveness Threshold

In view of the necessity to have CE threshold in decision making, various approaches have been used to define CE threshold, including league table approach, human capital approach and preference-elicitation approach (Shillcut *et al.*, 2009).

In league table approach, the interventions are distributed in a table from the most to the least cost effective. Then, CE threshold is revealed as the CE ratio of the last intervention to be approved as funds are exhausted. The drawback in this approach is the requirement to evaluate all the potential programmes to determine the last intervention in which this theoretical ideal is unachievable (Shillcut *et al.*, 2009, Newall *et al.*, 2014).

By defining a person's life according to the average income of individuals within the society, a human capital approach is implied. In this approach, individuals are entitled to their 'fare share' of a nation's wealth (Shillcut *et al.*, 2009, Newall *et al.*, 2014). Defining CE threshold based on the economic activity of individuals is gaining recognition in economic evaluations in low- and middle-income countries (LMIC). The Commission for Macroeconomics and Health applied per capita income and the World Health Organisation (WHO) applied gross domestic product (GDP) initiated from World Health Organisation's CHOosing Interventions that are Cost-Effective (WHO-CHOICE) project as their thresholds (Shillcut *et al.*, 2009). In WHO-CHOICE threshold, interventions with CE ratio below the GDP per capita are deemed 'very cost-effective' and those with CE ratio below three times the per capita GDP are deemed 'cost-effective' (Newall *et al.*, 2014).

The reference to the WHO-CHOICE threshold has been widely used for economic evaluation in LMIC (Newall *et al.*, 2014). However, the acceptance of WHO-recommended threshold value remains controversial because it depends on the robustness of the assumptions behind the estimation of the regional GDP per capita. The use of such a generalised threshold value may not be entirely relevant in every country, as different countries may have distinct socio-demographic and disease burdens despite having similar GDPs per capita. By using GDP per capita as an indicator, a nation's average wealth does not necessarily indicate the state of wealth of every member of society. For example, the incomes of some individuals in rural areas may remain low, although they live in high-income countries, and may not fairly represent a nation's wealth (Decision Support Unit of National Institute for Health and Clinical Excellence., 2007). Moreover, the threshold values estimated in this human capital approach may not reflect the budget available or the preferences of society (Newall *et al.*, 2014).

In preference-elicitation approach, CE thresholds are established either by the evaluation of real world decisions (revealed preference) or hypothetical scenarios (stated preference) of individuals within the society for the outcome of interest (Shillcut *et al.*, 2009, Newall *et al.*, 2014). In this approach, the threshold values estimated may help inform on questions of how much to spend on health but may be of limited use where budget are constrained. Because the threshold are not linked to the available budget, their use can lead to continual growth in healthcare spending as more interventions emerge that meet the cost-effectiveness criteria (Birch and Gafni, 2006). To date, there is no accepted standard method to estimate the CE threshold value (Eichler *et al.*, 2004). However, Thavorncharoensap et al. mentioned that the value of a QALY estimated in terms of a society's WTP per QALY should be adopted as the ceiling threshold (Thavorncharoensap *et al.*, 2013). Hence, in this study, preference-elicitation approached was chosen in the exploration of CE threshold.

The main outline in the determination of CE threshold is by eliciting the monetary value of a QALY. Estimating the value of a QALY is complicated and it involves some extent of methodological challenges. Due to the absence of typical 'buy and sell' transaction of a marketed good when placing a monetary value of a QALY, actual cost or sales information is seldom available. Having this situation, it is necessary to have an accurate valuation technique to elicit the economic value of a QALY that public places on it (Carson, 2000).

A variety of valuation techniques have been developed by economists to value non-marketed goods in consistent with the valuation of marketed goods. These valuation techniques are usually based on either revealed preference approach (observing behaviour toward some marketed goods in connection with the non-marketed goods of interest) or stated preference approach (using ‘stated’ information concerning preferences for non-marketed goods) (Carson, 2000, Carson and Hamemann, 2005).

In environmental economics literatures, examples of revealed preference approach are hedonic pricing and the household production function approach while the stated preferences approach is frequently referred to contingent valuation (CV) (Carson and Hamemann, 2005). Hedonic pricing is a method used to estimate economic values for environmental services that directly affect market prices. The basis of this approach is that the price of a marketed good is related to its characteristics. Thus, valuation can be done by looking at how people are willing to pay for a good based on its characteristics (Vanslembrouck *et al.*, 2005).

In household production function approach, the households typically produce commodities through combining goods usually purchased in the market place with their own time. This approach is mainly focusing on the demand of commodities as functions of commodity prices, which, depend on goods prices and the household’s technology. As a basis, the price of a unit of commodity is the sum of prices of goods purchased and of the time used per unit of commodity. For instance, in the process of baking a bread, the price of the bread is the sum of the prices of flour, eggs, sugar purchased and of the time used to bake the bread (Becker, 1965, Pollak and Wachter, 1975).

On the other hand, the contingent valuation method (CVM) is a survey method used to measure people’s WTP or willingness-to-accept (WTA) and is one of the most widely used approaches developed for the measurement of the value of non-marketed goods through hypothetical survey questions (Johannesson and Jonsson, 1991, Food and Agriculture Organization Information Division, 2000). CVM is the most popular method used in recent years as it can cover wide range of themes such as measuring project benefits in monetary terms, or assessing social impacts of environmental conservation policies (Fujita *et al.*, 2005).

2.2.1 Contingent Valuation Method

CV is a non-market valuation method used to estimate the value that a person places on a good using stated preference information (Food and Agriculture Organization Information Division, 2000, Hoyos and Mariel, 2010). This method is commonly used to elicit the individual's preference for a non-marketed good where it measures directly a respondent's WTP to obtain a specific good, or WTA to give up a good through the survey instrument (Arrow *et al.*, 1993, Hoyos and Mariel, 2010).

In CV, it creates a hypothetical marketplace for valuation of non-marketed goods in which no actual transactions are made. It has proven useful when implemented alone or jointly with other valuation technique for non-marketed goods (Food and Agriculture Organization Information Division, 2000). It remains as a technique capable of directly eliciting a monetary (Hicksian) measure of welfare (Hoyos and Mariel, 2010).

Using CVM, the values obtained are contingent on the information given. It means that the values are obtained from the respondents in a survey by using stated information concerning preferences for the good (Carson and Hamemann, 2005). In other words, CVM is a way of simulating a missing market where individual expresses his or her valuation for a good, contingent on a certain scenario (Berger *et al.*, 2003).

CV surveys have been used to value a discrete change in the provision of an environmental good, the value associated with substituting one good for another, or the marginal substitution of different attributes of an existing good (Carson and Hamemann, 2005, Hoyos and Mariel, 2010). Now, this technique has also been used for measuring the value of life, or the value of health improvements (Berger *et al.*, 2003).

2.2.1.1 Economic Theory of Contingent Valuation Method

CV is deeply rooted in welfare economics, which is the study of well-being of members of a society as a group. In this context, the sum of social benefits requires the aggregation of individual's benefits, and that an individual is the best judge of his or her own welfare (Berger *et al.*, 2003, Hoyos and Mariel, 2010).

CV applies surveys to measure an economic concept of value (Carson and Hamemann, 2005). To estimate an individual's benefits, CV triggers the measurement of the net change in the income of an individual that relates to a change in the quality and quantity of a non-marketed good. In other words, it links the survey instrument and economic theory, where CV survey provides information to elicit the WTP (or WTA) distribution for a change in a good (Hoyos and Mariel, 2010). From the perspective of WTP, the maximum amount a person would be willing to offer for a good can be elicited. On the other hand, for WTA, the value attributed by CVM is evaluated by the minimum monetary amount required for an individual to forgo some good, or to bear some harm (Martin-Fernandez *et al.*, 2010).

The relationship between WTP and WTA is vital in CEA because the rules of decision are based on the acceptability of the incremental cost per unit of effectiveness. This could be interpreted to mean that this cost would be acceptable or not is depending on the estimated societal WTP for an additional unit of health effect. A WTA/WTP ratio greater than one carries the meaning that the utility perceived by the loss is greater than that perceived by an equivalent gain. Consecutively, this has implications for the CE threshold at which to declare an intervention to be cost-effective, depending on whether it represents an increase in the utility with a cost increase (quadrant I of CE plane) or a loss of utility with lower costs (quadrant III of CE plane) (O'Brien *et al.*, 2002, Martin-Fernandez *et al.*, 2010).

Besides, CV combines economic theory with the utility function. The utility theoretical model provides a basic framework for interpreting the responses to a CV study. As these responses are usually treated as random variables, the economic model needs to incorporate a stochastic component and the WTP/WTA distributions need to be linked to the survey response probability with the assumption that an individual maximises his or her utility (Hoyos and Mariel, 2010). In welfare economics, it seeks to reveal whether the potential change in utility resulting from a change in an economic variable. The welfare implications are usually expressed in terms of a change in an index (the monetary amount), which would need to be taken from or given to an individual to keep the individual's overall level of utility constant (Carson *et al.*, 2001).

2.2.1.2 Survey Design in Contingent Valuation Studies

The design of questionnaire is a key aspect in CVM. To have a high quality CV study, the information described in the questionnaire should be consistent with scientific and expert knowledge. The hypothetical scenario constructed has to be as closely as possible to the real-world situation. In addition, the description of the good under valuation should be understandable and comprehensible to a respondent who might know little or nothing about the good. In CV study, face validity is considered as a desired property, where the information provided in the survey instrument should be clear and accurate, and the trade-off that the respondent is asked to make should be plausible in order to make a decision (Carson and Hamemann, 2005, Hoyos and Mariel, 2010).

In most of the CV surveys, the questionnaires are normally designed with general ("warm-up") questions introducing the purpose of the survey in opening session to make the respondents comfortable with participating in the survey. In second section, a clear and detailed description of the CV scenario and the good to be valued should be provided. This includes the current and baseline situation (status quo) and possible future states of the natural resource in the case of no implementation of the proposed policy, including the institutional context in which the good will be provided and the payment vehicle. Accompanying materials such as charts or photographs may be used to aid with the description of the scenario (Food and Agriculture Organization Information Division, 2000, Carson and Hamemann, 2005).

This is followed by elicitation section or payment question, which queries respondents about their maximum WTP for a commodity or the minimum WTA for giving it up. The next section is the analysis of the understanding and certainty of the answer provided by the respondents. In the last section, a set of questions regarding the