MEASURING PATIENT-REPORTED OUTCOMES IN TUBERCULOSIS

by

NA GUO

M.D., Shandong University, China, 2001
M.P.H., University of Alberta, Edmonton, Canada, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Pharmaceutical Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
September 2010

© Na Guo, 2010

ABSTRACT

Objectives: The primary objectives of this thesis were to: (1) measure health-related quality of life (HRQL) and health state utility values (HSUVs) among patients with active tuberculosis (TB) disease and latent tuberculosis infection (LTBI); (2) investigate the relationship between HRQL and adverse drug reactions (ADR) among active TB patients; (3) quantify patients' preferences for LTBI preventive treatment.

Methods: Two groups of patients were administered questionnaires: (1) Short-Form 36 (SF-36), Health Utility Index (HUI) and a Visual Analog Scale (VAS) were administered to 119 LTBI and 114 active TB patients at baseline and 3 months of their treatment. (2) A discrete choice experiment (DCE) survey was developed and administered among 214 LTBI patients. Conditional logit and latent class analysis were conducted to quantify respondents' preferences toward six treatment attributes (i.e. treatment length, clinic visit frequency, and risk of developing active TB, liver damage, skin rash and fatigue).

Results: The baseline SF-36, HUI-2, HUI-3, Short-Form 6D (SF-6D) and VAS scores from active TB patients were significantly lower than those from LTBI patients. Major ADRs were shown to have significant impacts on active TB patients' HRQL and patients with lower baseline SF-36 scores were more likely to develop ADRs during the treatment. The three health utility instruments (HUI-2, HUI-3, and SF-6D) displayed acceptable construct validity when applying among TB population. However, they did not generate identical HSUV scores for the same individual.

The DCE study results showed that all six attributes significantly influenced respondents' treatment decision and preference estimates were reasonable and consistent with our hypotheses. Substantial preference heterogeneity was observed among respondents. Latent class analysis assigned respondents into three groups and five socio-demographic factors significantly predicted the class assignment (i.e. origin of birth, education, employment, had children or not, and use of over-the-counter medications).

Conclusions: Active TB disease and the treatment associated ADR have substantial impacts on patients' HRQL. HRQL measurements might have the potential to predict patients' treatment outcomes. The DCE technique provides a useful tool of understanding patients' preferences surrounding health care products. This work demonstrates the value and importance of incorporating patient-reported outcome measurements into clinical research and practice.

PREFACE

This thesis incorporates four original papers that have been previously published or submitted for publication in peer reviewed journals. Permissions from the copyright owners have been obtained to include these published materials in this thesis.

The manuscript in Chapter 2 has been published in *Health and Quality of Life Outcomes* under the title "*Measuring health-related quality of life in tuberculosis: a systematic review*". As the first author, the candidate's role was the conception and design of the study, data collection and syntheses, and writing of the manuscript.

The content of Chapter 3 has been published in the journal of *Value in Health* under the title "*Health state utilities in latent and active tuberculosis*". The ethics approval of this study was obtained from the University of British Columbia Behavioral Research Ethics Board and the certificate number is *B04-0048* (APPENDIX 1). As the first author, the candidate developed the research hypotheses, performed data analyses, and wrote the manuscript.

The content in Chapter 4 has been published in *European Respiratory Journal* under the title "*Impact of adverse drug reaction and predictivity of quality of life status in tuberculosis*". This study received the ethics approval from the University of British Columbia Behavioral Research Ethics Board with the certificate number of *B04-0048* (APPENDIX 1). As the first author, the candidate's role in this study was the development of research hypotheses, statistical analysis, and writing of the final manuscript.

The manuscript in Chapter 5 has been submitted for publication under the title "*Patients' preferences for latent tuberculosis infection treatment: a discrete choice experiment*". Two sub-studies were conducted. Ethics approvals were received from the University of British Columbia Behavioral Research Ethics Board and the certificate numbers were *H08-03057* (APPENDIX 2) and *H09-01442* (APPENDIX 3). As the first author, the candidate's role was in all aspects of this study including formulating research questions, questionnaire design, study subject recruitment, data collection, statistical analyses, and writing of the final manuscript.

The work presented in this thesis was conducted and disseminated by the Doctoral candidate. The contribution of co-authors was primarily through providing research direction and intellectual support, reviewing manuscripts prior to submission for publication and offering critical evaluations. The Doctoral candidate was responsible for the writing and the final content of the manuscripts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADR:	Adverse Drug Reaction
AIC:	Akaike Information Criteria
AIDS:	Acquired Immune Deficiency Syndrome
ANOVA:	Analysis of Variance
BC:	British Columbia
BCCDC:	British Columbia Center for Disease Control
BCG:	Bacille Calmette Guerin
BDQ:	Brief Disability Questionnaire
Beck-DI:	Beck Depression Inventory
BIC:	Bayesian Information Criteria
BP:	Bodily Pain
CA:	Conjoint Analysis
CAT:	Computerized Adaptive Testing
CEA:	Cost Effectiveness Analysis
CES-D:	Epidemiological Studies Depression Scale
CI:	Confidence Interval
CLR:	Conditional Logistic Regression
COPD:	Chronic Obstructive Pulmonary Disease
CTT:	Classical Test Theory
DCE:	Discrete Choice Experiment
DOT:	Directly Observed Therapy

EMB:	Ethambutol
EQ-5D:	Euro Quality of Life 5 Dimension
FVC:	Forced Vital Capacity
GH:	General Health
GHQ-12:	General Health Questionnaire 12
GQOLI-74:	General Quality of Life Interview
HII:	Hierarchical Information Integration
HIV:	Human Immunodeficiency Virus
HRQL:	Health-Related Quality of Life
HSUV:	Health State Utility Value
HUI-2:	Health Utility Index 2
HUI-3:	Health Utility Index 3
ICC:	Intra-Class Correlation
IGRA:	Interferon-Gamma Release Assay
IIA:	Independent Irrelevant Alternative
INH:	Isoniazid
IQR:	Inter-Quartile Range
IRT:	Item Response Theory
KPS:	Karnofsky Performance Status
LCA:	Latent Class Analysis
LL:	Log-likelihood
LTBI:	Latent Tuberculosis Infection
MCID:	Minimally Clinically Important Difference

MCS:	Mental Component Summary
MDR-TB:	Multidrug Resistant Tuberculosis
MH:	Mental Health
MHI-5	Mental Health Index
MOS:	Medical Outcome Study
OR:	Odds Ratio
OTC:	Over The Counter
PCS:	Physical Component Summary
PF:	Physical Functioning
PRO:	Patient-Reported Outcome
PTB:	Pulmonary Tuberculosis
PZA:	Pyrazinamide
QALYs:	Quality Adjusted Life Years
QFG-IT:	QuantiFERON-TB Gold In-Tube
QLI	Quality of Life Index
QLQ:	Quality of Life Questionnaire
RCT:	Randomized Controlled Trial
RE:	Role-Emotional
RMP:	Rifampin
RP:	Role-Physical (SF-36 subscale)
RP:	Revealed Preference
RUT:	Random Utility Theory
SCL-90:	Symptoms Checklist 90

SD:	Standard Deviation
SE:	Standard Error
SF:	Social Functioning
SF-36:	Short-Form 36
SF-6D:	Short-Form 6 Dimension
SG:	Standard Gamble
SGRQ:	St. George Respiratory Questionnaire
SP:	Stated Preference
SSRS:	Social Support Rating Scale
TB:	Tuberculosis
TST:	Tuberculin Skin Test
TTO:	Time Trade Off
VAS:	Visual Analogue Scale
VT:	Vitality
WHO:	World Health Organization
XDR-TB:	Extensively Drug Resistant Tuberculosis

ACKNOWLEDGEMENTS

I am heartily grateful to my co-supervisors *Dr. Carlo A. Marra* and *Dr. Fawziah Marra* for drawing the blueprint of this great project and helping me make it come true. Over the past four years, they helped me discover a great well of insights, ideas, and creativity and the resolve to express them. Without their dedication and unwavering faith in me and my work, none of this would have been possible.

I am extremely fortunate of having *Dr. J. Mark FitzGerald, Dr. Jacek Kopec*, and *Dr. Aslam H. Anis* as my Ph.D. research committee members. I sincerely appreciate their valuable guidance and constructive feedback, which significantly improved the quality of my work. Special thank to *Dr. Kathleen MacLeod*, the committee chair, who kindly dedicated her time to supervise my progress over the past four years. Much thank to *Dr. Larry Lynd* for his kind inputs in the DCE study. Also I would like to thank *Lindsey Colley* and *Maja Grubisic* for their kind help and constant statistical support during my program.

I am very grateful to *Dr. R. Kevin Elwood* and all the clinic staff at British Columbia Center for Disease Control tuberculosis clinics (Vancouver and New Westminster), for their generous cooperation during the patient recruitment. Also, my sincere thank to all the study participants for volunteering their valuable time and wish them all the best.

Thanks go to all the great people at Collaboration for Outcomes Research and Evaluation, who shared their knowledge and time during the completion of this work. Best regards and thanks to

all my friends who always stay beside me and give me the strength to face the complexities of life and to follow my goals.

Lastly, my deepest gratitude goes to my parents, who have always pushed me ahead with the best gift in my life, education. To them, I dedicate this thesis.

CHAPTER 1 INTRODUCTION

1.1 TUBERCULOSIS

Tuberculosis (TB) is a bacterial infection caused by Mycobacterium tuberculosis (*M. tb*). *M. tb* can affect almost any part of the human body, but the lung is the most common site of infection. Without proper treatment, people with pulmonary TB (PTB) disease may be infectious and spread the disease. *M. tb* is transmitted from person to person primarily through inhalation and rarely through ingestion or percutaneous inoculation [1].

When healthy people are exposed to *M. tb* bacteria, 10-30% of them are likely to become infected [1]. Among immunocompetent individuals infected, approximately 5% would develop active TB disease within 18 to 24 months; another 5%, after various period of latency, would progress to active disease sometime during their lifetime; and the remaining 90% of infected people would live with the latent TB infection (LTBI) and never develop active disease [1]. The risk of developing active TB disease is determined by a wide range of factors, e.g. the time since infection, older age, close contact with an active TB case, and underlying immunocompromising conditions (e.g. HIV co-infection, use of immunosuppressive drugs, and substance abuse). Among immunocompromised individuals with LTBI, the risk of progressing to active TB disease could approach 10-20% every year [2].

The bacille Calmette-Guérin (BCG) is a widely used vaccine directed against TB. However, its mechanism of protection and effectiveness against TB is not well understood. BCG vaccination appears to be able to effectively prevent TB meningitis and miliary TB, for which reason the World Health Organization (WHO) recommends that BCG be given to children born in countries with a high burden of TB [3,4]. Epidemiologic and autopsy evidences suggest that BCG vaccination does not prevent the establishment of TB infection among exposed individuals [5,6]. However, BCG may be able to limit the subsequent multiplication and dissemination of the bacteria and thus may prevent the development of lesions [7].

1.1.1 Epidemiology of Tuberculosis

Despite the available effective treatment and prevention strategies, TB continues to be a major public health threat worldwide [8-10]. The burden and impact is so significant that the WHO claimed TB a "global emergency" in 1993 [9]. It is estimated that 32% of the world's population could be latently infected with TB [9, 11]; about 9 million people develop TB disease every year and approximately 2 million people die from this disease [9,12]. The emergence of HIV/AIDS epidemic has contributed significantly to the global TB burden because the two infections are closely connected and reinforce one another [2, 13]. Globally, among the HIV-infected individuals, 30% are estimated to be also latently infected with TB. And TB accounted for 23% of AIDS-related deaths worldwide [2,13]. The emergence of multidrug-resistant TB (MDR-TB) and extensively drug resistant TB

(XDR-TB) and the world population growth and migration is further complicating the situation. It has been projected that, between 2000 and 2020, about one billion people will be infected, 200 million people will develop TB disease, and 35 million will die from TB if control and prevention strategies are not further developed and implemented [14].

TB has been intimately linked with poverty, low socioeconomic status, and inadequate access to health care. Globally, TB is most often found in developing countries. Over half of global TB cases were reported in Asia (South-east Asia and Western Pacific regions) and one third from the African region. Among the fifteen countries with the highest estimated incidence rates of TB, thirteen were from Africa [12]. In developed countries, such as Canada, TB was once a major cause of morbidity and mortality early in the 20th century. However, with the improvement in living conditions, the development of public health infrastructures and the availability of anti-TB medications, the incidence of TB has been steadily declining and has reached a plateau over the past several decades [15]. Now, Canada has one of the lowest national incidence rates of TB in the world, at 4.7 per 100,000 population in 2007 [15]. However, high incidence rates are often observed among some sub-populations within Canada, e.g. foreign-born immigrants, Aboriginal Canadians, and marginalized inner city populations, especially injection drug users, which makes TB more a social problem than a medical one [15-18]. In 2007, the TB incidence rate was 25.7 per 100,000 population among the Canadian-born Aboriginals and 14.1 per 100,000 population among foreign-born immigrants. In contrast, among non-aboriginal Canadians, the incidence rate of TB was only 0.7 per 100,000 population [15].

1.1.2 Treatment of Active Tuberculosis Disease

If left untreated, an active TB patient could infect about 20% of people who had contact with him/her [1]. Therefore, prompt diagnosis and treatment of active TB patients is the most critical strategy in TB control and prevention [1,19]. In Canada, people diagnosed with active TB disease are required to be treated with the standard regimen for 6 to 9 months or until a minimum of 80% of the prescribed doses have been completed. This includes a combination of four first-line anti-TB medications, isoniazid (INH), rifampin (RMP), pyrazinamide (PZA) and ethambutol (EMB), for 2 months and, if the bacterium is confirmed to be sensitive, 4 more months of INH, RMP and PZA. EMB is included in the initial treatment phase primarily to prevent the emergence of drug resistance [1].

Although the standard anti-TB therapy is effective, most anti-TB medications are associated with various significant adverse drug reactions (ADRs). It has been reported that about 30% of people treated with anti-TB medications could develop one or more major ADR(s) and the majority of these ADRs occur within the first 3 months of the treatment [20-22]. Commonly reported ADRs include gastrointestinal system symptoms, liver damage, rash, pruritus, paresthesia, visual disturbance, muscle pain, and joint pain [21,22]. ADRs can be serious and life-threatening and lead to treatment discontinuation, as well as necessitating a prolongation of treatment as medications are stopped and then restarted. The discontinuation of first-line anti-TB medications due to ADRs often requires patients to switch to less effective and usually more toxic second-line

medications. As such, patients might be left at a higher risk of treatment failure or reactivation of active TB disease [1,23].

1.1.3 Treatment of Latent Tuberculosis Infection

Since 60~90% of new active TB cases arise from people with LTBI, identification and treatment of LTBI is another essential strategy in preventing the spread of TB infection [1,19,23,24]. Identification of LTBI often relies on "targeted screening" in order to select individuals with increased risk of developing active TB disease [23,24]. In British Columbia (BC), people are screened for LTBI for various reasons, e.g. contact with an active TB case, prior to school admission, prior to employment at health care and child care facilities, and admission to detoxification centers [23]. Until recently, the only diagnostic test for detecting *M. tb* infection is the tuberculin skin test (TST). Using this test, the diagnosis of LTBI has been defined as a positive TST without clinical or radiological evidences of active disease. However, the sensitivity and specificity of TST has long been criticized. False positives are common because past BCG vaccination may confound the TST result, i.e. a positive TST may reflect the BCG vaccination rather than the *M. tb* infection. Recently available interferon-gamma release assay (IGRA) blood tests, e.g. QuantiFERON-TB Gold In-Tube (QFG-IT, Cellestis Ltd, Australia) and T-SPOT.TB (Oxford Immunotec, Abingdon, UK), are not confounded by BCG and may offer more sensitive and specific diagnostic tools for *M. tb* infection [25,26].

With proper treatment of LTBI, the number of people who would develop active TB disease can be significantly decreased. Monotherapy with INH has been the most commonly used strategy for treating LTBI for over 40 years and its efficacy has been well established through multiple clinical trails [1,23,27,28]. Treatment with 6 months of INH have achieved a 69% reduction of developing active TB disease and 12 months of INH therapy has resulted in a 93% reduction [1,27,28]. However, the effectiveness of the preventive treatment has been limited by poor adherence among people with LTBI.

In Canada, 9 months of INH treatment is routinely offered to people diagnosed with LTBI who are at increased risk of developing active TB disease. RMP-containing regimens (RMP alone or combined with INH) can also be used as alternatives if INH drug toxicity is significant, or the patient was exposed to INH-resistant source active TB cases or longer treatment duration is not feasible [1]. Since treatment of LTBI leads to significant benefits both at the individual and public health levels, an important objective of Canadian TB control program is to reach 80% acceptance of preventive treatment among high-risk individuals with LTBI [1]. However, the treatment acceptance and completion rate in practice is well below the targeted level [19,29]. In 2004, in BC, 2,370 individuals were offered LTBI treatment and 1,166 (49.2%) of them accepted and started the treatment. The majority people were prescribed INH, alone or combined with other TB medications. Among people who started the LTBI treatment, 50% successfully completed the prescribed doses [30].

1.2 PATIENT-REPORTED OUTCOMES

In modern health care, the evaluation of health has shifted from merely using traditional biomedical measures to including patient-reported outcomes (PROs) [31,32]. The importance and value of measuring PROs has been well recognized by health care professionals and policy-makers. In clinical practice, patients are increasingly involved in making decisions about their own health and treatment. Also, patients' preferences and values have been more considered in justifying health care resource allocation and policy making.

PRO measures consist of a wide variety of subjective health outcomes that are provided directly by patients, e.g. symptoms, physical functioning status, psychological well-being, global health perception, and treatment satisfaction [31]. PRO measures offer more insights into human concerns and perceptions, such as pain, depression, well-being and satisfaction, that could be missed by traditional biomedical data. Therefore, PRO measurements add values to traditional physical measurements and laboratory results when evaluating the impact of illness and the effectiveness of treatment. This thesis focused on measuring health-related quality of life (HRQL), health state utility values (HSUVs) and preferences of patients undergoing treatment for TB.

1.2.1 Health-related Quality of Life

HRQL is an inherently complex and multidimensional concept that intends to describe the quality rather than merely the quantity of health. Although there is no universal

definition for HRQL, it can be defined as the individual's subjective perception of the impact of the disease and its treatment(s) on his/her physical health, mental well-being, and social functioning. HRQL reflects not only the individual's health problems, but also his/her attitude and emotional response to these problems. Structured and well-tested generic and disease-specific questionnaires or instruments are usually used to quantitatively measure HRQL [31-34], Chapter 2 provides more information about HRQL measurements and commonly used instruments in current TB research. These score-based instruments cover multiple health dimensions and can provide a descriptive profile of patients' health status and therefore they are often used by clinical researchers, along with traditional biological measures.

1.2.2 Utility and Preference

Although 'preference', 'utility' and 'value' have often been used interchangeably in the literature, they are different concepts in health economics. 'Preference' is the umbrella term used to describe the overall desirability, while 'utility' and 'value' are two subtypes of 'preference' [33]. 'Utility' refers to preference that is measured under uncertain circumstances [33]. For example, in a traditional standard gamble (SG) exercise, respondents could be asked to compare two outcomes, where at least one contains uncertainty, i.e. risk or probability. In contrast, 'value' is the preference derived under certainty. For example, using a visual analog scale (VAS), respondents may be asked to place a desirability score on a given health outcome that is assumed to occur with certainty [33].

In health economics, SG, time trade-off (TTO) and rating scales like VAS, are commonly used to directly measure utility or value [33,35]. However, to avoid the inconvenience and respondent burden associated with applying these direct approaches, preference-based generic utility instruments, e.g. Health Utility Index (HUI), Short Form 6D (SF-6D) and Euro Quality of Life 5D (EQ-5D), have been developed to indirectly measure utility. These instruments consist of a health status classification system and a scoring formula that is used to assign preference weights to the health states defined by the classification system. The preference weights were initially obtained by using direct approaches from general population.

In contrast to traditional HRQL instruments, where multiple scores are calculated to reflect various health dimensions, health state utility instruments summarize HRQL into one single index score, where a score of '1' often represents 'perfect health' and '0' represents 'death'. Health state utility value (HSUV) can be utilized as the quality-adjustment weights to calculate quality adjusted life years (QALYs), which combines the length of time spent in a particular health state and the quality of that health state. As an important measurement of health outcomes and benefits gained from health care interventions, QALYs have been incorporated into cost-effectiveness/cost-utility analyses, as the denominator in incremental cost-effectiveness ratios, to facilitate health care decision-making [33].

Until recently, the evaluation of benefits gained from health interventions has been mostly focused on patients' health outcomes. It has been recognized that individuals may

derive utility not merely from the improvement in their health status, but also from important aspects that do not directly affect their health outcomes, such as the treatment procedure and convenience factors. Therefore, stated preference (SP) techniques have been applied in health economics to elicit patients' or the public's preferences for health care interventions and to allow for the inclusion of both health outcomes and non-health outcomes when evaluating the benefits of health interventions [36].

Conjoint analysis (CA) is a large family of SP techniques, including ranking (rank a set of alternative scenarios), rating (score alternative scenarios on a scale of 1 to 10), paired comparison and choice-based exercises. The discrete choice experiment (DCE) is a choice-based CA technique. In the literature, DCE and CA are sometimes used interchangeably, but they are two different terminologies [37].

1.3 DISCRETE CHOICE EXPERIMENT

Public and patient preferences have been given more considerations in health policy decision-making and service delivery in modern health care. As a preference-elicitation technique, DCE has been widely used in marketing, transportation economics and environmental economics and recently gained popularity in the health care sector for a wide range of applications, e.g. understanding the decision-making process of patients and health care professionals, estimating the willingness to pay, and predicting immunization uptake rates [38-42].

In a typical DCE survey, the treatment or service of interest is defined by a few attributes with various levels assigned to them. Respondents are presented with a series of choice sets, each comprising two or more hypothetical treatment alternatives (i.e. different combinations of attribute levels). Respondents are asked to compare these alternatives and choose their preferred one in a way that they have to make trade-offs between these attributes. Consequently, the relative importance or contribution of each attribute to the overall utility of the treatment or service could be obtained by analyzing respondents' choices.

The DCE technique reflects an integration of a few economic theories [38,39]. DCE is based on the assumption that any commodity or product can be described by a few of its attributes or characteristics; individuals derive utility from these underlying attributes, rather than the commodity per se; individuals have preference or utility for those attributes; although utility/preference cannot be observed directly, it can be revealed and estimated through a series of choices the individual makes [38,39].

Choices made by respondents in a DCE survey are used to model utility under the random utility theory (RUT) framework. In RUT, the indirect utility (U) of individual i ($i=1,\ldots,n$) for a specific choice alternative j ($j=1,\ldots,J$), represented by U_{ij}, can be described by an explainable component, V_{ij}, and a non-explainable random component, ε_{ij}.

$$U_{ij} = V_{ij} + \varepsilon_{ij}, \quad \mathrm{j} = 1, \ldots, \mathrm{J}$$

(Eq.1.1)

The explainable component (V_{ij}) of the utility represents the attributes (X_{ij}) of the product of interest and the characteristics (Z_i) of individual *i* making the choice. It is modeled as Equation 1.2 where β and γ are the coefficients to be estimated:

$$V_{ij} = X_{ij}\beta + Z_i\gamma$$

(Eq. 1.2)

The non-explainable component (ε_{ij}) is to pick up the difference between the true utility and the utility modeled. It may reflect unobserved or unobservable factors affecting respondents' choices, preference variation, functional specification and/or measurement error.

The individual is assumed to have some construct of indirect utility for choice alternatives. The choice-making process in DCE can be seen as a comparison of the indirect utility of alternatives. When given two alternatives, a rational individual is expected to choose the one that leads to higher utility. Therefore, in a choice between two alternatives *j* and *k*, the individual *i* will choose alternative *j* over *k* if and only if:

$$V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}$$

(Eq. 1.3)

The probability that the individual *i* will choose alternative *j* over *k* is modeled as:

$$\begin{aligned} P_i(j) &= P(U_{ij} > U_{ik}) \\ &= P(V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik}) \\ &= P(V_{ij} - V_{ik} > \varepsilon_{ij} - \varepsilon_{ik}) \end{aligned}$$

(Eq. 1.4)

The choice model estimated would depend on the assumption made about the probability distribution of the random component (ε_{ij} - ε_{ik}). The assumption about the distribution

leads to a family of choice models with different statistical properties and suitable application situations. Commonly used choice models include multinomial logit, conditional logit, nested multinomial logit, mixed logit, and recently, latent class analysis [38,43].

1.4 VALIDATION OF PATIENT-REPORTED OUTCOMES MEASUREMENTS

PRO measurements are becoming increasingly important in clinical research and being frequently incorporated into health care decision-making. Prior to advocating the wide application of a PRO measure, it is important to go through the validation process, during which reliability, validity and responsiveness are commonly assessed in the specific population under investigation. Reliability, validity and responsiveness should not be regarded as a fixed property of the instrument. An instrument validated in one specific population will not necessarily perform well among other patient populations. Validation is also an evolving process to accumulate evidence to support the usefulness and creditability of PRO measurements in informing clinical practice and health policy making.

1.4.1 Reliability

Reliability concerns the extent to which a measurement is free of random measurement errors and can be regarded as the proportion of a measurement that is signal rather than

noise [44,45]. In psychometrics, commonly assessed reliability indicators include internal consistency reliability and test-retest reliability [44]. For a traditional HRQL instrument, internal consistency assesses how well items within the same dimension or subscales within an instrument measure the same concept, usually indicated by the overall correlation between those items or subscales. Test-retest reliability, reproducibility or stability over time, assesses whether an instrument could produce the same result over repeated administrations when the respondents have not changed in terms of the dimension measured.

1.4.2 Validity

Validity commonly refers to the extent to which an instrument measures what it purports to measure [44,45]. Validity is the basis of the interpretability and meaningfulness of the measurements derived from the instrument [44]. Three forms of validity are often considered in the literature: face or content validity, criterion validity and construct validity.

Evaluation of face or content validity is often a subjective judgment of whether the instrument is adequate to perform its proposed application. A commonly used procedure to test content validity often involves critical assessment and judgment from experts in the field or from individuals who belong to the target population. For HRQL and health utility instruments, content validity refers to how adequately the items or questions within the instrument address the health dimension of interest [44]. In the context of DCE,

content validity is concerned about the issues mainly involved in its experimental design, e.g. selection of attributes and levels, number and framing of choice questions, and the presentation of the questionnaire [38-40].

Criterion validity assesses the degree of agreement between a tested instrument and another measure that is regarded as a "gold standard". Although criterion validity would provide the strongest evidence to support the validity of the tested instrument, a criterion or "gold standard" measure for HRQL measurements rarely exists [44,45]. When measuring stated preferences using the DCE technique in economics, the "gold standard" is often considered to be the choices that individuals make in actual market choice situations (known as revealed preferences, as opposed to stated preferences) [42,46]. However, the opportunity of accessing actual choice behavior in health care field is relatively limited. In the absence of a "gold standard", construct validity of PRO measures is often assessed.

Construct validity involves comparisons between the tested instrument and other external measures, such as other well-established HRQL instruments and clinical criteria of physical and mental health status. Two aspects of construct validity are often examined: (1) convergent validity, which tests whether the measurements will correlate well with other methods that measure the similar concept; (2) divergent validity, which holds if the measurements will not correlate well with methods that measure different concepts. For example, measurements obtained from a valid HRQL instrument should be able to distinguish between patient groups differing in physical and/or mental health status and

severity categorized by clinical criteria. To evaluate the construct validity of the DCE technique, van der pol et al. [42] compared preferences obtained from direct open-ended questions with those derived from a DCE survey designed to obtain preferences for perinatal community based care.

1.4.3 Responsiveness

Responsiveness usually refers to the instrument's ability to reflect an underlying change when it occurs over time. There is a discussion in the literature whether responsiveness should be considered a separate psychometric property of an instrument or just a form of validity, i.e. longitudinal construct validity [47]. Responsiveness is an important psychometric property when considering the interpretability of changes in HRQL scores. However, there is a considerable inconsistency in the approaches for evaluating responsiveness [48]. Commonly used statistical indicators include effect size and standardized response mean. In addition, some external criteria for determining that patients have changed are often used to evaluate an instrument's responsiveness, e.g. patients' global rating of changes (e.g. self-rated health change over time as 'better', 'same' or 'worse') and observed change in clinical diagnostic criteria.

1.5 PROBLEM STATEMENT AND STUDY JUSTIFICATION

1.5.1 Measurement of Health-related Quality of Life in Tuberculosis

The focus of TB management has been primarily focused on achieving microbiological 'cure'. Little work has been done to understand the impact of the disease and its treatment from patients' perspectives. A literature review by Chang et al. in 2004 did not locate studies that measured HRQL experienced by TB patients using standardized instruments. To update our knowledge of HRQL research on TB, we conducted a systematic review in 2008 (Chapter 2) and found an increasing effort has been dedicated to this research field over the past few years. A number of studies attempted to investigate HRQL or relevant issues among TB patients using generic instruments, a battery of condition-specific instruments and qualitative methodologies as well, such as focus group discussions.

Among active TB patients, the literature review revealed that many aspects associated with active TB disease can lead to significant decrement in patients' HRQL. Besides various clinical symptoms, many patients also reported experiencing anxiety and depression. The social stigma attached to TB may affect patients' social functioning and life roles in their community. Even after being microbiologically "cured", some patients may still be left with residual anatomic changes and various degrees of functional damage. Most of these impacts cannot be adequately captured by traditional clinical indicators [31,34].

A major concern in active TB patient management is the treatment associated ADRs. Current knowledge of these ADRs has been limited to their frequency and clinical nature [20-22]. Their impacts on HRQL experienced by TB patients have not been investigated. In a recent study conducted in HIV-infected patients on combination antiretroviral therapy, patients' self-reported side effects were associated with a substantial decrement in HSUVs [49]. Previous studies have measured the HRQL experienced by TB patients before, during, and after the anti-TB treatment and have shown that the treatment has resulted in overall improvement in patients' HRQL [50-52]. However, none have separated the benefit and the potential harm of the treatment. It would be interesting to know, among TB patients who developed major ADRs, whether the anti-TB treatment led to more benefits or harms on patients' HRQL. The study in Chapter 4 was the first study to examine the specific impact of anti-TB treatment induced ADRs on HRQL.

An important research interest on HRQL measurement is its predictive value for patients' future health outcomes and prognosis. Self-reported health status may capture patients' known or unknown underlying health problems, risk behaviors, personality and psychosocial coping mechanism, which altogether may provide some insights into their health status and future health outcomes [53-57]. In the study in Chapter 4, we also investigated the relationship between patients' baseline HRQL status and the likelihood of developing ADRs during the following three months of anti-TB treatment.

Although generally considered to be clinically normal and otherwise healthy, people with LTBI may share the same psychological and social impact from TB, as those with active

TB disease. The recommended INH preventive treatment takes 9 months and may have various ADRs. It would be interesting to know how the infection and the preventive treatment affect the HRQL among LTBI patients. One specific objective in Chapter 3 study was to describe and compare the HRQL status among patients with LTBI and those with active TB disease.

1.5.2 Generation of Health State Utility Values in Tuberculosis

Due to the exponential increase in health care needs and limited resources, it is crucial to use analytical methods to guide the efficient and equitable resource allocation among competing demands. Economic analysis to demonstrate the comparative cost-effectiveness of different health interventions can provide recommendations to inform decision makers, e.g. how to maximize the benefits from a given health budget or find the least costly way to obtain a desired level of benefits.

In addition to its impact on humanistic outcomes such as quality of life, TB exerts a significant economic burden on society through the cost of diagnosis and treatment, the resources spent on prevention and control and, indirectly, the loss of productivity. Socio-demographically, TB affects all age groups, but it is responsible for deaths in the most productive ages between 15 and 44 [58]. Economic evaluations can have an impact on TB through justifying the most effective strategies and practices or informing budget development in TB prevention and control [59]. Appropriate HSUV to calculate QALYs is critical in conducting proper economic evaluation analysis.

A few economic analyses have recently been done in TB treatment, screening and control strategies [59-61]. However, none of these studies incorporated QALYs as the health outcome measures, primarily because of the lack of HSUVs among TB patients in the literature. To fill in this research gap, one objective of the study in Chapter 3 was to generate needed health utility measurements in TB that could be utilized in future health economic evaluations.

Another issue around measuring health utility is that there are a number of indirect utility instruments commonly used in the literature, including SF-6D, HUI-2, HUI-3, and EQ-5D [33]. Each of these instruments covers multiple health dimensions and uses different techniques to obtain preference weights from different populations. There has been a major concern about whether these instruments can be used interchangeably and whether they lead to comparable outcomes and eventually reach to similar policy decisions. There is also a need to assess these instruments' construct validity and investigate their strengths and weaknesses in the application among TB population.

1.5.3 Understanding Patients' Preference toward the Preventive Treatment

In the management of LTBI patients, one major challenge is to successfully advocate the preventive treatment to people with LTBI and to improve the treatment completion rate. Although the preventive treatment can provide both individual benefits and public health protection, its effectiveness in practice is significantly compromised due to the low acceptance rate and poor adherence among patients [1]. So far, the majority of research

has been focused on finding the socio-demographic and other modifiable predictors associated with poor adherence behavior among patients on preventive treatment. There has been limited empiric research in understanding patients' treatment decision-making process. Decision-making has been recognized to be a complex behavior, not only depending on the probability of risks and benefits, but also how patients perceive and value these outcomes and uncertainties. Better understanding patient preferences for treatment options can provide important information to help improve the acceptance of and adherence to the preventive treatment. The DCE technique offers a tool to quantify patient's preferences when making decisions on whether to accept the preventive treatment.

1.6 OBJECTIVES AND THESIS ORGANIZATION

The overall objectives of this thesis were: (1) to measure and compare the HRQL among active TB patients and LTBI patients; (2) to generate health state utility values that could be utilized in future economic evaluations involved TB population; (3) in active TB patients, to investigate the relationship between HRQL measurements and the treatment associated ADRs; and (4) in LTBI patients, to quantify their preferences regarding the prophylactic treatment.

This thesis consists of six chapters. Chapter 1 provides a general introduction to: (1) the epidemiology of TB; (2) the treatment of active TB disease and the potential risk of adverse drug reactions; (3) the treatment of LTBI and the low acceptance rates among

patients; (4) PRO measures, including HRQL, health utility measurements, and stated preference; (5) the theoretical background of DCE, a stated preference-elicitation technique; and (6) validation of PRO measures.

Chapter 2 is a systematic review on measuring HRQL and health utility in TB population. Specific research questions that were answered in this chapter include: (1) structured HRQL instruments used in current TB research; (2) the impact of TB disease or infection and the associated treatment on patients' HRQL; and (3) the demographic, socio-economic, and clinical factors associated with HRQL outcomes in TB patients.

Three original studies were conducted to address the thesis objectives, from Chapter 3 to Chapter 5. Chapter 3 presents the results of a cross-sectional evaluation of HRQL and health state utility among latent and active TB patients. The specific objectives of this study were to: (1) compare the HRQL status between people with active TB disease and those with LTBI; (2) generate health state utility values for people with active TB disease or LTBI; (3) examine the construct validity of SF-36, SF-6D, HUI-2, and HUI-3 in TB population; (4) compare if HUI-2, HUI-3 and SF-6D can generate interchangeable utility scores for the same individual; and (5) investigate the ceiling and floor effect problems for the application of the three utility instruments in TB patient population.

Chapter four is a longitudinal study to examine the relationship between HRQL and ADRs among active TB patients on treatment. The specific research objectives of this study were to: (1) investigate the impact of ADRs of different severity on HRQL; and (2)

examine the correlation between baseline HRQL measurements and the possibility of developing ADRs during the treatment.

In Chapter 5, a DCE survey was developed and administered among LTBI patients to measure patients' preferences when making decisions on whether to accept LTBI preventive treatment. Specific objectives of this study included: (1) to quantify patients' preferences for six key treatment attributes; (2) to examine the preference heterogeneity among respondents; and (3) to explore the socio-demographic factors associated with preference heterogeneity.

The final chapter provides a summary of major study findings and discusses the strengths and limitations of this work. Potential implications of the findings and some recommendations for future research are discussed as well.

1.7 REFERENCES

1. Health Canada. Canadian tuberculosis standards. 6th edition. 2007. http://www.phac-aspc.gc.ca/tbpc-latb/pubs/pdf/tbstand07_e.pdf (accessed at June 8, 2010).

2. World Health Organization (WHO). Tuberculosis and HIV. http://www.who.int/hiv/topics/tb/en/index.html (accessed June 20, 2010)

3. Rodrigues LC, Diwan VK, Wheeler JG. Protective effect of BCG against tuberculous meningitis and miliary tuberculosis: a meta-analysis. Int J Epidemiol 1993;22:1154-1158.

4. WHO. Geneva. Weekly epidemiological record. No. 4, 2004, 79, 25-40. http://www.who.int/immunization/wer7904BCG_Jan04_position_paper.pdf (accessed at June 8, 2010).

5. Styblo K, Meijer J. Impact of BCG vaccination programmes in children and young adults on the tuberculosis problem. Tubercle 1976;57:17-43.

6. Sutherland I, Lindgren I. The protective effect of BCG vaccination as indicated by autopsy studies. Tubercle 1979;60:225-231.

7. Smith DW, McMurray DN, Wiegeshaus EH, Grover AA, Harding GE. Host-parasite relationships in experimental airborne tuberculosis. IV. Early events in the course of infection in vaccinated and nonvaccinated guinea pigs. Am Rev Respir Dis 1970;102:937-949.

8. Raviglione MD, Snider DE, Kochi A. Global epidemiology of tuberculosis: morbidity and mortality of a worldwide epidemic. JAMA 1995;273:220-226.

9. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Communicable Diseases Prevention and Control, World Health Organization (WHO). Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO global surveillance and monitoring project. JAMA 1999;282:677-686.

10. Blumberg HM, Burman WJ, Chaisson RE, Daley CL, Etkind SC, Friedman LN, Fujiwara P, Grzemska M, Hopewell PC, Iseman MD, Jasmer RM, Koppaka V, Menzies RI, O'Brien RJ, Reves RR, Reichman LB, Simone PM, Starke JR, Vernon AA; American Thoracic Society, Centers for Disease Control and Prevention and the Infectious Diseases Society. American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America: treatment of tuberculosis. Am J Respir Crit Care Med. 2003;167:603-662.

11. Dye C. Global epidemiology of tuberculosis. Lancet 2006;367:938-940.

12. World Health Organization (WHO). Global tuberculosis control: epidemiology, strategy, financing. WHO Report 2009. http://www.who.int/tb/publications/global_report/2009/en/ (accessed at June 17, 2010).

13. Raviglione MC, Narain JP, Kochi A. HIV-associated tuberculosis in developing countries: clinical features, diagnosis, and treatment. Bull WHO 1992;70:515-526.

14. World Health Organization (WHO). Tuberculosis. http://www.who.int/immunization/topics/tuberculosis/en/ (accessed at June 8, 2010).

15. Public health agency of Canada. Tuberculosis in Canada 2007 Pre-Release. http://www.phac-aspc.gc.ca/publicat/2008/tbcanpre07/index-eng.php (accessed at June 17, 2009).

16. FitzGerald JM, Wang L, Elwood RK. Tuberculosis: control of the disease among aboriginal people in Canada. Can Med Assoc J 2000;162:351-355.

17. FitzGerald JM. Optimizing tuberculosis control in the inner city. Can Med Assoc J 1999;160:821-822.

18. Hernandez E, Kunimoto D, Wang L, Rogrigues M, Black W, Elwood RK, FitzGerald JM. Predictors for clustering among TB cases in Vancouver: a four year molecular epidemiology study. Can Med Assoc J 2002;167:349-352.

19. Horsburgh CR Jr, Goldberg S, Bethel J, Chen S, Colson PW, Hirsch-Moverman Y, Hughes S, Shrestha-Kuwahara R, Sterling TR, Wall K, Weinfurter P. Tuberculosis Epidemiologic Studies Consortium. Latent tuberculosis infection treatment acceptance and completion in the United States and Canada. Chest 2010;137:401-409.

20. Yee D, Valiquette C, Pelletier M, Parisien I, Rocher I, Menzies D. Incidence of serious side effects from first-line antituberculosis drugs among patients treated for active tuberculosis. Am J Respir Crit Care Med 2003;167:1472-1477.

21. Forget EJ, Menzies D. Adverse reactions to first-line antituberculosis drugs. Expert Opin Drug Saf 2006;5:231-249.

22. Marra F, Marra CA, Bruchet N, Richardson K, Moadebi S, Elwood RK, FitzGerald JM. Adverse drug reactions associated with first-line anti-tuberculosis drug regimens. Int J Tuberc Lung Dis 2007;11:868-875.

23. British Columbia Center for Disease Control. communicable disease control manual.: tuberculosis. http://www.bccdc.ca/NR/rdonlyres/90E2CF31-8621-4082-81B9-BB67B8D1E4F3/0/TB_GF_manual_1999.pdf (accessed at June 17, 2010).

24. Targeted tuberculin testing and treatment of latent tuberculosis infection. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. This is a Joint Statement of the American Thoracic Society (ATS) and the Centers for Disease Control and Prevention (CDC). This statement was endorsed by the Council of the Infectious Diseases Society of America. (IDSA), September 1999, and the sections of this statement. Am J Respir Crit Care Med 2000;161(4 Pt 2):S221-247.

25. Richeldi L. An update on the diagnosis of tuberculosis infection. Am J Respir Crit Care Med 2006;174:736-742.

26. Lalvani A, Pareek M. Interferon gamma release assays: principles and practice. Enferm Infecc Microbiol Clin. 2010;28:245-252.

27. Comstock GW, Ferebee SH, Hammes LM. A controlled trial of community-wide isoniazid prophylaxis in Alaska. Am Rev Respir Dis 1967;95:935-943.

28. International Union Against Tuberculosis Committee on Prophylaxis. Efficacy of various durations of INH preventive therapy for TB: 5 years of follow-up in the IUAT trial. Bull World Health Organ 1982;60:555-564.

29. Hirsch-Moverman Y, Daftary A, Franks J, Colson PW. Adherence to treatment for latent tuberculosis infection: systematic review of studies in the US and Canada. Int J Tuberc Lung Dis 2008;12:1235-1254.

30. British Columbia Centre for Disease Control 2004 annual tuberculosis report. http://www.bccdc.org/downloads/pdf/tb/reports/2004_TB_Annual.pdf (accessed at June 17, 2010).

31. Guyatt GH, Feeny DH, Patrich DL. Measuring health-related quality of life. Ann Intern Med 1993;118:622-629.

32. Sherbourne CD, Sturm R, Wells KB. What outcomes matter to patients? J Gen Intern Med 1999;14:357-363.

33. Drummond MF, Sculpher MJ, Torrance GW, eds. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford University Press, 2005.

34. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care 1989;27(Suppl 3):217-232.

35. Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. J Clin Epidemiol 2003;56:317-325.

36. Ryan M, Netten A, Skåtun D, Smith P. Using discrete choice experiments to estimate a preference-based measure of outcome--an application to social care for older people. J Health Econ 2006;25:927-944.

37. Orme BK. Getting started with conjoint analysis: strategies for product design and pricing research. (2nd ed). Madison, WI: Research Publishers LLC, 2006.

38. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. Cambridge: Cambridge University Press, 2005.

39. Ryan M, Gerard K, Amaya-Amaya M (eds.). Using discrete choice experiments to value health and health care. Chapter one: discrete choice experiments in a nutshell. pp:13-46. Dordrecht: Springer, 2008.

40. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26:661-677.

41. Hall J, Kenny P, King M, Louviere J, Viney R, Yeoh A. Using stated preference discrete choice modelling to evaluate the introduction of varicella vaccination. Health Econ 2002;11:457-465.

42. Van der Pol M, Shiell A, Au F, Johnston D, Tough S. Convergent validity between a discrete choice experiment and a direct, open-ended method: comparison of preferred attribute levels and willingness to pay estimates. Soc Sci Med 2008;67:2043-2050.

43. Boxall PC, Adamowicz WT. Understanding heterogeneous preferences in random utility models: a latent class approach. Environmental and Resource Economics 2002;23:421-446.

44. Streiner GL, Norman RD. Health measurement scales: a practical guide to their development and use. 2nd ed. Oxford: Oxford University Press, 1995.

45. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247-263.

46. Urama KC, Hodge ID. Are stated preferences convergent with revealed preferences? Empirical evidence from Nigeria. Ecological Economics 2006;59:24-37.

47. Hays RD, Hadorn D. Responsiveness to change: An aspect of validity, not a separate dimension. Qual Life Res 1992; 1:73–75.

48. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349-362.

49. Braithwaite RS, Goulet J, Kudel I, Tsevat J, Justice AC. Quantifying the decrement in utility from perceived side effects of combination antiretroviral therapies in patients with HIV. Value Health 2008;11:975-979.

50. Chamla D. The assessment of patients' health-related quality of life during tuberculosis treatment in Wuhan, China. Int J Tuberc Lung Dis 2004;8:1100-1106.

51. Dhingra VK, Rajpal S. Health related quality of life (HRQL) scoring (DR-12 score) in tuberculosis--additional evaluative tool under DOTS. J Commun Dis 2005;37:261-268.

52. Marra CA, Marra F, Colley L, Moadebi S, Elwood RK, Fitzgerald JM. Health-related quality of life trajectories among adults with tuberculosis: differences between latent and active infection. Chest 2008;133:396-403.

53. Mossey JM, Shapiro E. Self-rated health: a predictor of mortality among the elderly. Am J Public Health 1982;72:800-808.

54. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. Am J Public Health 1990;80:446-452.

55. Gold M, Franks P, Erickson P. Assessing the health of the nation. The predictive validity of a preference-based measure and self-rated health. Med Care 1996;34: 163–177.

56. Franks P, Gold MR, Fiscella K. Sociodemographics, self-rated health, and mortality in the US. Soc Sci Med 2003;56:2505–2514.

57. Sehlen S, Lenk M, Hollenhorst H, Schymura B, Aydemir U, Herschbach P, Dühmke E. Quality of life (QoL) as predictive mediator variable for survival in patients with intracerebral neoplasma during radiotherapy. Onkologie 2003;26:38–43.

58. Grimard F, Harling G. The impact of tuberculosis on economic growth. Department of Economics, McGill University. http://neumann.hec.ca/neudc2004/fp/grimard_franque_aout_27.pdf (accessed at June 17, 2010).

59. Floyd K. Costs and effectiveness - the impact of economic studies on TB control. Tuberculosis 2003;83:187-200.

60. Dasgupta K, Schwartzman K, Marchand R, Tennenbaum TN, Brassard P, Menzies D. Comparison of cost-effectiveness of tuberculosis screening of close contacts and foreign-born populations. Am J Respir Crit Care Med 2000;162:2079-2086.

61. Khan K, Muennig P, Behta M, Zivin JG. Global drug-resistance patterns and the management of latent tuberculosis infection in immigrants to the United States. N Engl J Med 2002;347:1850-1859.

CHAPTER 2 MEASURING HEALTH-RELATED QUALITY OF LIFE IN TUBERCULOSIS: A SYSTEMATIC REVIEW

2.1 FOREWORD

This chapter is a literature review on measuring health-related quality of life in tuberculosis patients. The content of this chapter has been published in the journal *Health and Quality of Life Outcomes.*[1] The candidate is the first author on this manuscript which is coauthored by *Dr. Fawziah Marra* and *Dr. Carlo A. Marra*. The candidate's role in this manuscript was the conception and design of the review, data collection and analyses, and writing of the manuscript.

2.2 INTRODUCITON

The assessment of patient reported outcomes (PROs) has become more accepted and valued in the disease management and outcome evaluation. Health-related quality of life (HRQL) is a complex type of PRO that evaluates health status. HRQL broadly describes how well individuals function in daily lives and their own perception of well-being in physical, psychological, and social aspects [1,2]. Although traditional clinical and biological indicators are often intrinsically related to patients' quality of life, they fail to

[1] A version of this chapter has been published in *Health and Quality of Life Outcomes*. Guo N, Marra F, Marra CA. Measuring health-related quality of life in tuberculosis: a systematic review. *Health Qual Life Outcomes* 2009;7:14.

represent one's self-perceived function and well-being in everyday life settings. It is known that patients with chronic diseases place a high value on their mental and social well-being as well as pure physical health [3]. As a result, HRQL has become an area of increasing interest and has been evaluated in many diseases, including tuberculosis (TB). To measure HRQL, two kinds of instruments are often used: generic and disease-specific [1,2,4]. Generic instruments are developed to cover the common and important aspects of health and can be used to assess and compare HRQL across different health conditions and sub-populations [1,4]. In contrast, disease- or condition-specific instruments are designed to reflect unique problems most relevant to a given disease and/or its treatment [1,4]. Theoretically, disease-specific instruments are more precise and more sensitive to small but potentially important differences or changes on HRQL, compared to generic instruments [1,4]. One special category of generic HRQL instruments assesses "preferences" for certain health states [2]. These instruments summarize quality of life into a single utility score, reflecting the 'value' people place on a health state, anchored at 0 (death) and 1 (full health) [2]. Health utility measurements are often used in health economic studies.

Although effective therapy has long been available, TB remains a major public health threat globally, with one third of the world's population infected [5,6]. Many aspects of TB along with its treatment could potentially compromise patients' HRQL. For example, the standard anti-TB therapy consists of four medications and takes at least 6 to 9 months to complete, with serious risks of adverse reactions [6-8]. In some communities, TB patients are perceived as a source of infection and the resultant social rejection and

isolation leads to a long-term impairment on patients' psychosocial well-being [9-14]. Many TB patients also report to experience negative emotions, such as anxiety and fear [13,14]. However, the current goal of TB management is to achieve microbiological 'cure' and there has been little effort taken to consider patients' HRQL. In 2004, Chang et al. published a review summarizing the English medical literature on the quality of life in TB patients [15]. At that time, the authors were unable to locate studies measuring HRQL using standardized instruments. Over the past few years, more effort has been dedicated to this research field. Therefore, the present review was performed to identify published original studies utilizing structured HRQL instruments.

2.3 OBJECTIVES

The objectives of this review were: (1) to identify HRQL instruments used in TB research; (2) to better understand the impact of TB disease or infection and the associated treatment on patients' HRQL; and (3) to examine demographic, socio-economic, and clinical factors associated with HRQL outcomes in TB patients.

2.4 METHODS

2.4.1 Search Strategies for Identification of Potential Studies

A systematic literature search was performed using the following electronic databases: Medline (1950-present), EMBASE (1980-present), Cochrane Register of Controlled

Trials (CENTRAL), CINAHL, PsycINFO, and HaPI (1985-present). Key word searching and/or subject searching were performed, if applicable. The following keywords were used: tuberculosis (TB), Quality of Life (QoL), Quality Adjusted Life Years (QALY), health utility, health status, life quality, and well-being. The limit feature was used to select human studies published between 1981 and 2008 written in English or Chinese (traditional or simplified). The last time electronic database search was conducted during July 22, 2008. The reference sections of the following key journals were manually searched for relevant articles: International Journal of Tuberculosis and Lung Disease, Chest, Quality of Life Research, and Health and Quality of Life Outcomes. Reference lists of included studies, review articles, letters, and comments were checked afterwards. We did not contact the authors of identified studies or relevant experts to locate unpublished studies. Each stage of the literature searching process is illustrated in FIGURE 2.1.

2.4.2 Inclusion and Exclusion Criteria

All clinical trials and observational studies where multi-dimensional HRQL was evaluated, either as a primary or secondary outcome, using structured HRQL instruments were considered in this review. Participants were those diagnosed with active TB disease or latent TB infection (LTBI), regardless of the site and stage of the disease and the treatment status. There were no limitations on age, gender, race, the origin of birth, and other socio-economic status.

For the purpose of this review, HRQL was defined as patients' self-evaluations of the impact of either active TB disease or LTBI and the associated treatments on their physical, mental, and social well-being and functioning. The following requirements for HRQL measurement were set a priori for studies to be included in this review: (1) one multi-dimensional HRQL instrument or a combination of single-dimensional instruments had to be used to capture the broad framework of HRQL; (2) the HRQL instruments could be either generic or disease (or condition) -specific; (3) the origin of the applied instruments had to be identifiable and traceable; (4) the HRQL instruments had to have psychometric properties such as reliability and validity reported from previous studies or were assessed in the specific study being reviewed; (5) HRQL outcomes had to be self-reported by the specific participant, but HRQL measurement that were completed with help from proper proxies, such as family members and caregivers, were also accepted.

Studies were excluded if (1) HRQL was evaluated using qualitative methodologies, such as focus groups; or (2) only one single dimension of HRQL (e.g. depression) was assessed; or (3) HRQL was assessed using instruments designed for the specific study without psychometric properties evaluated and reported; or (4) a modified version of a previously validated instruments (e.g. SF-36) was used as the psychometric properties of the original instrument could be changed by the modification.

2.4.3 Data Extraction

If the study was included in this review, the following information was collected: study design, inclusion and exclusion criteria of subjects, included subjects' socio-demographic characteristics and clinical features, HRQL instrument(s) used, the origin and structure of HRQL instrument(s), administration of HRQL instrument(s), and HRQL outcomes and validation results.

2.5 RESULTS

The literature search identified 2540 articles which were narrowed to 26 [9-14,16-35] (FIGURE 2.1). After reviewing the full texts, 14 studies were further excluded for various reasons: 6 studies used qualitative methodologies [9-14]; 2 studies measured only one single dimension of HRQL [16,17]; 1 study [18] used the Short-Form 36 (SF-36) but the response options of SF-36 were modified to 3 levels (i.e. the same as before, better, and worse) without providing validation data; 1 study [19] used one single question from a structured instrument; 1 study was a duplicate and the earlier version was excluded [20,21]; 1 study [22] used a generic instrument, the General Quality of Life Interview (GQOLI-74), however, no relevant references were provided to track the origin and the psychometric properties of this instrument; 2 articles [23,24] were published from the same study, and therefore only included as one study for the review; another 2 articles, Marra et. al. [25] and Guo et al. [26], reported longitudinal and cross-sectional results from one same study respectively, and thus only one study was counted for the review.

Therefore, a total of 12 original studies were included in this review [21,23,25,27-35] and an overview is presented in TABLE 2.1.

Of the 12 included studies, one was published in 1998 [27] and the remaining 11 were published after 2001 [21,23,25,28-35]. Nine studies were published in English and 3 in Chinese [27,29,33]. The included studies were carried out within different countries: 3 in China [27-29]; 1 in both China and southern Thailand [33]; 2 in India [21,35]; 2 in Turkey [30,31]; 2 in Canada [23-26]; and 2 in the USA [32,34]. Seven of the included studies were cross-sectional [27,29-31,33-35] and 4 were prospective cohort studies [21,23,25,28]. The remaining one study was a randomized controlled trial (RCT) [32], but only baseline HRQL assessment data was reported in the published article. Among the 12 studies, three studies included a comparison group either from the general population [28] or from a "healthy" non-TB sample [27,29]; one study used the normative data from the Canadian population as the reference group [23,24]; two studies included people with LTBI as controls [25,34]; one study compared TB patients with a group of chronic obstructive pulmonary disease (COPD) patients [31]; and the remaining 5 studies did not include proper comparison groups. Sample size (i.e. number of subjects included in the statistical analysis) varied among the 12 studies, from 46 to 436. Only one study [23] reported how the sample size was estimated statistically. A wide range of TB patients were included in this review: pulmonary TB and extra-pulmonary TB, active TB disease and LTBI, and current TB and previously treated TB.

To measure multiple-dimensional HRQL, a variety of instruments were involved in the included studies (TABLE 2.2). As a result, it was not possible to statistically summarize the results and thus a qualitative synthesis approach was taken for this review.

2.5.1 HRQL Instruments Used in the Included Studies

Nine studies included generic multi-dimensional instruments with or without specific single-dimensional ones, one study used a newly developed TB-specific multi-dimensional instrument [21], and two studies used a battery of single-dimensional instruments [31,33].

2.5.1.1 Generic HRQL instruments

The SF-36 was used in 6 studies with different language versions [23-28,33]. It consists of 36 items which are aggregated into 8 subscales, including physical functioning (PF), role-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role-emotional (RE), and mental health (MH) [36]. From the 8 subscales, the physical component summary (PCS) and mental component summary (MCS) scores can be also calculated [36]. Duyan et al. used the 24-item Quality of Life Questionnaire (QLQ), which covers 7 domains, including living conditions, finances, leisure, family relations, social life, health, and access to health care [30]. The 24-item QLQ was first presented by Greenley et al. in 1997 [37]. Finally, the long Medical Outcome Study (MOS) core questionnaire was used in Pasipanodya et al. [34]. This is a

generic instrument covering multiple dimensions, including physical function, social function, general health, vitality, and limitations due to physical and emotional functioning [38]. The well-known SF-36 was developed and evolved based on a subset of items from the MOS core questionnaire [38].

2.5.1.2 Specific HRQL Instruments

Dhingra and Rajpal measured HRQL with the DR-12, a new TB-specific instrument, which was developed in India and first published in 2003 [20]. It is composed of 12 items, among which 7 cover TB symptoms (i.e. cough and sputum, haemoptysis, fever, breathlessness, chest pain, anorexia, and weight loss) and 5 relate to socio-psychological and exercise adaptation (i.e. emotional symptoms/depression, interest in work, household activities, exercise activities, and social activities) [20,21]. All response options are presented on 3-point scales and equal weights are given to each item when calculating the two domain scores and the total score [20,21]. The St. George Respiratory Questionnaire (SGRQ) used in Pasipanodya et al. [34] is a widely used specific instrument designed for measuring HRQL in patients with COPD and other types of respiratory diseases. Three domain (symptom, activity, and impacts) scores and a total score can be generated [39]. It was developed at the St. George's Hospital Medical School at the UK and has been translated into various languages [39].

Yang et al. used two single-dimensional instruments, the Chinese version Symptoms Checklist 90 (SCL-90) and Social Support Rating Scale (SSRS) [29]. The SCL-90 is a

90-item symptom inventory designed mainly to evaluate a broad range of psychological problems and symptoms, including 9 dimensions: somatization, obsessive-compulsive behaviour, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism [40]. The 10-item SSRS was used to measure the self-perceived availability and use of social support services [27]. The study by Aydin and Ulusahin used two single-dimensional instruments, the General Health Questionnaire 12 (GHQ-12) and Brief Disability Questionnaire (BDQ) [31]. GHQ-12 is a short version of the GHQ-60, which was developed for screening non-psychotic psychiatric disorders in the general population [41]. The BDQ, derived from the MOS short form general health survey, is used to measure patients' physical and social disability level [42]. Marra et al. [25] used the Beck Depression Inventory (Beck-DI), along with the SF-36 and a couple of health utility instruments. The Beck-DI is a 21-item instrument, designed to measure the symptoms and degree of depression [43].

In the USA study, a series of instruments or questions were used to assess TB-infected homeless individuals' self-perceived physical health, psychological profile, emotional well-being, social support, and health care access and use [32]. Examples included the 5-item Mental Health Index (MHI-5) and the Center for Epidemiological Studies Depression Scale (CES-D) [32].

2.5.1.3 Health Utility Instrument

Health utility, one generic measure of HRQL, reflects subjective preferences for health states and also provides quantitative estimates of HRQL under certain health states [2]. The two studies [23-26] conducted in Canada applied various health utility assessment techniques among TB patients, including the Health Utility Index (HUI), EuroQol (EQ-5D), Short-Form 6D (SF-6D), Visual Analogue Scale (VAS), and Standard Gamble (SG). HUI, SF-6D, and EQ-5D are multi-attribute health status classification systems that indirectly measure preferences for health states [2]. SG and VAS are to directly obtain individuals' preferences using different techniques.

HUI currently consists of HUI-2 and HUI-3 [44]. HUI-2 and HUI-3 are derived from the same questionnaire but HUI-2 has 7 domains (sensation, mobility, emotion, cognition, self-care, pain, and fertility) and HUI-3 contains 8 domains (vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain). EQ-5D consists of 5 domains, including mobility, self-care, usual activity, pain, and anxiety/depression [2]. SF-6D is derived from a subset of SF-36 questions. It has 6 dimensions including physical functioning, role limitations, social functioning, pain, mental health, and vitality [45].

The SG is a classic technique to obtain individual preferences for health outcomes, based on the theory of von Neumann and Morgenstern [2]. In the study by Dion et al., the respondent was offered a choice between the certain outcome of a particular health state and a hypothetical gamble, with relative possibilities of perfect health and immediate

death varying. The gamble was terminated when the respondent was indifferent to the choice between the given health state and the gamble. The VAS used by Dion et al. was a 100 cm "feeling thermometer", marked at each end by word descriptions as "immediate death" and "perfect health". The respondents were asked to put a mark at the point that represents their current health status [23,24]. Similarly, a 10 cm length of horizontal line (anchored at 0cm=death and 10cm=perfect health) was used by Marra et al. [25] as VAS.

2.5.2 Psychometric Properties of HRQL Instruments in Tuberculosis

The SF-36 was used in 6 studies, and overall it showed acceptable validity and reliability. Chamla [28] validated the Chinese version SF-36 among active pulmonary TB patients and the general population in China. The reliability was tested by Cronbach's α, ranging form 0.88 to 0.97 for the eight SF-36 subscales. All 36 questions of the SF-36 had internal item consistency coefficients between 0.56 and 0.86. In Dion et al. [23,24], the reliability of SF-36 was evaluated among a mixture of TB patients, including 25 with LTBI, 17 with active TB on treatment, and 8 with previously treated TB. The internal consistency of the SF-36 responses was strong, with coefficients of 0.86-0.92 for the two summary scores and 0.73-0.94 for the subscale scores. The test-retest reliability (2-week interval) of SF-36 was tested by calculating Intraclass Correlation (ICC) coefficients: 0.66-0.79 for the two SF-36 summary scores. He et al. [33] also reported good reliability of the Chinese version SF-36 (Cronbach' $\alpha > 0.7$) among the two groups of TB patients from China and Thailand.

Validity of the SF-36 was evaluated by examining the correlations between SF-36 outcomes with other external variables, including clinical criteria, responses from other HRQL measures, and physician's evaluations. It was reported that SF-36 scores were able to discriminate between TB patients with different severity levels [21,26] and between patients at different stages of treatment (i.e. the start, middle, and end of the treatment) [21,25,28]. In Guo et al. [26], the correlations between SF-36 summary scores (PCS and MCS) and four utility instruments (SF-6D, HUI-2, HUI-3, and VAS) were tested by calculating Spearman's coefficients. SF-6D scores were strongly correlated with both PCS and MCS (0.79, 0.80), and HUI-2, HUI-3, and VAS scores were more strongly correlated with PCS (0.59, 0.66, and 0.67) than with MCS (0.37, 0.48, and 0.59). Similarly, in the study by Dion et al. [23,24], SF-36 scores were observed moderately correlated with EQ-5D and VAS scores, but poorly correlated with SG scores (Pearson coefficients < 0.2). Wang et al. [27] reported that patient-reported SF-36 scores were well correlated with physician proxy-reported Quality of Life Index (QLI) and Karnofsky Performance Status (KPS) scores, with correlation coefficients of 0.78 and 0.89 respectively. However, it was not reported which type of correlation coefficient was calculated.

The structural validity of SF-36 was tested in two studies, but the results were not consistent. In Chamala [28], factor analysis was applied to evaluate the 2-dimensional model of the SF-36. Two factors (physical health and mental health) were extracted and subjected to orthogonal rotation using the Varimax method. The observed pattern of correlations between the 8 subscales and the 2 factors supported the authors' prior

hypothesis. For example, it was reported that the 4 physical subscales (PF, RP, BP, and GH) were correlated strongly with the physical health factor, but only poorly correlated with the mental health factor. On the other hand, the 4 mental subscales (MH, RE, SF, and VT) were strongly correlated with the mental health factor, but not the physical factor. He et al. [33] used principle component analysis to test the structural validity of SF-36. However, the results showed that the 8 subscales were not well independent, and there were overlapping items between different subscales. For example, RE and RP subscales were both strongly correlated among the two groups of patients (correlation coefficient 0.82 and 0.77). Based on their findings, the authors concluded that the SF-36 did not show satisfactory construct validity in the studied TB patients.

The application of SF-36 among TB patients also revealed some problems. In the study by Dion et al. [23,24], SF-36 subscales demonstrated a remarkable ceiling effect problem. Over 50% participants with concurrent or previous TB reported the highest scores for 5 of SF-36 subscales (PF, RP, RE, BP, and SF).

Ceiling and floor effects are a common problem for the application of health utility instruments in TB. In Dion et al. [23,24], 42-53% participants reported the best possible EQ-5D health state. Guo et al. also observed ceiling and/or floor effect problems with three commonly used health utility instruments. HUI-2 and HUI-3 suffered from a serious ceiling effect problem, both in global score and single dimension level. For example, 25% of active TB patients scored 1.0 (perfect health) using the HUI-2 and 98% of them reported the best level of hearing for HUI-3. SF-6D, on the other hand, was

primarily limited by its narrow range of available utility values, from 0.30 to 1.0. Health states at the lower end may not be adequately represented by the SF-6D. Despite these problems with the application among TB patients, some positive aspects of these utility instruments were also observed. For example, these utility instruments showed moderate to strong correlations with the SF-36 responses as stated before [23,24,26]. Guo et al. [26] also reported moderate to strong agreement among SF-6D, HUI-2, HUI-3, and VAS, using ICC: the overall ICC coefficient among these 4 instruments was 0.65 and paired ICC coefficients ranged from 0.53 to 0.67. In addition, these four utility instruments were all able to discriminate between TB patients with different severity levels.

Pasipanodya et al. [34] administered the lung disease-specific SGRQ among people with treated pulmonary TB disease or LTBI. Test-retest reliability of the SGRQ was examined by ICC coefficients, 0.93 for the total score and 0.83-0.91 for subscale scores. Internal consistency was tested by Cronbach's α, at 0.93. To evaluate its validity, SGRQ responses were correlated with a previously validated MOS core questionnaire and a couple of clinical pulmonary function tests, such as the forced vital capacity (FVC). Overall, SGRQ scores and MOS scores agreed on similar health constructs and diverged on dissimilar constructs. Low but significant correlations were observed between SGRQ scores and pulmonary function test results (-0.12 to -0.29, $p<0.05$). On the other hand, a ceiling effect problem for SGRQ was observed. In both treated pulmonary TB patients and people with LTBI, the distribution of SGRQ scores was skewed toward higher HRQL. In addition, considering varied levels of reading and understanding in English in

respondents, different language versions of SGRQ were used, but the potential impact of combining results from these on HRQL outcomes was not known.

Dhingra and Rajpal [21] applied the new TB-specific instrument, DR-12, among TB patients under directly observed therapy (DOT). It was reported that, at the beginning of treatment, DR-12 scores demonstrated significant differences between pulmonary and extra-pulmonary TB patients, and between sputum positive and sputum negative patients. Over the treatment period, higher DR-12 score gains were observed among patients who positively responded to the treatment compared to those who did not. Based on these evidences, the authors came to the conclusion that DR-12 had strong construct validity in the studied population. However, the clinical criteria or indicators were not well defined in the published work. All comparisons were performed by using paired or unpaired t-tests. Potential confounders such as socio-demographic and clinical variables were not controlled in the final data analysis.

2.5.3 Impact of Tuberculosis on HRQL

Overall, active TB disease had significant and encompassing impacts on patients' HRQL. Using the SF-36, Chamla [28] found that, compared to the general population, people with active TB disease scored significantly lower on PF, RP, GH, BP, and VT ($p<0.05$), but no significant differences were observed on RE, SF, and MH subscales ($p>0.05$). In general, physical health subscales were more affected than mental ones. Dion et al. [23,24] also found active TB patients scored significantly lower in SF-36 PCS scores, but

not in MCS scores, when compared to people with LTBI and those with previously treated TB disease. In terms of health utility outcomes, Dion et al. found that active TB patients scored significantly lower in VAS (median 92.5 VS. 97.5, p=0.02) and SG (median 80.0 VS. 90.0, p=0.002) than others at the baseline assessment. However, no significant difference was observed in EQ-5D scores between active TB patients and others. It is likely that the small sample size and the heterogeneous composition of subjects could have prevented the authors from detecting the small but important differences in the sample. Wang et al. [27] found that active TB patients reported lower scores ($p<0.01$) across all SF-36 subscales than healthy non-TB people, with RP and RE being most affected. Marra et al. and Guo et al. [25,26] found that, compared to those with LTBI, people with active TB scored significantly lower at all SF-36 subscales, SF-6D, HUI-2, HUI-3, and VAS. In contrast, SF-36 scores among people with LTBI before the preventative therapy were very similar to the U.S. norm references.

In the study by Marra et al. [25], Beck-DI scores showed substantial impairment on mental well-being in active TB patients, compared to people with LTBI. However, many aspects of the Beck-DI (such as fatigue) can also be symptoms of TB and might not be necessarily indicative of mental health impairments. Aydin and Ulusahin [31] compared TB patients to COPD patients and found that TB patients had a lower prevalence of depression and anxiety and a lower level of disability, suggested by GHQ-12 and BDQ scores. The authors postulated that the chronic duration of COPD and the older age of the COPD patients may result in a higher prevalence of psychological impairments. Within TB patients, multi-drug resistant TB patients reported the worst disability level,

according to BDQ outcomes. Yang et al. [29] found that pulmonary TB patients reported more psychological symptoms listed in the SCL-90 and a lower degree of social support using SSRS compared to healthy controls. However, SCL-90 scores did not show significant correlation with SSRS scores, which is not consistent with the established relationship between social support and health [46], as discussed by the authors.

The impaired HRQL experienced by TB patients may be a reflection of socio-demographic status (e.g. age, gender, and socio-economic status) and other underlying co-morbid conditions, besides TB and its treatment. A few included studies explored the relationship between socio-demographic features and clinical factors and HRQL in TB patients. In general, the findings were consistent, but some discrepancies existed. Yang et al. [29] and Nyamatihi et al. [32] observed that females were more likely to report poorer health than males, especially on mental health problems, such as depression and anxiety. Chamla [28] and Guo et al. [26] found older people tended to have poorer HRQL than younger ones. But Duyan et al. [30] did not find significant associations between gender, age and HRQL in TB patients. On the other hand, they [30] found that better HRQL was correlated with higher income, higher education, better housing conditions, better social security, and closer relationships with family members and friends. Some clinical factors that were observed to correlate with poorer HRQL in TB patients include size of pulmonary TB infection, duration of TB disease, reactivation of previous TB infection, number of symptoms before treatment, development of hemoptysis, hospitalization, underlying chronic conditions, anemia, and count of white blood cells before treatment [27,28].

2.5.4 Effect of Anti-Tuberculosis Treatment on HRQL

Chamla [28], Dhingra and Rajpal [21], and Marra et al. [25] prospectively measured active TB patients' HRQL at the start, middle, and end of treatment. In the study by Chamla [28], after the anti-TB treatment, significant improvement was observed in all physical health subscales of the SF-36 (PF, RP, BP, and GH, $p<0.05$); two mental health subscales, RE and SF ($p<0.05$), improved significantly, but not VT and MH ($p>0.05$). During the treatment, RP, VT and MH scores decreased after the initial 2 months and but showed overall improvement at the end of the treatment, while all other subscale scores showed gradual increase over the treatment [28]. Dhingra and Rajpal [21] observed a gradual improvement on DR-12 scores in active TB patients over the course of the treatment. Overall, a more identifiable improvement was observed in symptom scores than that in socio-psychological and exercise adaptation scores. Consistently, Marra et al. [25] also found significant HRQL improvement in active TB patients over the 6 months of treatment, using SF-36 and Beck-DI.

Although anti-TB treatment improved HRQL overall, active TB patients still had poorer HRQL at the end of the treatment compared to the general population or people with LTBI, especially in psychological well-being and social functioning. Chamla [28] observed that, at the end of the treatment, active TB patients still scored significantly lower at RP, VT, and MH subscales compared to general population comparisons. Marra et al. [25] found that, after the 6 month of treatment, active TB patients scored significantly lower at SF-36 PCS and MCS summary scores compared to people with

LTBI. An interesting finding by Marra et al. [25] is that, after the preventive treatment, MCS scores among people with LTBI decreased significantly, while PCS scores remained unchanged. Pasipanodya et al. [34] measured HRQL among pulmonary TB patients who completed at least 20 weeks of treatment, using the SGRQ. Compared with those with LTBI, treated TB patients had lower SGRQ scores. Those with better lung functions and/or born in the U.S. (against foreign-born) tended to have better HRQL outcomes. No gender difference was observed in SGRQ scores.

Muniyandi et al. [35] assessed the HRQL in a sample of previous TB patients one year after successful completion of treatment. 40% of these people reported persistent symptoms, such as breathlessness, cough, chest pain, and occasional fever. The authors calculated three SF-36 component scores: the physical well-being, mental well-being, and social well-being. Based on their results, there was no gender difference on physical well-being score; but females scored much lower at mental and social well-being scores. Compared with younger people, older ones had significantly lower physical and mental well-being scores, but not the social score. They also presented the U.S. general population norms for the three component scores and concluded that TB patients' HRQL returned to normal level one year after the completion of treatment. However, the way of calculating the three SF-36 component scores is not commonly seen in literatures, and the reference regarding the U.S. general population norms provided in the published paper cannot be located.

2.6 DISCUSSION

HRQL has been appreciated as an important health outcome measure in clinical research. We identified 12 original studies where multi-dimensional HRQL was assessed among people with TB disease or infection using structured instruments around the world. We found that TB and its treatment have a significant impact on patients' quality of life from various aspects and this impact tends to persist for a long time even after the successful completion of treatment and the microbiological 'cure' of the disease.

The results suggest that TB disease has a negative and encompassing impact on active TB patients' self-perceived health status in physical, psychological, and social aspects. Overall, the anti-TB treatment showed positive effect on improving patients' HRQL. It appeared that physical health seemed to be more affected by the disease but improved more quickly after the treatment, while the impairment on mental well-being tended to persist for a longer term [21,28]. However, even after the active TB patients successfully completed the treatment and were considered microbiologically 'cured', their HRQL remained poor as compared to the general population [23-25,28]. The ongoing HRQL impairment may be partly due to the persistent physical symptoms and residual physiological damages from the disease and/or the treatment. Furthermore, a few qualitative studies [9-14,16-18] have shown that the social stigma attached to the diagnosis of TB in some cultures is significant. People with TB may feel isolated from their family and friends or experience the fear and anxiety of being known by others

about their diagnosis. All these consequential impairments also need to be ‘cured’ and may take a long recovery time.

Most studies have focused on assessing HRQL in active TB patients. Although people with LTBI do not present with clinical disease or symptoms, they are likely to be subjected to the same social and psychological impacts as active TB patients. The knowledge of a deadly and stigmatized disease lying dormant in his/her body may also induce anxiety and fear. As Marra et al. [25] observed that, after receiving 6 months of preventive therapy with isoniazid, the mental well-being of people with LTBI decreased significantly.

HRQL assessment in TB research is still a new area, and a valid and reliable TB-specific instrument is much needed. Currently, a wide range of HRQL instruments were utilized in the literature. The SF-36 was the most frequently used instrument and it appeared to be a valid and reliable tool to be used in TB. Although the SF-36 has been used extensively to assess both population health and specific health conditions for various medical conditions, as a generic health assessment instrument, it offers little information to help understand the unique experiences among TB patients, such as social stigma and anti-TB treatment related ADRs.

Our review identified one TB-specific HRQL instrument, DR-12, which was developed in India [20,21]. Unfortunately, its validation study was not conducted in a systematic fashion and the current evidence provided was not convincing. Further applications and

appropriate methodologies are needed to show DR-12 is a psychometrically sound HRQL instrument feasible and valid for TB patients. In addition, the DR-12 is actually designed specifically for pulmonary TB patients, judging from its item content. TB can affect almost any part of the human body, and in Canada, about 40% of active TB diseases would present as extra-pulmonary TB [47]. Different types of TB disease would have very different clinical presentations and affect people's function differently. This may be a challenge when developing a TB-specific HRQL instrument.

It should be also noted that most TB patients have very different cultural and socio-demographic backgrounds compared with the population in which many of these instruments were originally developed. Also, in the studies done in Canada and the USA [24-26,32,34], most TB patients were foreign-born and the instruments were normally self-administered in the English language which would not have been the respondents' first language. Thus, the results of these studies may not be valid if careful translation and cultural adaptation of the instrument was not done to accommodate the multi-cultural population.

Particular attention should be given to some methodological issues on assessing HRQL among people with active TB disease or LTBI. To comprehensively examine the impact of TB and its treatment on patients' HRQL, it is very important to include a proper comparison group from a similar demographic and socio-economic background. When conducting the study, researchers are recommended to seek statistical consultation regarding proper sample size estimating, missing data handling, and adjusting for

potential confounders, such as socio-demographic status and presence of co-morbidities. Another concern is the lack of interpretation of HRQL outcomes in terms of clinical meaningfulness. Statistical significance is a useful way to interpret the result, but it fails to relate the HRQL outcome with clinical relevance. As such, more work needs to be done to relate changes in HRQL assessment in TB to concepts such as the minimal clinical difference [48,49].

2.7 CONCLUSIONS

Our review of the literature shows that TB diminishes patients' HRQL, as measured by various instruments. However, due to the heterogeneity of HRQL measurements, it was difficult to assimilate results across studies. A few studies used the SF-36 which appeared to be a valid instrument in the measurement of HRQL in TB. Other instruments require further psychometric testing to determine their suitability in measurement in this context. Our review suggests that HRQL assessment in people with TB is a growing research area and a psychometrically sound TB-specific HRQL instrument is lacking. A critical step in the future would be to design an applicable, reliable, and valid TB-specific HRQL instrument. Particular attention should be given to address the methodological issues when conducting a HRQL assessment study in TB patients.

TABLE 2.1: Overview of Included Studies

Reference	Publication Year	Study Design	Study Location and year	Patients	Comparison group	HRQL assessment
21	2005	Prospective cohort	India (2002)	76 pulmonary or extra-pulmonary TB patients on DOTS	No.	DR-12; at baseline, 2 months and the end of treatment
23,24	2002, 2004	Prospective cohort	Montreal, Canada (1999-2000)	46-50 patients (mixture of active TB, latent TB and previously treated TB)	Normative data for Canadian population used as reference.	SF-36 (English and French version), EQ-5D, VAS and SG; at 0, 1 week and 2 weeks
25,26	2008	Prospective cohort	Vancouver, Canada (2005-2006)	84 active TB patients for the cross-sectional study; 75-85 active TB patients for the longitudinal study	78 latent TB patients for the cross-sectional study; 70-75 latent TB patients for the longitudinal study	SF-36 (and SF-6D), HUI-2, HUI-3, Beck-DI and VAS; at baseline, 3 months and completion of treatment
27	1998	Cross-sectional	China (1996)	228 hospitalized pulmonary TB patients	228 healthy controls (matched for age and sex).	SF-36 (Chinese version)
28	2004	Prospective cohort	China (2001-2002)	102 newly diagnosed TB patients	103 non-TB controls from the general population (matched for age and sex).	SF-36 (Chinese version); at baseline, 2 months and the end of treatment
29	2003	Cross-sectional	China (2001)	132 registered pulmonary TB patients	71 healthy controls	SCL-90 and SSRS
30	2005	Cross-sectional	Turkey (2003-2004)	120 TB patients hospitalized for at least 1 month	No.	A 24-item QLQ

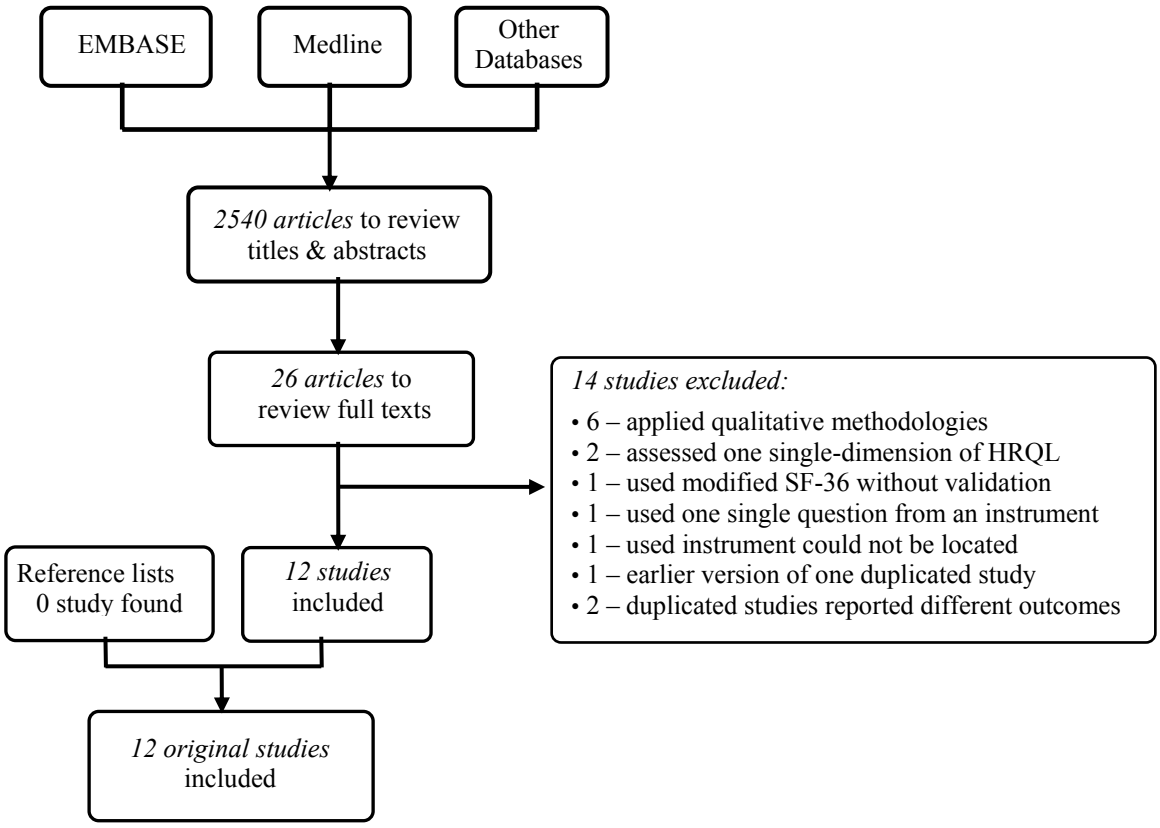
TABLE 2.1: Overview of Included Studies (Continued)

Reference	Publication Year	Study Design	Study Location and year	Patients	Comparison group	HRQL assessment
31	2001	Cross-sectional	Turkey (1999)	42 newly diagnosed TB, 38 defaulted TB, and 39 multi-drug resistant TB patients	38 COPD patients (chronic bronchitis and emphysema)	GHQ-12 and BDQ
32	2005	RCT baseline HRQL assessment	Los Angeles, USA (1997-2002)	415 homeless adults with latent TB infection enrolled in a TB-adherence trial	No.	Multiple instruments and questions
33	2005	Cross-sectional	China and Thailand	84 pulmonary TB patients (52 from China, 32 from Thailand)	No.	SF-36 (Chinese and Thai language version)
34	2007	Cross-sectional	Texas, USA (2005-2006)	105 pulmonary TB patients who completed at least 20 weeks of treatment	207 people with latent TB infection.	MOS core questionnaire and SGRQ
35	2007	Cross-sectional	India	436 TB patients were assessed one year after they successfully completed treatment and cured	No.	SF-36

TABLE 2.2: HRQL Instruments Used by Included Studies

HRQL Instruments	References
Generic	
Short-Form 36 (SF-36)	23-28,33,35
MOS core questionnaire	34
A 24-item Quality of Life Questionnaire (QLQ)	30
Specific	
DR-12	20,21
St. George Respiratory Questionnaire (SGRQ)	34
Symptoms Checklist 90 (SCL-90)	29
Social Support Rating Scale (SSRS)	29
General Health Questionnaire 12 (GHQ-12)	31
Brief Disability Questionnaire (BDQ)	31
Beck Depression Inventory (Beck-DI)	25,26
Mental Health Index (MHI-5)	32
Center for Epidemiological Studies Depression Scale (CES-D)	32
Health utility	
Health Utility Index – 2 & -3 (HUI-2 & -3)	25,26
Short-Form 6D (SF-6D)	25,26
EuroQol (EQ-5D)	23,24
Visual Analogue Scale (VAS)	23-26
Standard Gamble (SG)	23,24

FIGURE 2.1: Process of Literature Searching



2.8 REFERENCES

1. Guyatt GH, Feeny DH, Patrich DL. Measuring health-related quality of life. Ann Intern Med 1993;118:622-629.

2. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (Eds). Methods for the economic evaluation of health care programmes. 3rd edition. Oxford University Press; 2005.

3. Sherbourne CD, Sturm R, Wells KB. What outcomes matter to patients? J Gen Intern Med 1999;14:357-363.

4. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care 1989;27(Suppl 3):217-232.

5. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. JAMA 1999;282:677-686.

6. Blumberg HM, Burman WJ, Chaisson RE, Daley CL, Etkind SC, Friedman LN, Fujiwara P, Grzemska M, Hopewell PC, Iseman MD, Jasmer RM, Koppaka V, Menzies RI, O'Brien RJ, Reves RR, Reichman LB, Simone PM, Starke JR, Vernon AA. American Thoracic Society/Centers for Disease Control and

Prevention/Infectious Diseases Society of America: treatment of tuberculosis. Am J Respir Crit Care Med 2003;167:603-662.

7. Forget EJ, Menzies D. Adverse reactions to first-line antituberculosis drugs. Expert Opin Drug Saf 2006;5:231-249.

8. Marra F, Marra CA, Bruchet N, Richardson K, Moadebi S, Elwood RK, FitzGerald JM. Adverse drug reactions associated with first-line anti-tuberculosis drug regimens. Int J Tuberc Lung Dis 2007;11:868-875.

9. Liefooghe R, Michiels N, Habib S, Moran MB, De Muynck A. Perception and social consequences of tuberculosis: a focus group study of tuberculosis patients in Sialkot, Pakistan. Soc Sci Med 1995;41:1685-1692.

10. Kelly P. Isolation and stigma: the experience of patients with active tuberculosis. J Community Health Nurs 1999;16:233-241.

11. Long NH, Johansson E, Diwan VK, Winkvist A. Fear and social isolation as consequences of tuberculosis in VietNam. A gender analysis. Health Policy 2001;58:69-81.

12. Macq J, Solis A, Martinez G, Martiny P, Dujardin B. An exploration of the social stigma of tuberculosis in five 'municipios' of Nicaragua to reflect on local interventions. Health Policy 2005;74:205-217.

13. Hansel NN, Wu AW, Chang B, Diette GB. Quality of life in tuberculosis: patient and provider perspectives. Qual Life Res 2004;13:639-652.

14. Marra CA, Marra F, Cox VC, Palepu A, Fitzgerald JM. Factors influencing quality of life in patients with active tuberculosis. Health Qual Life Outcomes 2004; 20: 58.

15. Chang B, Wu AW, Hansel NN, Diette GB. Quality of life in tuberculosis: a review of the English language literature. Qual Life Res 2004;13:1633-1642.

16. Natani GD, Jain NK, Sharma TN, Gehlot PS, Agrawal SP, Koolwal S, Gupta RB, Agnihotri SP. Depression in tuberculosis patients: correlation with duration of disease and response to anti-tuberculous chemotherapy. Indian Journal of Tuberculosis 1985;32:195-198.

17. Kelly-Rossini L, Perlman DC, Mason DJ. The experience of respiratory isolation for HIV-infected with TB. J Assoc Nurses AIDS Care 1996;7:29-36.

18. Rajeswari R, Muniyandi M, Balasubramanian R, Narayanan PR. Perceptions of tuberculosis patients about their physical, mental and social well-being: A field report from south India. Soc Sci Med 2005;60:1845-1853.

19. de Figueiredo AA, Lucon AM, Srougi M. Bladder augmentation for the treatment of chronic tuberculous cystitis. Clinical and urodynamic evaluation of 25 patients after long term follow-up. Neurourol Urodyn 2006;25:433-440.

20. Dhingra VK, Rajpal S. Health related quality of life (HRQL) scoring in tuberculosis. Indian J Tuberc 2003;50:99-104.

21. Dhingra VK, Rajpal S. Health related quality of life (HRQL) scoring (DR-12 score) in tuberculosis--additional evaluative tool under DOTS. J Commun Dis 2005;37:261-268.

22. Ji Q, Sun Y-J, Liu Q-G, Gong F-Q, Jin Y, Wu J. Correlation of mental status and social function with quality of life in patients with tuberculosis. Zhongguo Linchuang Kangfu 2006;10:21-24.

23. Dion MJ, Tousignant P, Bourbeau J, Menzies D, Schwartzman K. Feasibility and reliability of health-related quality of life measurements among tuberculosis patients. Qual Life Res 2004;13:653-665.

24. Dion MJ, Tousignant P, Bourbeau J, Menzies D, Schwartzman K. Measurement of health preferences among patients with tuberculous infection and disease. Med Decis Making 2002;22(Suppl 5):102-114.

25. Marra CA, Marra F, Colley L, Moadebi S, Elwood RK, Fitzgerald JM. Health-related quality of life trajectories among adults with tuberculosis: differences between latent and active infection. Chest 2008;133:396-403.

26. Guo N, Marra CA, Marra F, Moadebi S, Elwood RK, Fitzgerald JM. Health state utilities in latent and active tuberculosis. Value Health 2008;11:1154-1161.

27. Wang Y, Lii J, Lu F. Measuring and assessing the quality of life of patients with pulmonary tuberculosis. Chin J Tuberc Respir Dis 1998;21:720-723.

28. Chamla D. The assessment of patients' health-related quality of life during tuberculosis treatment in Wuhan, China. Int J Tuberc Lung Dis 2004;8:1100-1106.

29. Yang L, Wu DL, Guo HG, Liu JW. A study of the psychological and social factors in patients with pulmonary tuberculosis. Chin J Tuberc Respir Dis 2003;26:704-707.

30. Duyan V, Kurt B, Aktas Z, Duyan GC, Kulkul DO. Relationship between quality of life and characteristics of patients hospitalised with tuberculosis. Int J Tuberc Lung Dis 2005;9:1361-1366.

31. Aydin IO, Ulusahin A. Depression, anxiety comorbidity, and disability in tuberculosis and chronic obstructive pulmonary disease patients: applicability of GHQ-12. Gen Hosp Psychiatry 2001;23:77-83.

32. Nyamathi A, Berg J, Jones T, Leake B. Predictors of perceived health status of tuberculosis-infected homeless. West J Nurs Res 2005;27:896-910.

33. He CY, He LM, Li MH. Application of SF-36 scale on pulmonary tuberculosis patients in Yunnan province of China and southern Thailand. Chin J Epidemiol 2005;26:187-189.

34. Pasipanodya JG, Miller TL, Vecino M, Munguia G, Bae S, Drewyer G, Weis SE. Using the St. George respiratory questionnaire to ascertain health quality in persons with treated pulmonary tuberculosis. Chest 2007;132:1591-1598.

35. Muniyandi M, Rajeswari R, Balasubramanian R, Nirupa C, Gopi PG, Jaggarajamma K, Sheela F, Narayanan PR. Evaluation of post-treatment health-related quality of life (HRQoL) among tuberculosis patients. Int J Tuberc Lung Dis 2007;11:887-892.

36. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med 2001;33:350-357.

37. Greenley JR, Greenberg JS, Brown R. Measuring quality of life: a new and practical survey instrument. Soc Work 1997;42:244-254.

38. Hays RD, Sherbourne CD, Mazel R. User's manual for the Medical Outcomes Study (MOS) core measures of health-related quality of life. http://www.rand.org/pubs/monograph_reports/MR162/ (accessed at June 23, 2010).

39. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. Respir Med 1991;85(Suppl B):25-31.

40. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale--preliminary report. Psychopharmacol Bull 1973;9:13-28.

41. Golderberg D, Williams P. A user's guide to the General Health Questionnaire. Windsor, UK: NFER-Nelson, 1988.

42. Stewart AL, Hays RD, Ware JE. The MOS short form general health survey: reliability and validity in a patient population. Med Care 1988;26:724–732.

43. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation, 1996.

44. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health Qual Life Outcomes 2003;16:54.

45. Brazier J, Roberts J, Deveril M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271-292.

46. House JS, Landis KR, Umberson D. Social relationships and health. Science 1988;241:540-545.

47. Health Canada. Canadian tuberculosis standards. 6th ed. 2007. http://www.phac-aspc.gc.ca/tbpc-latb/pubs/pdf/tbstand07_e.pdf (accessed at June 23, 2010).

48. Sloan JA, Symonds T, Vargas-Chanes D, Friedly B. Practical guidelines for assessing the clinical significance of heath related quality of life changes within clinical trials. Drug Inf J 2003;37:23–31.

49. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999;15:141-155.

CHAPTER 3 HEALTH STATE UTILITIES IN LATENT AND ACTIVE TUBERCULOSIS: A CROSS-SECTIONAL ANALYSIS

3.1 FOREWORD

This chapter is a cross-sectional analysis of health state utility in a sample of latent and active tuberculosis patients and additional comparison of four utility instruments (SF-6D, HUI-2, HUI-3, and VAS). The content of this chapter has been published in the journal of *Value in Health*.[2]

As the first author, the candidate developed the hypotheses, performed data analyses, and wrote the final manuscript. Co-authors of this study included *Dr. Carlo A. Marra, Dr. Fawziah Marra, Suzana Moadebi, Dr. R. Kevin Elwood,* and *Dr. J. Mark FitzGerald*.

3.2 INTRODUCTION

Although effective chemotherapy is available, worldwide tuberculosis (TB) remains a major public health problem, with approximately 1/3 of the world's population infected [1]. In North America, since mortality due to TB is rare, the avoidance of morbidity and maintenance of patients' quality of life has become the goal of TB management [2, 3].

[2] A version of this chapter has been published in *Value in Health*. Guo N, Marra CA, Marra F, Moadebi S, Elwood RK, Fitzgerald JM. Health state utilities in latent and active tuberculosis. *Value Health* 2008; 11:1154-1161.

Health-related quality of life (HRQL) has been introduced into medical practices and evaluated in many diseases including TB. There is increasing evidence to show that TB has substantial and encompassing impacts on patients' quality of life [4-9]. Furthermore, clinical measures often correlate poorly with daily well-being and function, the areas in which patients are most interested and familiar.

Two basic approaches to assessing HRQL are available: generic and disease-specific instruments [10]. Generic instruments capture the common and important aspects of health, while specific instruments are designed to focus on unique problems associated with a disease. One category of generic instruments is "preference-based" or "utility measure" and examples of such are the EuroQol 5D (EQ-5D), the Health Utilities Index 2 and 3 (HUI-2 and HUI-3) and the Short-Form 6D (SF-6D). These instruments summarize HRQL into a single index number, anchored at 0 (death) and 1.0 (perfect health) and are based on societal health preferences. Scores can then be incorporated into cost-utility analyses (a special type of cost-effectiveness analysis or "CEA") through the calculation of quality adjusted life years (QALYs) that combine the utility score for a health state (quality of life) with the duration of time spent in the particular health state (quantity of life).

Recently, there have been many CEAs of the treatment of, and/or screening for, TB infection in populations such as intravenous drug users [11, 12], the elderly [13], tuberculin reactors stratified by age [14, 15], immigrants [16-19], and new blood-based diagnostic strategies [20-22]. So far, none have evaluated TB health outcomes in the

holistic way recommended by the influential U.S. Panel on Cost-Effectiveness in Health and Medicine [23], namely, QALYs constructed from utilities elicited from TB patients. The main reason for the lack of QALYs as outcome measures in TB economic evaluations is the dearth of TB related utility data in the literature.

Considering the void in the literature regarding health state utility values for active TB and latent tuberculosis infection (LTBI), the primary purpose of this paper was to characterize and compare utility scores between the HUI-2, HUI-3, and SF-6D in a sample of participants with either active TB or LTBI.

3.3 METHODS

3.3.1 Study Setting and Subjects

Participants were recruited from the TB Clinic at the BC Centre for Disease Control (BCCDC). To be eligible for this study, the participant had to have been diagnosed with either active TB or LTBI within 2 months of study entry. Subjects were excluded if: (1) they were less than 18 years old; (2) they were not taking medication for active TB or LTBI; or (3) they were judged to be incapable of answering the questionnaires. This study received ethics approval from the University of British Columbia Behavioral Research Ethics Board and informed signed consent was obtained from each participant.

During the first clinic visit, relevant participants' information was collected, including: (1) socio-demographics, e.g. age, gender, education, and co-morbid conditions; (2) TB symptoms, medication use, adverse events and self-reported TB symptom severity and control; (3) responses to the HRQL questionnaires (SF-36, HUI-2/HUI-3, and a visual analogue scale). All questionnaires were self-administered and participants were permitted to take them home for completion. Investigators followed up by phone-call to ensure that questionnaires were returned within a week.

3.3.2 Assessment of Health-related Quality of Life

3.3.2.1 Short Form 36 and Short Form 6D

The Short-Form 36 Health Survey (SF-36) is one of the most widely used generic health profile instruments [24]. Eight domain scores (physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health) and two summary scores, the physical component summary (PCS) and the mental component summary (MCS), are calculated to describe health status. However, no single, preference-based index score can be derived directly from this widely used HRQL instrument. To obtain a preference-based utility score from it, Brazier et al. restructured the SF-36 into a health state classification system with 6 attributes called the SF-6D (which includes physical functioning, role limitations, social functioning, pain, mental health, and vitality). Each

attribute has 4 to 6 levels, and therefore the SF-6D describes 18,000 health states. Scoring is based on responses from a random sample drawn from the UK general population.

3.3.2.2 Health Utilities Index

The HUI is a family of multi-attribute health status classification systems which is currently composed of the HUI-2 and HUI-3 [25]. HUI-2 and HUI-3 utility scores are derived from the same questionnaire, but they contain different health dimensions with various levels in each, and they employ different samples and scoring models to obtain preference scores. HUI-2 has seven health attributes (sensation, mobility, emotion, cognition, self-care, pain, and fertility), each with 3 to 5 levels [26], while HUI-3 consists of eight attributes with 5 to 6 levels within each. In total, HUI-2 defines 18,000 health states and HUI-3 defines 972,000 health states [27]. Preference scores for both HUI-3 and HUI-2 health states were based on Canadian samples using the standard gamble and a visual analogue scale (VAS) approach.

In this study, both the HUI and SF-36 questionnaires used “the past 4 weeks” as the response time frame to minimize potential differences in recall.

3.3.2.3 Visual Analogue Scale

In this study, a horizontal line, 10cm in length, anchored at “0” (Death) at the left-end and “1” (Perfect health) at the right-end was used. Each respondent was asked to mark on

the line the point that they felt represented their current health state.

3.3.3 Statistical Analysis

Scores for the SF-6D, HUI-3 and HUI-2 were calculated for each patient from the two questionnaires, SF-36 and HUI. All patients who had the completed the SF-6D, HUI-3, HUI-2 and VAS scores were included in the analysis. Descriptive summary statistics were used to describe demographic characteristics. Chi square tests, student's t-test or ANOVA, as appropriate, were used to examine the demographic differences between active and latent patients. In each group, patients who were included and those who were excluded were compared, as well.

The SF-36 domain scores and utility scores were summarized by groups numerically and graphically. Because the utility scores were usually not normally distributed, median value and inter-quartile range (IQR) were also presented. Differences between latent and active patients on SF-36 scores and utility values were tested using simple linear regression, adjusting for the available demographic variables. Active TB patients were then categorized by self-graded TB severity and control, and each instrument's ability to discriminate between severity subgroups was tested.

Two-way ANOVA and paired two-sample t-tests were performed to examine the differences among the 4 utility instruments within individuals. The ceiling and floor effects for the three indirect utility instruments (HUI-2, HUI-3 and SF-6D) were

examined from two aspects: the frequency of possible minimum and maximum scores for each instrument; and the proportion of respondents who reported the best and worst level within each single health attribute for each instrument.

Spearman's rho was calculated to assess the correlation between various measures. The correlation coefficients were explained as follows. -0.30 to 0.30: weak correlation; -0.49 to -0.30 and 0.30 to 0.49: moderate correlation; <-0.50 and >0.50: strong correlation [28]. Agreement among the utility scores was assessed using the intra-class correlation coefficient (ICC) generated from a two-way mixed-effect ANOVA model where the patient effect was random and the instrument effect was fixed. The single measure ICCs were used and interpreted according to the following guidelines. ICC <0.40: poor agreement; 0.40 to 0.75: moderate to good agreement; >0.75: excellent agreement [29]. All p values were two-sided, and $p < 0.05$ was considered to be statistically significant.

3.4 RESULTS

3.4.1 Description of Study Sample

In a consecutive fashion, over a twelve month period we approached 147 LTBI and 133 active TB patients to participate in the study. Of these, 119 LBTI and 114 active tuberculosis patients were deemed eligible and consented to participate in the study. Of these, 27 (20 LBTI and 7 active) withdrew from the study prior to completing the questionnaires. Subsequent to enrollment , three patients reported as having active TB were excluded as they did not fit the entry criteria of having active TB (under medical

surveillance for previously active TB; not currently active; no documentation of prior medications; no growth on smear and cultures). Thus 104 active TB and 99 LBTI participants were enrolled in the study. Of these, 25 (12.3%) participants failed to fully complete the SF-36, 15 (7.4%) did not fully complete the HUI, and 13 (6.4%) had no VAS results. In total, 162 (79.8%) TB patients (78 LTBI and 84 active TB) with available SF-36 and all four utility scores (HUI-2, HUI-3, SF-6D and VAS) were included in the analysis. In each group, there were no significant differences on the studied demographic characteristics between patients included and those excluded.

The demographic features of both the active TB and LTBI participants are described in TABLE 3.1. Active TB participants were significantly older than LTBI participants (49.0 vs. 36.3 yrs, $p<0.01$). Subjects with active TB were more likely to have co-morbid conditions (45% vs. 24%, $p=0.02$) and report alcohol use compared to those with LTBI (86% vs. 69%, $p=0.01$). Overall, of the 162 TB patients in this study, 43% were males, 87% were born outside Canada, and 56% were originally Asians or Pacific Islanders.

3.4.2 SF-36 Scores

In comparison to those with LBTI, subjects with active TB had a much lower average scores on all 8 domains of the SF-36 (TABLE 3.2, all p-values < 0.05), after adjusting for the demographic differences between the two groups. Overall, older participants tended to score lower than those who were younger, and no gender difference was observed on SF-36 domain scores in this sample.

3.4.3 Global Utility Scores

TABLE 3.3 summarizes the four sets of global utility scores for LBTI and active TB participants, respectively. For each utility instrument assessment, significantly lower average scores were observed in active TB participants compared to those with LTBI, after accounting for the impact of the demographic differences (all $p<0.01$). In active TB participants, the four preference-based instruments (the HUI-2, HUI-3, SF-6D and VAS) yielded significantly different global score within individuals, with HUI-2 having the highest average score overall, followed by HUI-3, SF-6D and VAS.

FIGURE 3.1~3.4 show the distributions of the global scores among active TB participants. SF-6D scores were distributed normally, but with limited available range, and the observed lowest score was 0.32. In contrast, HUI-2 and -3 scores covered a wider range: -0.26-1.00 for HUI-3 and 0.13-1.00 for HUI-2. However, HUI-2 and -3 scores were highly skewed toward 1.00 (perfect health), demonstrating a ceiling effect. For the HUI-2 and HUI-3, respectively, 25% and 21% of active participants reported scores of 1.00 (perfect health). In contrast, 2.4% participants scored 1.00 for SF-6D. VAS scores were not as skewed as HUI-2 and -3, and spanned the entire scale range, from 0.08 to 1.00.

TABLE 3.4 describes the proportion of respondents at the best and worst level within each single health dimension for HUI-2, HUI-3 and SF-6D. For HUI-2 and -3, less than 4% of active TB participants reported the worst level in all health dimensions, but large

proportions of participants were at the best level. There are three common health dimensions in the HUI-2 and HUI-3 which are emotion, cognition, and pain. In the pain dimension, about 50% of the respondents chose the best possible level, while in the cognition and emotion dimensions, respondents chose the best level of the HUI-3 domain much less than the best level of the HUI-2 domain. For SF-6D dimension scores, except for vitality, all had at least 22% at the best level. For example, 42% and 33 % were at the best level in the pain and role limitations dimensions, respectively. On the other hand, 63%, 29% and 14% were at the worst level in role limitation, vitality, and social functioning, respectively. Even though the role limitation dimension has four levels in total, 96% of participants reported either the best or worst level, leaving a gap at the intermediate levels.

In TABLE 3.5, SF-36 summary scores and each of the utility scores were summarized by the level of self-reported disease severity. Each instrument demonstrated a clear monotonically decreasing trend in scores with increasing severity level. Among the four sets of utility scores, only HUI-3 scores were not statistically different between adjacent severity subgroups.

TABLE 3.6 presents the Spearman's coefficients for correlation between SF-36 summary scores and the 4 global preference-based scores. Strong correlation was observed between the 4 utility instrument scores (0.52-0.86, p-values<0.0001). In terms of the relationship with SF-36, SF-6D scores were strongly correlated with both PCS and MCS (0.79, 0.80),

while HUI-3 and -2 were strongly correlated with PCS (0.59, 0.66), but only moderately correlated with MCS (0.37, 0.48).

The overall ICC among the 4 utility instrument scores was 0.65, indicating good agreement. ICCs (and 95% CIs) between each paired utility scores were presented in TABLE 3.7. There was excellent agreement between HUI-2 and HUI-3 (0.84). The agreement between all other paired utility scores was good (0.53-0.67).

3.5 DISCUSSION

This study has generated much needed evidence regarding preference-based utility scores for subjects with active TB and LTBI that can be integrated into cost-utility analyses of TB diagnostic and treatment interventions. In addition, this study provides some comparative information regarding the different methods of estimating health utilities and aids in the choice of measure to be included in primary studies. Finally, the results show that there are significant problems with all of the instruments in terms of floor and ceiling effects especially when each domain is considered.

Other authors have estimated preferences for health states in TB either through expert opinion [15, 30] or by surveying the general public [31]. However, only one other study has attempted to measure health preferences from patients with LTBI and active or recently cured, active TB [7]. These authors relied on a small, heterogeneous sample of LTBI (n=25), active TB (n=17) and previous active TB (n=8) participants to derive

health state utility values thus limiting the generalizability of their findings. Furthermore, the authors utilized the standard gamble (SG), a visual analog scale (VAS), and the EQ-5D to elicit health state preferences and the SF-36 as a general health measure. In general, their findings were that those with concurrent active TB rated hypothetical health states (mild, moderate and severe) lower than those with LBTI and previously active TB. Also SG scores correlated poorly with the EQ-5D and the SF-36, and moderately with the VAS. The authors concluded that the SG was likely a more comprehensive measure of health-related quality of life since it is not limited to a small number of domains and levels like questionnaire based instruments are and provides a more holistic assessment of quality of life. However, there are recent publications that put the validity and the reliability of the SG into question when used in specific patient populations [32, 33]. As such, it is not clear that the SG is appropriate for the elicitation of utility values in all cases.

From the same dataset, the authors published another paper whose purpose was to evaluate the feasibility and reliability of the EQ-5D and the SF-36 [6]. They found that the proportion of acceptable SF-36 questionnaires was 78% versus 90% for the EQ-5D. In addition, they found a high degree of internal consistency for the SF-36 and high test-retest reliability. Similar to our study, the authors were concerned about the use of health status questionnaires in a multicultural sample which are not adapted and tested in the respondent's native language. However, since the authors found a low frequency of inconsistent responses and high internal consistency, they concluded that it was valid to use these questionnaires. Similarly, authors of other papers [34, 35] found similar high

feasibility and strong psychometric properties for the SF-36 in non-native English language speakers.

It was difficult to compare the VAS scores from our sample to that of Dion et al's due to the lack of stratification in the presentation of their results into subjects with LTBI and active TB. However, we found the VAS to be much more highly correlated with the PCS (r=0.65) and the MCS (r=0.58) than Dion et al. whose correlations were 0.54 and 0.32, respectively.

Within our results, the utility global scores were significantly different with the HUI-2 generating the highest and the SF-6D providing the lowest mean scores. The distribution of SF-6D scores was much more symmetrical than that of HUI-3 or HUI-2, but within a very limited scale range, from 0.32 to 1.00. In contrast, HUI-3 scores spanned approximately twice the range of SF-6D scores, and HUI-2 distributed from 0.13 to 1.00. Although having a wider scale range than SF-6D, HUI-3 and HUI-2 scores were highly skewed toward 1.00, with 25% and 21% participants reporting a HUI-2 and HUI-3 score of 1.0, respectively, indicating a possible ceiling effect in active TB participants. There was no evidence to demonstrate a ceiling or floor effect for SF-6D global utility scores, since only 2.4% of participants reported perfect health and none reported the worst health state. However, a dimension specific analysis SF-6D yielded possible floor effects. The most notable result was in the role limitation dimension. In the pain dimension, all three instruments demonstrated a significant ceiling effect, with 42% at the best level for SF-6D and 50% for HUI-2 and HUI-3. Compared with SF-6D, all HUI-2 and HUI-3 single

dimensions demonstrated ceiling effects with very limited floor effects.

Discrepancies between scores of the indirect utility instruments have been observed and discussed in other patient populations [36, 37]. These discrepancies may arise due to the internal differences of the utility instruments, such as the health dimensions covered, the methods of obtaining preferences (such as the standard gamble or the time trade-off techniques), and the application of the instruments. Each instrument focuses on different health dimensions with various levels within each dimension and can have different recall periods.

However, despite these differences some similarities were observed as well. The correlation between these instruments was from moderate to strong, suggesting that they are likely measuring similar concepts of health. The four utility instruments were able to discriminate TB participants with different severity levels. In general, the average utility scores showed a decrease with the increasing disease severity.

In conclusion, health state utility values for active and LTBI have been determined using different instruments. The three measures, the HUI-2, HUI-3 and SF-6D, did not generate identical utility scores in active or LTBI participants. The biggest concern with the HUI-2 and HUI-3 was the ceiling effect, and SF-6D was limited by its narrow range of utility values available and the potential floor effect. As such, the HUI-2 and HUI-3 may be more able to evaluate HRQL in more severe participants as the SF-6D doesn't appear to adequately capture health states at the lower end of the scale.

TABLE 3.1: Demographic Description of Participants

	Latent (n=78)	Active (n=84)	p-value
Age (mean in yrs, SD)	36.3 (11.1)	49.0 (19.0)	**< 0.01 ***
Gender (% male)	40%	45%	0.48
Foreign-born (%)	85%	89%	0.38
Race (%)			0.50
Asian or Pacific Islander	51%	60%	
East Indian (South Asian)	17%	14%	
Caucasian	13%	15%	
Black	5.1%	3.6%	
Aboriginal	3.8%	1.2%	
Others	10.1%	6.2%	
Marital Status (%)			0.55
Married	47%	56%	
Single	33%	27%	
Others	20%	17%	
Smoking (% yes)	26%	23%	0.60
Alcohol use (% yes)	69%	86%	**0.01 ***
Drug use (% yes)	10%	4.8%	0.27
Co-morbidity (% yes)	24%	45%	**0.02 ***

TABLE 3.2: SF-36 Scores

	Latent (n=78)			Active (n=84)		
	Mean (95% CI)	**Median**	**IQR**	**Mean (95% CI)**	**Median**	**IQR**
PF *	54.3 (52.7-55.9)	57.0	0.0	41.2 (38.1-44.4)	40.2	25.3
RP	51.9 (49.9-53.8)	56.9	9.2	38.1 (34.7-41.5)	37.3	30.0
BP	55.7 (53.4-58.0)	62.1	11.0	50.1 (47.3-53.0)	51.1	20.7
GH	52.5 (50.7-54.4)	52.9	9.5	44.5 (42.3-46.7)	44.1	14.3
EN	52.7 (50.7-54.8)	52.1	12.5	44.0 (41.0-46.9)	42.7	18.7
SF	52.7 (50.6-54.7)	56.8	5.5	39.5 (36.3-42.7)	40.5	23.2
RE	50.7 (48.6-52.9)	55.9	2.9	35.0 (30.9-39.1)	32.6	35.0
EM	50.9 (48.7-53.0)	52.8	14.1	44.4 (41.8-47.0)	43.0	16.9
PCS	54.7 (53.2-56.1)	56.8	6.3	44.8 (42.1-47.5)	45.9	18.9
MCS	50.3 (48.5-52.0)	52.8	8.9	40.1 (37.1-43.1)	41.4	22.0

* PF: Physical Functioning; RP: Role Limitations due to Physical Health; BP: Bodily Pain; GH: General Health; VT: Vitality; SF: Social Functioning; RP: Role Limitations due to Emotional Problems; MH: Mental Health; PCS: Physical Component Summary; MCS: Mental Component Summary

TABLE 3.3: Global Utility Scores

		Mean (95% CI)	Median	Min	Max	IQR
	SF-6D	0.82 (0.80-0.85)	0.85	0.40	1.00	0.12
Latent	**HUI-3**	0.90 (0.86-0.94)	0.97	0.08	1.00	0.10
(n=78)	**HUI-2**	0.93 (0.90-0.95)	0.95	0.49	1.00	0.08
	VAS	0.87 (0.84-0.90)	0.90	0.39	1.00	0.20
	SF-6D	0.68 (0.65-0.72)	0.64	0.32	1.00	0.24
Active	**HUI-3**	0.76 (0.70-0.82)	0.90	-0.26	1.00	0.31
(n=84)	**HUI-2**	0.85 (0.80-0.89)	0.93	0.13	1.00	0.19
	VAS	0.66 (0.61-0.71)	0.70	0.08	1.00	0.32

TABLE 3.4: Frequency of Reporting the Best and Worst Levels in Active Participants (n=84)

Utility Instruments	Best level, No. (%)	Worst level, No. (%)
HUI-2		
Sensation	51 (60.7%)	3 (3.6%)
Mobility	60 (71.4%)	0
Emotion	59 (70.2%)	0
Cognition	61 (72.6%)	1 (1.2%)
Self-care	77 (91.7%)	2 (2.4%)
Pain	42 (50.0%)	1 (1.2%)
HUI-3		
Vision	53 (63.1%)	2 (2.4%)
Hearing	82 (97.6%)	1 (1.2%)
Speech	77 (91.7%)	0
Ambulation	60 (71.4%)	0
Dexterity	70 (94.0%)	0
Emotion	35 (41.7%)	0
Cognition	19 (22.6%)	1 (1.2%)
Pain	42 (50.0%)	2 (2.4%)
SF-6D		
Physical function	22 (26.2%)	5 (6.0%)
Role limitation	28 (33.3%)	53 (63.1%)
Social function	25 (29.8%)	12 (14.3%)
Pain	35 (41.7%)	6 (7.1%)
Mental health	19 (22.6%)	4 (4.8%)
Vitality	6 (7.1%)	24 (28.6%)

TABLE 3.5: SF-36 and Utility Scores by Self-reported TB Severity and Control of TB Symptoms (Mean and SD)

	PCS	MCS	SF-6D	HUI-3	HUI-2	VAS
Severity of TB symptoms						
Very mild	54.5 (6.6)	52.4 (7.6)	0.84 (0.11)	0.93 (0.10)	0.94 (0.07)	0.84 (0.14)
Mild	50.0 (5.2)	44.6 (10.1)	0.68 (0.05)	0.84 (0.19)	0.93 (0.08)	0.75 (0.09)
Moderate	44.6 (9.5)	34.5 (12.6)	0.64 (0.10)	0.73 (0.28)	0.83 (0.19)	0.65 (0.13)
Severe	35.6 (12.0)	33.8 (12.4)	0.59 (0.15)	0.65 (0.36)	0.76 (0.25)	0.54 (0.22)
Very severe	32.5 (11.8)	28.9 (10.8)	0.54 (0.08)	0.53 (0.37)	0.71 (0.28)	0.35 (0.21)
Control of TB symptoms						
Very well controlled	55.8 (6.0)	51.8 (7.8)	0.85 (0.09)	0.90 (0.16)	0.94 (0.07)	0.84 (0.15)
Well controlled	49.3 (6.3)	47.7 (9.6)	0.73 (0.13)	0.92 (0.08)	0.95 (0.07)	0.74 (0.18)
Adequately controlled	40.0 (12.7)	32.3 (11.7)	0.60 (0.11)	0.68 (0.30)	0.78 (0.21)	0.60 (0.18)
Not well controlled	33.9 (8.7)	30.2 (11.9)	0.56 (0.11)	0.52 (0.38)	0.74 (0.27)	0.42 (0.22)
Not controlled at all	-	-	-	-	-	-

TABLE 3.6: Correlation (Spearman's rho) Between SF-36 and Utility Instruments

	PCS	MCS	SF-6D	HUI-3	HUI-2	VAS
PCS	1.00					
MCS	0.45	1.00				
SF-6D	0.79	0.80	1.00			
HUI-3	0.66	0.48	0.71	1.00		
HUI-2	0.59	0.37	0.59	0.90	1.00	
VAS	0.67	0.59	0.65	0.58	0.52	1.00

All p-values < 0.0001

TABLE 3.7: Intra-class Correlations (ICCs) Between the Four Utility Instruments

	SF-6D	HUI-3	HUI-2	VAS
SF-6D	1.00			
HUI-3	0.63 (0.53-0.72)	1.00		
HUI-2	0.67 (0.58-0.75)	0.84 (0.79-0.88)	1.00	
VAS	0.66 (0.56-0.74)	0.56 (0.45-0.66)	0.53 (0.41-0.63)	1.00

FIGURE 3.1: Distribution of SF-6D Scores in Active Participants

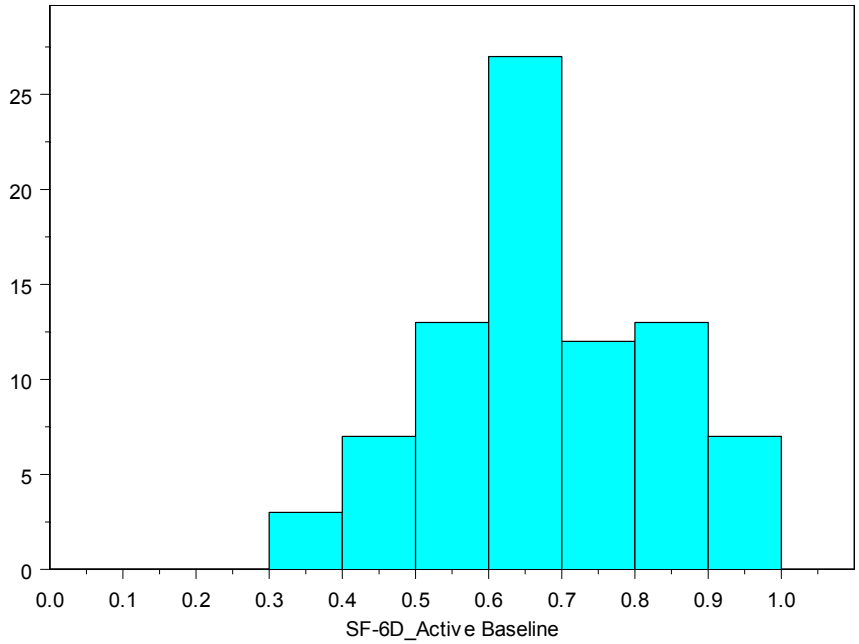


FIGURE 3.2: Distribution of HUI-3 Scores in Active Participants

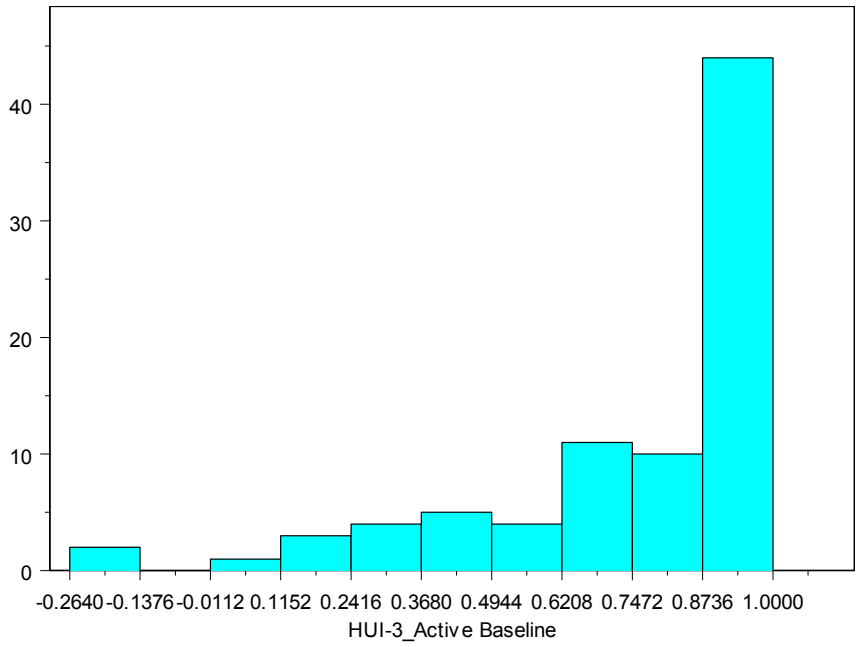


FIGURE 3.3: Distribution of HUI-2 Scores in Active Participants

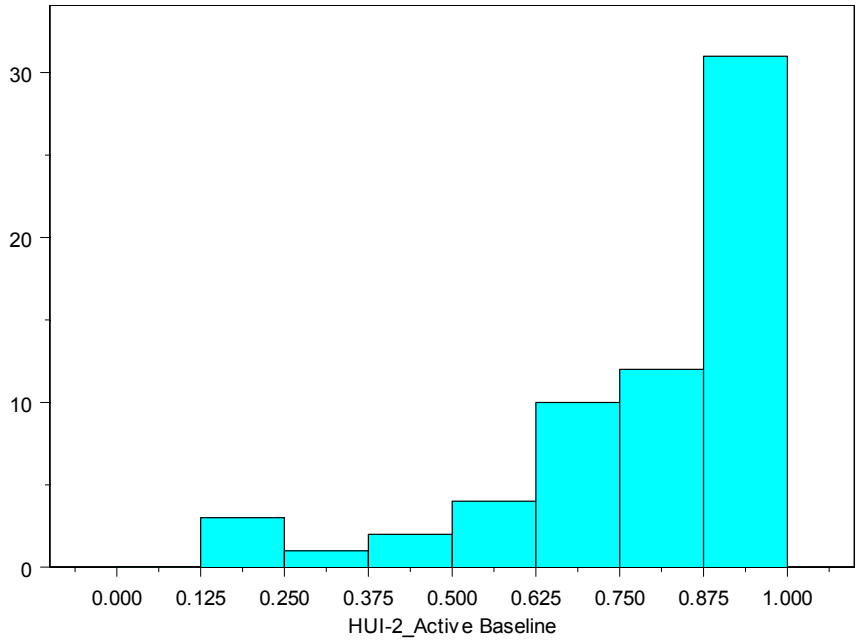
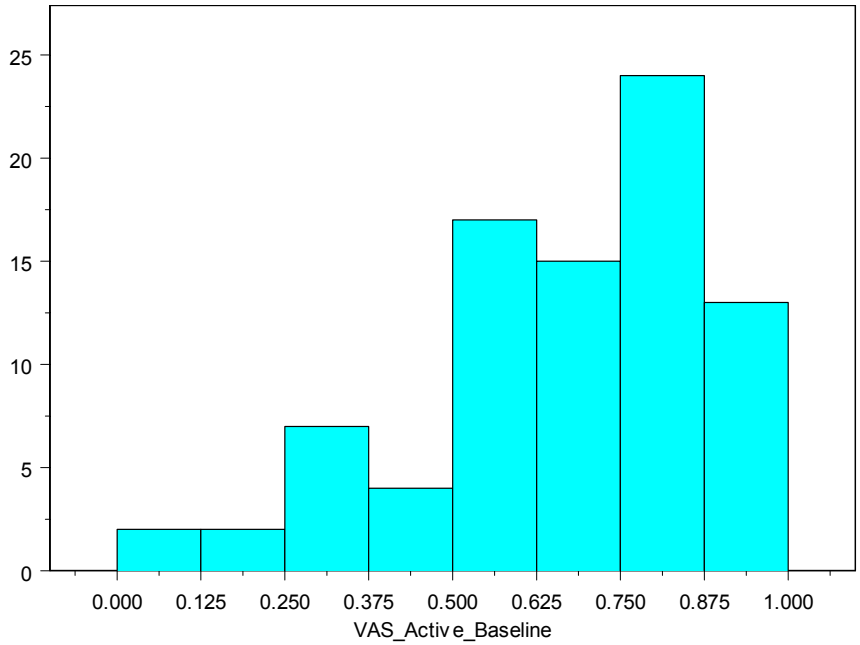


FIGURE 3.4: Distribution of VAS Scores in Active Participants



3.6 REFERENCES

1. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. JAMA 1999;282:677-686.

2. Blumberg HM, Burman WJ, Chaisson RE, Daley CL, Etkind SC, Friedman LN, Fujiwara P, Grzemska M, Hopewell PC, Iseman MD, Jasmer RM, Koppaka V, Menzies RI, O'Brien RJ, Reves RR, Reichman LB, Simone PM, Starke JR, Vernon AA; American Thoracic Society, Centers for Disease Control and Prevention and the Infectious Diseases Society. American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America: treatment of tuberculosis. Am J Respir Crit Care Med 2003;167:603-662.

3. Long R, Njoo H, Hershfield E. Tuberculosis: 3. Epidemiology of the disease in Canada. CMAJ 1999;160:1185-1190.

4. Chamla D. The assessment of patients' health-related quality of life during tuberculosis treatment in Wuhan, China. Int J Tuberc Lung Dis 2004;8:1100-1106.

5. Hansel NN, Wu AW, Chang B, Diette GB. Quality of life in Tuberculosis: patient and provider perspectives. Qual Life Res 2004;13:639-652.

6. Dion MJ, Tousignant P, Bourbeau J, Menzies D, Schwartzman K. Feasibility and reliability of health-related quality of life measurements among tuberculosis patients. Qual Life Res 2004;13:653-665.

7. Dion MJ, Tousignant P, Bourbeau J, Menzies D, Schwartzman K. Measurement of health preferences among patients with tuberculous infection and disease. Medical Decision Making 2002;22(Suppl.):S102-114.

8. Duyan V, Kurt B, Aktas Z, Duyan GC, Kulkul DO. Relationship between quality of life and characteristics of patients hospitalised with tuberculosis. Int J Tuberc Lung Dis 2005;9:1361-1366.

9. Rajeswari R., Muniyandi M, Balasubramanian R, Narayanan PR. Perceptions of tuberculosis patients about their physical, mental and social well-being: A field report from south India. Soc Sc Med 2005;60:1845-1853.

10. Drummond MF, Sculpher MJ, Torrance GW, eds. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford University Press, 2005.

11. Gourevitch MN, Alcabes P, Wasserman WC, Arno PS. Cost-effectiveness of directly observed chemoprophylaxis of tuberculosis among drug users at high risk for tuberculosis. Int J Tuberc Lung Dis 1998;2:531-540.

12. Perlman DC, Gourevitch MN, Trinh C, Salomon N, Horn L, Des Jarlais DC. Cost-effectiveness of tuberculosis screening and observed preventive therapy for active drug injectors at a syringe-exchange program. J Urban Health 2001;78:550-567.

13. Marchand R, Tousignant P, Chang H. Cost-effectiveness of screening compared to case-finding approaches to tuberculosis in long-term care facilities for the elderly. Int J Epidemiol 1999;28:563-570.

14. FitzGerald JM, Gafni A. A cost-effectiveness analysis of the routine use of isoniazid prophylaxis in patients with a positive Mantoux skin test. Am Rev Respir Dis 1990;142:848-853.

15. Salpeter SR, Sanders GD, Salpeter EE, Owens DK. Monitored isoniazid prophylaxis for low-risk tuberculin reactors older than 35 years of age: a risk-benefit and cost-effectiveness analysis. Ann Intern Med 1997;127:1051-1061.

16. Dasgupta K, Schwartzman K, Marchand R, Tennenbaum TN, Brassard P, Menzies D. Comparison of cost-effectiveness of tuberculosis screening of close contacts and foreign-born populations. Am J Respir Crit Care Med 2000;162:2079-2086.

17. Floyd K. Costs and effectiveness--the impact of economic studies on TB control. Tuberculosis (Edinb) 2003;83:187-200.

18. Khan K, Muennig P, Behta M, Zivin JG. Global drug-resistance patterns and the management of latent tuberculosis infection in immigrants to the United States. N Engl J Med 2002;347:1850-1859.

19. Schwartzman K, Menzies D. Tuberculosis screening of immigrants to low-prevalence countries. A cost-effectiveness analysis. Am J Respir Crit Care Med 2000;161:780-789.

20. Diel R, Wrighton-Smith P, Zellweger JP. Cost-effectiveness of interferon-gamma release assay testing for the treatment of latent tuberculosis. Eur Respir J 2007;30:321-332.

21. Diel R, Nienhaus A, Loddenkemper R. Cost-effectiveness of interferon-gamma release assay screening for latent tuberculosis infection treatment in Germany. Chest 2007;131:1424-1434.

22. Oxlade O, Schwartzman K, Menzies D. Interferon-gamma release assays and TB screening in high-income countries: a cost-effectiveness analysis. Int J Tuberc Lung Dis 2007;11:16-26.

23. Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.

24. Brazier J, Roberts J, Deveril M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271-292.

25. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI ®): concepts, measurement properties and applications. Health Qual Life Outcomes 2003;1:54.

26. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. Med Care 1996;34:702-722.

27. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, Denton M, Boyle M. Multiattribute and single-attribute utility functions for the Health Utilities Index Mark 3 system. Med Care 2002;40:113-128.

28. Cohen J. A power primer. Psychol Bull 1992;112:155-159.

29. Colton T. Statistics in medicine. Boston: Little, Brown, 1974.

30. Schechter CB, Rose DN, Fahs MC, Silver AL. Tuberculin screening: cost-effectiveness analysis of various testing schedules. Am J Prev Med 1990;6:167-175.

31. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chronic Dis 1978;31:697-704.

32. Puhan MA, Guyatt GH, Montori VM, Bhandari M, Devereaux PJ, Griffith L, Goldstein R, Schünemann HJ. The standard gamble demonstrated lower reliability than the feeling thermometer. J Clin Epidemiol 2005;58:458-465.

33. Witney AG, Treharne GJ, Tavakoli M, Lyons AC, Vincent K, Scott DL, Kitas GD. The relationship of medical, demographic and psychosocial factors to direct and indirect health utility instruments in rheumatoid arthritis. Rheumatology (Oxford) 2006;45:975-981.

34. Arocho R, McMillan CA, Sutton-Wallace P. Construct validation of the USA-Spanish version of the SF-36 health survey in a Cuban-American population with benign prostatic hyperplasia. Qual Life Res 1998;7:121-126.

35. Thumboo J, Fong KY, Ng TP, Leong KH, Peng PH, Thio ST, Boey ML. Validation of the MOS SF-36 for quality of life assessment of patients with systemic lupus erythematosus in Singapore. J Rheumatol 1999;26:97-102.

36. Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, Esdaile JM, Anis AH. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D,

and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sc Med 2005;60:1571-1582.

37. O'Brien BJ, Spath M, Blackhouse G, Severens JL, Dorian P, Brazier J. A view from the bridge: agreement between the SF-6D utility algorithm and the Health Utilities Index. Health Econ 2003;12:975-981.

CHAPTER 4 IMPACT OF ADVER DRUG REACITONS AND PREDICTIVITY OF QUALITY OF LIFE STATUS

4.1 FOREWORD

The content of this chapter has been accepted for publication in *European Respiratory Journal* under the same title. [3] As the first author, the candidate developed the research hypotheses, performed data analyses, and wrote the final manuscript. Co-authors of this study included *Drs*. *Fawziah Marra*, *J. Mark FitzGerald*, *R. Kevin Elwood*, and *Carlo A. Marra*. APPENDIX 4 presents supplementary contents to this manuscript.

4.2 INTRODUCITON

In Canada, people diagnosed with active tuberculosis (TB) disease are routinely treated with isoniazid, rifampin, pyrazinamide and ethambutol. Although effective, the treatment is associated with significant adverse drug reactions (ADRs). Current knowledge of these ADRs has focused on their frequency and clinical natures [1,2]. However, reporting clinical natures of ADRs is not adequate to fully reflect their impacts on patients' health status. In recent years, patient-reported outcomes, such as health-related quality of life (HRQL), are increasingly appreciated in evaluating the impact of illnesses and the effectiveness of medical interventions [3]. Therefore, we intended to investigate the

[3] A version of this chapter has been accepted for publication in the *European Respiratory Journal*. Guo N, Marra F, Fitzgerald JM, Elwood RK, Marra CA. Impact of adverse drug reaction and predictivity of quality of life status in tuberculosis. *Eur Respir J* 2010;36:206-208.

impact of anti-TB treatment-induced ADRs on patients' HRQL and to examine the association between baseline HRQL status and the likelihood of developing ADRs during the subsequent treatment.

4.3 METHODS

In British Columbia, TB patients are seen monthly through TB control clinics managed by the British Columbia Center for Disease Control (BCCDC). Monthly laboratory tests are done to monitor patients' liver and renal function and haematological status. Tolerance to medications is evaluated during clinic visits. Medical records are kept in the *integrated* Public Health Information System (*i*PHIS).

This study partially represents a longitudinal HRQL study, where English-speaking adults with newly initiated treatment for active TB disease or latent TB infection were recruited from BCCDC during 2005 and 2006. Ethics approval was obtained from the University of British Columbia's Behavioural Research Ethics Board. Participants provided informed consent. For the present study, active TB patients were considered eligible if they further met the criteria: (1) had no pre-existing liver problems or no other known severe health conditions before the treatment; (2) completed at least 3 months of treatment; and (3) completed both baseline and 3-month HRQL assessments. HRQL was measured at baseline, 3 months and 6 months of the treatment using the Short-Form 36 version 2 (SF-36 v2) [4], as there is no well-validated TB-specific questionnaire in the literature. SF-36 measures eight dimensions: physical functioning (PF), role-physical

(RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role-emotional (RE) and mental health (MH). From these subscales, a physical component summary (PCS) and a mental component summary (MCS) can be determined [4]. As the majority of ADRs occurred within the first 3 months of anti-TB treatment [1, 2], only baseline and 3-month SF-36 outcomes were used for analyses.

From the *i*PHIS database, physician narratives and nurse notes were retrospectively reviewed to identify anti-TB medication related ADRs. The criteria we have used previously to assess the possibility and severity of ADRs were applied here [2]. All *definite*, *possible* and *probable* ADRs during the first 3 months of the treatment were included and classified by severity: an ADR was defined as *major* if it led to discontinuation of routine treatment and/or required additional treatment for the symptoms; an ADR was *minor* if no additional medical interventions were taken. Patients were then categorized into three groups: (1) no ADR: never had any ADR; (2) minor ADR: only developed minor ADRs; and (3) major ADR: developed at least one major ADR.

ANOVA, Pearson's Chi-squared test or Fisher's exact test was used to compare the baseline socio-demographic differences between the three groups. SF-36 scores were calculated and transformed using norm-based scoring that resulted in a mean of 50 and standard deviation of 10 in the general population [4]. A higher SF-36 score indicates a better HRQL outcome. Multiple linear regression models were constructed to explore the impact of ADRs on the 3-month HRQL. The dependent variable was each one of the SF-

36 scales (eight subscales, PCS and MCS). In each model, two independent variables were included to represent the presence/absence of major ADRs and minor ADRs respectively. Baseline SF-36 score, age, sex, race, marital status, smoking and the presence of comorbidities were adjusted for in each model. Ordinal logistic regression was applied to examine the association between baseline SF-36 scores and the occurrence of ADRs during the first 3 months. The dependent variable indicating the occurrence of ADRs has three levels: no ADR, minor ADR and major ADR. Then we added each of the baseline SF-36 scores, as a predictor variable, to separated models. Age, sex, race, marital status, smoking and the presence of comorbidities were included in each model. Considering their clinical importance, all covariates were kept in final fitted regression models, regardless of their statistical significance. All analyses were performed using software SAS (version 9.0; SAS Institute Inc, Cary, NC).

4.4 RESULTS

Among the 104 active TB patients recruited in the HRQL study, 89 further met the present study's inclusion criteria and were included. Their average age was 49.2 yrs, with 56% being female. Seventy-one per cent were ethnically Asian. Of the 89 subjects, 41 (46%) reported having other health conditions, such as high blood pressure, diabetes, back problems, arthritis and asthma. During the first 3 months of treatment, 21 patients (24%) developed at least one major ADR and another 29 (33%) experienced only minor ADRs. Common ADRs included skin rash and/or pruritus (35%), gastrointestinal symptoms (30%), liver damage (including mild liver enzyme elevation and hepatitis)

(23%), paresthesia (18%), fatigue/weakness (18%), visual disturbance (17%) and joint/muscle pain (17%). The most common ADR leading to treatment discontinuation was liver enzyme elevation (11%). Sociodemographic differences were observed across the three groups (no ADR, minor ADR and major ADR), e.g. females were more likely to develop ADRs than males. However, none of these differences were statistically significant.

Baseline and 3-month SF-36 scores are summarised in TABLE 4.1. At 3 months, all SF-36 scores were monotonically decreasing across the three groups. After adjusting for socio-demographic factors and baseline SF-36 scores, compared to those who had no ADRs, subjects who developed major ADRs had significantly lower 3-month scores on PF ($p=0.03$), VT ($p=0.01$), MH ($p=0.01$) and MCS ($p=0.03$). No significant differences were found between subjects who experienced only minor ADRs and those who had no ADRs. As seen in TABLE 4.1, subjects who developed major ADRs during the 3 months of treatment scored lowest at baseline, whereas those who had no ADRs scored highest. Ordinal logistic regression analyses showed that subjects who scored lower on SF-36 at baseline were significantly associated with a higher risk of developing ADRs during the treatment, after adjusting for socio-demographic factors. For example, for a one-unit increase in baseline MCS score, the odds of developing ADRs (minor or major) would be reduced by 6% (OR=0.94, $p=0.001$), and the odds of having major ADRs versus having no ADRs and minor ADRs would be reduced by 6% as well (the proportional odds assumption).

4.5 DISCUSSIONS AND CONCLUSIONS

This study is the first to examine the anti-TB treatment induced ADRs in terms of HRQL experienced by patients. We found that anti-TB treatment ADRs had substantial impacts on patients' HRQL, and that poor baseline HRQL status was associated with a higher possibility of developing ADRs during the subsequent treatment.

Our results showed that developing major ADRs led to significant reduction on two mental health subscales (VT and MH), the mental health summary (MCS) and one physical subscale (PF) of SF-36. Experiencing minor ADRs showed some negative impacts, but none were statistically significant. This suggests that severe ADRs result in more decrements in HRQL; experiencing ADRs seems to be more of a mental wellbeing burden than a physical one. We also observed that patients with lower baseline SF-36 scores had a higher risk of developing ADRs during the first 3 months of treatment, after adjusting for important socio-demographic factors. HRQL measures capture patients' self-perceived health status, functioning and well-being. Poor baseline HRQL status may reflect known or unknown physical and mental health problems and individual psychological coping mechanisms, which would predispose the individual to a higher risk of unfavourable outcomes. Our results suggest that baseline HRQL measurements could potentially help identify those patients who would eventually develop an ADR to the anti-TB treatment. This finding could help plan quality healthcare management and justify resource allocation, e.g. improving the baseline HRQL of those high-risk patients through lifestyle modification consultation and social-psychological support would

promote the outcome to medical treatment. Future prospective studies are needed to better understand the predictive value of HRQL status and to determine whether it could be manipulated for therapeutic purpose [5–7].

There are some limitations to our study. First, ADRs were retrospectively collected by reviewing administrative medical records, which usually are not established for the purpose of research and may not be comprehensive. Second, the TB population may be exposed to various risk factors. Although we controlled for some important socio-demographic factors in our analyses, there may still be unmeasured confounders that we did not capture. Finally, our inadequate sample size might prevent us from exploring more significant findings. Despite these limitations, as the first study on this topic, we believe the current report is a good starting point for future research.

TABLE 4.1: SF-36 Scores by Severity of ADRs

SF-36 v2 Scales	N	Baseline	3 month	p-value [§]
Physical functioning (PF)				
No ADR	38	45.2 ± 14.4	46.9 ± 13.5	-
Minor ADR	29	42.6 ± 12.3	44.7 ± 11.9	0.75
Major ADR	20	33.6 ± 13.2	33.8 ± 11.9*	**0.03**
Role-physical (RP)				
No ADR	36	45.0 ± 13.4	47.1 ± 13.0	-
Minor ADR	29	36.2 ± 15.1	41.0 ± 14.7	0.70
Major ADR	18	29.4 ± 13.6	36.7 ± 11.1	0.95
Bodily pain (BP)				
No ADR	39	53.9 ± 11.4	55.9 ± 9.6	-
Minor ADR	28	52.8 ± 9.2	53.2 ± 8.8	0.87
Major ADR	20	44.1 ± 15.6	45.5 ± 13.7	0.26
General health (GH)				
No ADR	39	46.2 ± 9.5	47.7 ± 9.0	-
Minor ADR	28	44.2 ± 8.5	47.1 ± 8.7	0.44
Major ADR	20	41.2 ± 8.0	41.1 ± 8.2	0.07
Vitality (VT)				
No ADR	39	46.9 ± 12.0	50.1 ± 9.5	-
Minor ADR	28	42.4 ± 12.0	46.2 ± 8.6	0.41
Major ADR	20	38.5 ± 10.6	39.1 ± 10.8*	**0.01**
Social functioning (SF)				
No ADR	39	44.8 ± 11.6	46.5 ± 11.1	-
Minor ADR	28	35.6 ± 14.8	43.4 ± 11.4	0.60
Major ADR	20	32.6 ± 13.6	36.4 ± 13.1	0.12
Role-emotion (RE)				
No ADR	37	40.6 ± 16.4	43.9 ± 16.4	-
Minor ADR	29	32.1 ± 18.9	38.8 ± 15.8	0.66
Major ADR	18	22.4 ± 15.8	31.7 ± 13.7	0.68
Mental health (MH)				
No ADR	38	47.8 ± 9.0	50.3 ± 8.5	-
Minor ADR	26	41.0 ± 12.7	46.2 ± 7.6	0.83
Major ADR	20	39.2 ± 9.5	39.2 ± 13.2*	**0.01**
Physical summary (PCS)				
No ADR	35	49.2 ± 11.6	50.6 ± 10.5	-
Minor ADR	26	47.0 ± 9.1	48.6 ± 9.7	0.32
Major ADR	18	38.4 ± 10.9	40.3 ± 10.2	0.52
Mental summary (MCS)				
No ADR	35	44.7 ± 9.5	47.7 ± 9.1	-
Minor ADR	26	36.2 ± 14.8	43.0 ± 8.7	0.96
Major ADR	18	32.5 ± 11.6	36.3 ± 12.5*	**0.03**

Data are presented as mean¡SD, unless otherwise stated. [§] p-values are based the multiple regression analyses, adjusting for baseline SF-36 score, age, sex, race, marital status, smoking, comorbidity. Subjects who had no ADRs are the reference. * p-value<0.05.

4.6 REFERENCES

1. Forget EJ, Menzies D. Adverse reactions to first-line antituberculosis drugs. Expert Opin Drug Saf 2006;5:231–249.

2. Marra F, Marra CA, Bruchet N, Richardson K, Moadebi S, Elwood RK, Fitzgerald JM. Adverse drug reactions associated with first-line anti-tuberculosis drug regimens. Int J Tuberc Lung Dis 2007;11:868–875.

3. Guyatt GH, Feeny DH, Patrich DL. Measuring health-related quality of life. Ann Intern Med 1993;118:622–629.

4. Ware JE Jr., Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. User's manual for the SF-36v2™ health survey (2nd ed). Lincoln, RI: QualityMetric Incorporated, 2007.

5. Gold M, Franks P, Erickson P. Assessing the health of the nation. The predictive validity of a preference-based measure and self-rated health. Med Care 1996;34:163–177.

6. Franks P, Gold MR, Fiscella K. Sociodemographics, self-rated health, and mortality in the US. Soc Sci Med 2003;56:2505–2514.

7. Sehlen S, Lenk M, Hollenhorst H, Schymura B, Aydemir U, Herschbach P, Dühmke E. Quality of life (QoL) as predictive mediator variable for survival in patients with intracerebral neoplasma during radiotherapy. Onkologie 2003;26:38–43.

CHAPTER 5 PATIENTS' PREFERENCE FOR LATENT TUBERCULOSIS INFECTION PREVENTIVE TREATMENT: A DISCRETE CHOICE EXPERIMENT

5.1 FOREWORD

The content of this chapter has been submitted for publication under the same title. [4] As the first author, the candidate's role included formulating research questions, questionnaire design, study subject recruitment, data collection, performance of statistical analysis, and the writing of the final manuscript. Co-authors of this study included *Drs. Carlo A. Marra, J. Mark FitzGerald*, *R. Kevin Elwood*, *Aslam H. Anis*, and *Fawziah Marra*.

5.2 INTRODUCTION

Tuberculosis (TB) continues to be a major public health threat with one third of world's population infected [1,2]. Since new active TB cases mostly arise from people with latent TB infection (LTBI), preventing LTBI from developing into clinical disease has both individual benefit and public health significance [3,4]. Isoniazid (INH) has been long used for treating LTBI with well-established efficacy from multiple clinical trials [5-7].

[4] A version of this chapter has been submitted for publication. Guo N, Marra CA, FitzGerald JM, Elwood RK, Anis AH, Marra F. *Patients' preference for latent tuberculosis infection preventive treatment: a discrete choice experiment*.

However, the effectiveness of INH treatment has been limited by the low acceptance rate and poor adherence among patients [8-11]. In 2004, in British Columbia (BC), among people who were offered LTBI preventive treatment, 49.2% accepted and initiated the treatment; and then 50% of those who accepted the treatment eventually completed it [4]. The challenge for health care professionals is to advocate a preventive treatment to people who do not have clinical disease and are otherwise healthy, not to mention that the treatment has potential significant side effects and long-term commitment.

A large body of research has intended to identify patient socio-demographic profiles and other modifiable predictors associated with poor adherence among patients on LTBI preventive treatment [8-11]. However, patients' treatment decision-making behavior has rarely been closely investigated [12]. Treatment decision-making is a complex behavior, not only depending on the probability of risks and benefits, but also on how patients perceive and value these uncertainties. Therefore, the primary objective of this study was to quantify patients' preferences when making decisions on whether to accept the preventive treatment, using a discrete choice experiment (DCE) technique. Originated from marketing, transport economics and environmental economics, DCE has been recently adopted in the health care sector to elicit patient or public preferences on health products and services [13,14].

5.3 METHODS

5.3.1 Development of Discrete Choice Experiment Survey

In BC, individuals diagnosed with LTBI are requested to see a physician to discuss preventive treatment with 9 months of INH through TB clinics, managed by BC Centre for Disease Control (BCCCD). The treatment is publicly funded and is optional to the patients. Once the patients initiate the treatment, they come to the clinic monthly to pick up their medications until the treatment is completed.

To identify treatment attributes for the DCE survey, semi-structured individual interviews (see APPENDIX 5) were conducted at the BCCDC TB clinic. Twenty newly diagnosed LTBI patients were purposively chosen to ensure that half of them had accepted the treatment and half had declined it. After discussing with the physician, the patients were invited to talk about their perceived motivating factors and barriers of initiating the treatment. Based on interview results and expert opinions, 6 key treatment attributes with various levels (TABLE 5.1) were chosen: (1) *length of treatment*; (2) *frequency of clinic visit*; (3) *risk of developing active TB disease after treatment*, which was used as an indicator of the treatment effectiveness; (4) *chance of developing liver damage*; (5) *chance of developing skin rash*; (6) *chance of developing fatigue*.

Sawtooth® CBC/SSI Web V6.4.2 (Sawtooth Software, Inc. Sequim, WA, USA) was used to generate 12 versions of DCE questionnaires. A fractional factorial experiment design was applied where orthogonality, level balance, and minimal overlap were taken into

account [15]. Each version had 10 random choice sets plus 2 identical fixed choice sets. An example choice set is given in FIGURE 5.1. In each choice set, respondents were presented with 2 hypothetical treatment options (A and B) and asked to choose the one they preferred. If they preferred not to take any treatment, an opt-out option of 'Neither' was also given and it was explicitly described as "*a lifetime risk of 10% of developing active TB disease without taking preventive treatment and no risk of any side effects*" [16].

Located at question 1 and 11 in the questionnaire, the fixed choices were to check the validity and consistency of the responses and were not included in data analysis [17,18]. In the fixed choices, treatment option B had clearly dominant attribute levels than A, i.e. shorter length, more effective but fewer side effects. Respondents were expected to choose option B over A if they understood the task and made rational decisions (or choose 'Neither' if they were determined not to take any preventive treatment at all).

In the random choice sets, treatment options were carefully checked to limit the number of dominant scenarios where no trade-offs would be required. Prior to final administration, the DCE survey was pilot-tested among 60 LTBI patients to ensure: (1) attributes and levels are plausible, meaningful and understandable to the respondents; and (2) the format of the survey and the presentation of questions are clear and easy to understand.

5.3.2 Subject Recruitment and Data Collection

Subjects were recruited through BCCDC TB clinics, if they met the following criteria: (1) diagnosed with LTBI as per clinical practice; (2) 19 years or older; and (3) able to understand and complete the English-language questionnaires. Ethics approval was obtained from the University of British Columbia Behavioural Research Ethics Board.

Data was collected using a questionnaire that included the DCE survey, socio-demographics and other information relevant to TB, such as reason of doing tuberculin skin test (TST) and Bacille Calmette Guerin (BCG) vaccination status. An explanation of DCE choice tasks was given to each subject and a 'warm-up' choice set was completed. The subject could complete the questionnaire at the clinic or complete it elsewhere and bring it back on the next clinic visit.

5.3.3 Statistical Analysis

Subjects who failed the two consistency-check questions or had incomplete socio-demographic profiles were excluded. Included respondents' baseline characteristics were summarized with descriptive statistics. DCE choice data were analyzed using conditional logistic regression (CLR) and latent class analysis (LCA). All attribute variables were evaluated on a continuous scale except for the '*frequency of clinic visit*', which was effects coded [19].

CLR was conducted to estimate the overall mean preferences for the treatment attributes among respondents. Traditional approaches, such as CLR, assume homogeneous preferences among the study population. However, we hypothesized that preferences would vary by respondents' socio-demographic status and other characteristics. Therefore, segmented CLR analyses were performed to estimate the preferences in various subgroups. The Wald test was used to compare preference differences across these subgroups.

Unlike CLR, LCA accounts for preference heterogeneity. It assumes that preferences can be reflected by observed characteristics (e.g. treatment attributes and socio-demographic variables) and unobserved factors (e.g. attitudes) [20-22]. As such, respondents could be grouped into a finite number of latent classes. Preferences are relatively homogenous within each class. LCA also accounts for the fact that each respondent makes multiple choices which are correlated. We successively fitted LCA models from 2 classes up to 10 classes and calculated 3 model fit statistical indicators for each model, including Log-likelihood (LL), Akaike Information Criteria (AIC), and Bayesian Information Criteria (BIC). The optimal number of classes was determined when a point was reached where an additional class would not significantly improve the model fit. In addition, the practical interpretability of class membership was considered. Socio-demographic and other relevant variables were also investigated for inclusion in the final LCA model based on their influences on the class membership, using a manual stepwise selection method.

LCA was performed using Latent GOLD® version 4.5 (Statistical Innovations Inc., Belmont, MA, USA) and other analyses were conducted using SAS® version 9.1 (SAS Institute Inc., Cary, NC, USA). Statistical significance was defined as p value ≤ 0.05.

5.4 RESULTS

5.4.1 Respondents Characteristics

As shown in FIGURE 5.2, among the 214 respondents, 10 (4.7%) answered both fixed questions wrong and were excluded; 200 (93.5%) answered at least one fixed question correct; 4 (1.9%) chose 'Neither' option in all 12 choice sets whatever the level of attributes (i.e. non-demanders). After further excluding 10 respondents with incomplete socio-demographic profiles, 194 (90.7%) were included in final data analysis.

TABLE 5.2 summarizes the baseline characteristics of the included respondents. Their average age was 38.0 years, 61.9% were females, and 53.6% had children. 85.6% of the respondents were born outside Canada and 68.6% had Asian background. Among the 194 respondents, 80.9% had college/university or higher education, 65.0% were currently employed and 55.0% had annual family income below $40,000 CAD. 22.2% of our respondents reported having other health conditions, e.g. hypertension, arthritis, and diabetes. Reasons for having a TST screening test: 19.6% were exposed to an active TB case; 16.5% did the test for medical reasons or a doctor's recommendation; and 63.9% for school, employment or immigration screening. Half of the respondents believed they were BCG vaccinated in the past.

5.4.2 Preference Estimates from Conditional Logistic Regression

After excluding the fixed questions, responses to 1,940 random DCE choice sets from 194 respondents were included for choice analysis. The overall mean preferences derived from the general CLR are presented in TABLE 5.3.

All 6 attributes were found to significantly influence respondents' treatment choices. The negative preference estimates reflected that respondents were averse to longer treatment regimen (-0.04 per month longer) and higher risk of active TB (-0.23 per 1% risk increase), liver damage (-0.16 per 1% risk increase), skin rash (-0.02 per 1% risk increase) and fatigue (-0.03 per 1% risk increase). On average, respondents had a significantly positive preference on *every 1 month* clinic visit (0.17, p=0.01); they showed negative preferences for *no visit* (-0.08), *every 2 weeks* (-0.05) and *every 2 months* (-0.03), but none were statistically significant. Subgroup CLR analyses suggested that respondents' treatment preferences varied substantially by most socio-demographic and other TB-relevant variables. Therefore, LCA was conducted to closely investigate the preference heterogeneity among respondents.

5.4.3 Preference Estimates from Latent Class Analysis

Although the model fit statistics didn't show clear convergence on an optimal number of latent classes, they were all leveling off at 3- and 4-class models. After further considering the interpretability of class membership, a 3-class solution is presented and

discussed here. Five covariates (i.e. origin of birth, having children, education level, employment status, and on OTC medications) were included in the final fitted 3-class model, since they significantly influenced the class assignment. The preference estimates of the 3 classes are summarized in TABLE 5.4. The Wald statistic tests whether the preference was different across the 3 classes. The relative importance of the 6 treatment attributes to respondents in each class is shown in FIGURE 5.3.

Consistent with the CLR results, respondents' preference decreased with increased risk of active TB and side effects and longer treatment length (TABLE 5.4). Among the 6 attributes, *chance of skin rash*, *chance of fatigue* and *length of treatment* were of least importance to most respondents (FIGURE 5.3). Significant differences on preferences across the 3 classes existed in 2 attributes, *frequency of clinic visit* ($p<0.01$) and *chance of developing liver damage* ($p<0.01$), as indicated by the sign and magnitude of preference estimates (TABLE 5.4). Class-1 respondents considered the *risk of developing active TB* and *frequency of clinic visit* most important when choosing treatment and they preferred *every 1 month* clinic visit over other frequency options. Individuals in class-2 were most averse to the *chance of liver damage*. They preferred clinic visit of *every 2 months* over other options. In class-3, respondents placed similarly important values on the *risk of developing active TB*, *chance of developing liver damage* and *frequency of clinic visit* in treatment choices. They preferred no clinic visit.

The class probabilities indicated that 47% of the respondents were assigned in class-1, 32% in class-2, and 21% in class-3. TABLE 5.5 describes the socio-demographic profiles

of the 3 classes. Class-3 respondents were identified as those who had the highest probability of choosing 'Neither' option in the DCE choice tasks. They were more likely to be born outside Canada, have higher education, be unemployed, and take OTC medications. Individuals in class-1 had the lowest probability of choosing 'Neither' option in the survey and they were most likely to have children among the 3 groups. The employment rate in class-2 was the highest among the three classes.

5.5 DISCUSSION AND CONCLUSION

To our knowledge, this is the first study to measure patients' preferences for LTBI preventive treatment using a DCE technique with a robust latent class analysis approach. Results consistently suggested that patients were overall in favor of treatment that was more effective, required shorter period, and carried fewer side effects, but substantial preference heterogeneity existed among respondents. Considering socio-demographic variables in choice data analysis provided valuable information to understand the heterogeneity and respondents' decision-making behavior. This paper illustrates how DCE could be a useful tool to understand patients' preferences toward health products and interventions and to possibly inform health care practice.

All studied treatment attributes were found to have significant impact on patients' decision-making. The *risk of developing active TB, chance of developing liver damage,* and *frequency of clinic visit* were the most important determinants of treatment choices for most respondents. Patients were much more concerned about serious side effects (i.e.

liver damage) than mild/moderate ones (i.e. skin rash and fatigue). Although previous studies have found that the length of regimen was a critical predictor to treatment adherence and completion [10-12], our results revealed that treatment benefits and serious side effects were more concerned than its length when patients made the initial decision whether to start the preventive treatment. From what we observed in practice, a fair proportion of patients agree to start the 9 month INH regimen but stop it after a while, even without experiencing side effects. This suggests initial decision-making and treatment adherence are different behaviors and they need to be examined and targeted differently.

Another important finding is the existence of substantial preference heterogeneity among respondents. A 3-latent-class model including 5 socio-demographic covariates appeared to best fit our data. Class-1 individuals were most likely to have children compared to others. They seemed to adhere to the preventive treatment and the current practice (i.e. monthly clinic visit) at the TB clinics. Class-2 respondents were most concerned about the potential risk of liver damage and preferred a slightly less frequent clinic visit, every 2 months. The employment rate in class-2 was the highest. One possible motivator for class-2 individuals to accept the preventive treatment might be the obligation or pressure from their work, mostly in health care or early education. Class-3 respondents had the highest probability of choosing 'Neither' option in DCE tasks and might be more likely to refuse the preventive treatment in reality. They were more likely to be born outside Canada, be unemployed, be highly educated, or take OTC medications. Most of our respondents were born in Asian countries, where BCG vaccination is a routine practice.

The well-known fact that the LTBI diagnosis test, TST, is highly confounded by BCG vaccination may play a role in the high rate of declining the preventive treatment among the foreign-born. For the unemployed individuals, current living needs or being busy with school work possibly leave them little time or energy to consider preventing a disease that might only happen in the future. It is interesting that higher education was found to be associated with a tendency of refusing the preventive treatment. It is possible that people with higher education are more likely to choose a healthy lifestyle and believe they have more control over their own health. They would have less concern about an infectious disease such as TB, which is often associated with lower socio-economic status.

Our results underline the importance of accounting for preference heterogeneity when analyzing choice data. Model fit statistics suggested that LCA model fitted our data better than CLR model and including socio-demographic covariates in LCA models would further improve the model fit. LCA provides richer information that would not be observed with more traditional CLR approaches. Considering respondents' socio-demographic profiles would provide us informative insights in understanding their perceptions toward risks and benefits and how they react differently to treatment options. This information could possibly assist health care professionals in more effectively communicating with patients when advocating the preventive treatment, leading to improved treatment acceptance.

Although widely used in marketing, environmental economics and transportation economics, DCE is still relatively new in health economics. Methodological concerns on

its experimental design, sample size and selection, and choice data collection and analysis have been widely discussed. A large sample size is generally suggested in DCE studies in order to achieve a desirable level of precision. Although there are no definitive answers to the optimal sample size, a few general considerations are recommended, e.g. number of attributes and levels, complexity of choice tasks, statistical approach used to analyze choice data, and degree of desired precision, and research budget constraints [14, 23]. In addition, an often used rule in regression analysis suggests that a minimum of 10 dependent variable observations is required for each independent variable. Our sample size, 194, was enough to fit a latent class model with 6 attribute variables and 5 covariates.

We ensured the validity of our DCE design through carefully choosing treatment attributes with patient interviews and expert opinions and piloting the survey before final administration. To help respondents understand, a 'warm-up' DCE choice task was practiced with the researcher before they completed the survey on their own. We also incorporated two fixed questions in the survey to check the validity and quality of the response data. 4.7% of the respondents answered both consistency-check questions wrong, which suggests some respondents might not fully understand the DCE tasks. Making trade-offs across 6 attributes could be cognitively challenging for some people. Another 10.3% of the respondents answered one fixed question wrong. Potential reasons could be that they were learning about their preferences or the trade-offs as they went through the survey; or some just simply lost their patience toward the end of the survey. However, considering the fact that most respondents spoke English as their second

language, our DCE survey was well accepted and understood. In addition, the findings were clinically plausible and generally in line with our expectations, which greatly supported the theoretical validity of our methodology.

The most expressed concern surrounding DCE and other stated preference (SP) techniques has been that 'stated' preference may not reflect 'true' preference and therefore not be able to predict individuals' actual behavior in reality [24-26]. Stated preferences are elicited under well-controlled experimental context. The experiment may simplify real-life decision-making situations and respondents may be not willing to or not able to express their 'true' preferences. One approach to address this concern would be to compare the stated choices in DCE survey with decision-making behavior in real-life.

In conclusion, the DCE technique offers a promising tool to understand and quantify patients' preferences and decision-making processes surrounding health products and services. To extend the usefulness of DCE in informing health care practice, future efforts should be directed to validation of this technique.

TABLE 5.1: Attributes and Levels to Describe LTBI Preventive Treatment

Treatment attribute	Levels	Neither option
Length of treatment	4 months, 6 months, 9 months, 12 months	None
Frequency of clinic visit	every 2 months, every 1 month, every 2 weeks	None
Risk of developing active TB after treatment	0%, 1%, 2%, 4%	10%
Chance of developing liver damage	0%, 1%, 3%, 5%, 10%	0%
Chance of developing skin rash	0%, 5%, 10%	0%
Chance of developing fatigue	0%, 5%, 10%	0%

TABLE 5.2: Description of Respondent Baseline Characteristics

Baseline characteristics	N=194
Age (mean, SD*)	38.0 years (11.8)
Female (n, %)	120 (61.9%)
Born outside Canada (n, %)	166 (85.6%)
Ethnic origin (n, %)	
Asian	133 (68.6%)
Caucasian	50 (25.8%)
Others	11 (5.7%)
Marital status (n, %)	
Single	74 (38.1%)
Married	102 (52.6%)
Others	18 (9.3%)
Have children (n, %)	104 (53.6%)
Education (n, %)	
High school or less	37 (19.1%)
College/University	123 (63.4%)
Post-graduate	34 (17.5%)
Employed (n, %)	126 (65.0%)
Annual household income (n, %)	
0~19,999	48 (26.4%)
20,000~39,999	52 (28.6%)
40,000~59,999	33 (18.1%)
60,000~99,999	23 (12.6%)
>=100,000	26 (14.3%)
Co-morbidity (n, %)	43 (22.2%)
On prescription medications (n, %)	69 (35.8%)
On OTC[§] medications (n, %)	32 (16.5%)
BCG vaccination status (n, %)	
Yes	100 (51.6%)
No	37 (19.1%)
Not sure	57 (29.4%)
Reason for TST (n, %)	
Contact of active TB case	38 (19.6%)
Medical reason/Dr. referrals	32 (16.5%)
School/employment/immigration	124 (63.9%)
Know someone who had active TB (n, %)	51 (26.3%)

SD*: standard deviation; [§]OTC: over the counter

TABLE 5.3: Overall Preference Estimate from Conditional Logistic Regression

Treatment attribute	Preference	SE [‡]	p-value
Length of treatment [†]	-0.04	0.01	<.01*
Frequency of clinic visit			
No visit	-0.08	0.15	0.59
Every 2 months	-0.03	0.07	0.62
Every 1 month	0.17	0.07	0.01*
Every 2 weeks	-0.05	0.07	0.40
Risk of developing active TB after treatment [§]	-0.23	0.03	<.01*
Chance of developing liver damage [§]	-0.16	0.01	<.01*
Chance of developing skin rash [§]	-0.02	0.01	0.03*
Chance of developing fatigue [§]	-0.03	0.01	<.01*

[†] every 1 month increase**;** [§] every 1% increase in risk; [‡] SE: standard error; *statistical significance

TABLE 5.4: Preference Estimate from Latent Class Analysis

Treatment attribute	Class-1 mean (SE)	Class-2 mean (SE)	Class-3 mean (SE)	Wald p-value
Length of treatment	-0.08 (0.02)*	-0.01 (0.03)	-0.09 (0.04)*	0.17
Frequency of clinic visit				<0.01§
No visit	-2.29 (0.45)*	-1.73 (0.44)*	1.80 (0.54)*	
Every 2 months	0.67 (0.18)*	0.69 (0.19)*	-0.60 (0.24)*	
Every 1 month	0.99 (0.17)*	0.65 (0.20)*	-0.55 (0.24)*	
Every 2 weeks	0.63 (0.16)*	0.39 (0.16)*	-0.65 (0.23)*	
Risk of developing active TB	-0.33 (0.04)*	-0.21 (0.06)*	-0.31 (0.08)*	0.34
Chance of developing liver damage	-0.10 (0.03)*	-0.42 (0.05)*	-0.23 (0.04)*	<0.01§
Chance of developing skin rash	-0.02 (0.02)	-0.05 (0.02)*	-0.05 (0.03)	0.53
Chance of developing fatigue	0.05 (0.02)*	-0.03 (0.02)	-0.03 (0.03)	0.69

* significant preference within the class; § significant difference across the 3 classes

TABLE 5.5: Socio-demographic Profile of Each Latent Class

Covariate	Class-1	Class-2	Class-3	All respondents (N=194)
Born outside Canada	80.0 %	85.9 %	97.7 %	85.6 %
Education				
High-school or less	21.5 %	24.7 %	5.0 %	19.1 %
College/university	63.9 %	62.3 %	63.2 %	63.4 %
Post-graduate	14.6 %	12.5 %	31.8 %	17.5 %
Employed	66.1 %	78.3 %	41.8 %	65.0 %
Having children	62.6 %	43.5 %	48.6 %	53.6 %
On OTC medications	14.7 %	12.1 %	27.4 %	16.5 %

FIGURE 5.1: Example of a DCE Choice Set

Treatment Features	Treatment A	Treatment B	Neither
Length of treatment	**12 months** of 1 pill daily	**9 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Every 2 weeks**	**Every 2 months**	None
Risk of developing active TB after treatment (benefit)	**0 out of 100 (0%)**	**4 out of 100 (4%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**1 out of 100 (1%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**0 out of 100 (0%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? **(tick only one box)**	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

FIGURE 5.2: Respondent Recruitment

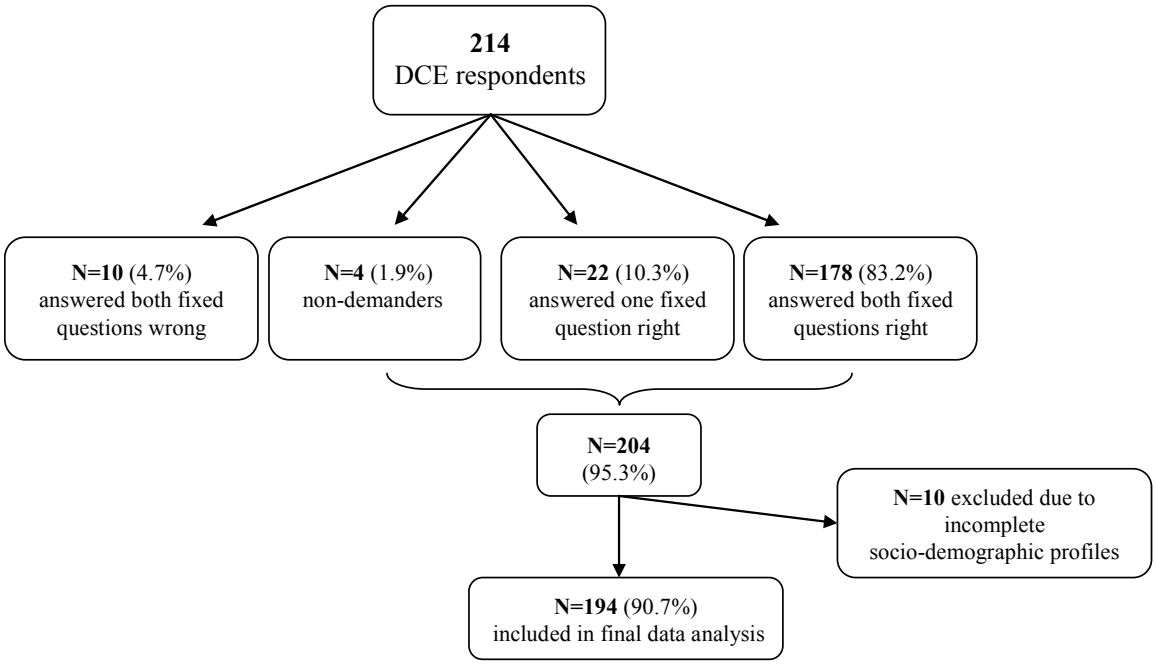
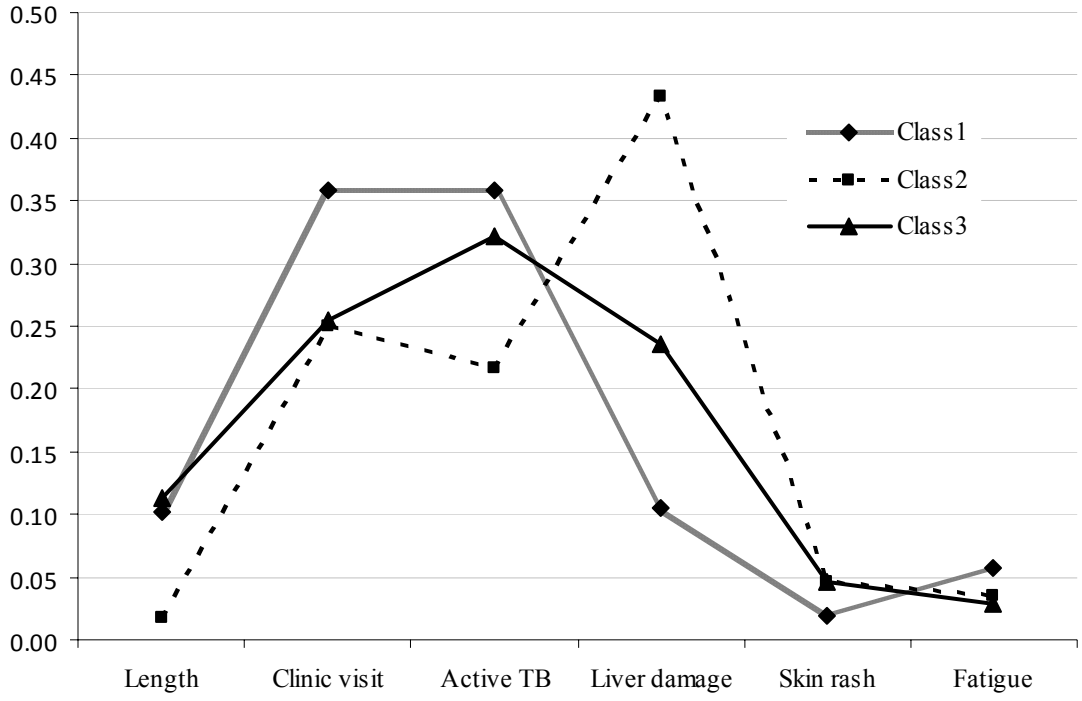


FIGURE 5.3: Relative Importance of Attributes



5.6 REFERENCES

1. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. JAMA 1999;282:677-686.

2. Blumberg HM, Burman WJ, Chaisson RE, Daley CL, Etkind SC, Friedman LN, Fujiwara P, Grzemska M, Hopewell PC, Iseman MD, Jasmer RM, Koppaka V, Menzies RI, O'Brien RJ, Reves RR, Reichman LB, Simone PM, Starke JR, Vernon AA; American Thoracic Society, Centers for Disease Control and Prevention and the Infectious Diseases Society. American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America: treatment of tuberculosis. Am J Respir Crit Care Med 2003;167:603-662.

3. Health Canada. Canadian tuberculosis standards. 6th edition. 2007. www.phac-aspc.gc.ca/tbpc-latb/pubs/pdf/tbstand07_e.pdf (accessed at June 23, 2010).

4. British Columbia Center for Disease Control communicable disease control manual: tuberculosis. http://www.bccdc.ca/NR/rdonlyres/90E2CF31-8621-4082-81B9-BB67B8D1E4F3/0/TB_GF_manual_1999.pdf (accessed at June 23, 2010).

5. Ferebee SH, Mount FW, Anastasiades AA. Prophylactic effects of INH on primary tuberculosis in children. Am Rev TB Pulm Dis 1957;76:942-963.

6. Comstock GW, Ferebee SH, Hammes LM. A controlled trial of community-wide INH prophylaxis in Alaska. Am Rev Respir Dis 1967;95:935-943.

7. International Union Against Tuberculosis Committee on Prophylaxis. Efficacy of various durations of INH preventive therapy for TB: 5 years of follow-up in the IUAT trial. Bull World Health Organ 1982;60:555-564.

8. Shukla SJ, Warren DK, Woeltje KF, Gruber CA, Fraser VJ. Factors associated with the treatment of latent tuberculosis infection among health-care workers at a midwestern teaching hospital. Chest 2002;122:1609-1614.

9. Morisky DE, Ebin VJ, Malotte CK, Coly A, Kominski G. Assessment of tuberculosis treatment completion in an ethnically diverse population using two data sources. Implications for treatment interventions. Eval Health Prof 2003;26:43-58.

10. Kwara A, Herold JS, Machan JT, Carter EJ. Factors associated with failure to complete isoniazid treatment for latent tuberculosis infection in Rhode Island. Chest 2008;133:862-868.

11. Hirsch-Moverman Y, Daftary A, Franks J, Colson PW. Adherence to treatment for latent tuberculosis infection: systematic review of studies in the US and Canada. Int J Tuberc Lung Dis 2008;12:1235-1254.

12. Horsburgh CR Jr, Goldberg S, Bethel J, Chen S, Colson PW, Hirsch-Moverman Y, Hughes S, Shrestha-Kuwahara R, Sterling TR, Wall K, Weinfurter P; Tuberculosis Epidemiologic Studies Consortium. Latent Tuberculosis Infection Treatment Acceptance and Completion in the United States and Canada. Chest 2010;137:401-409.

13. Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl Health Econ Health Policy 2003;2:55-64.

14. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26:661-677.

15. Hensher DA, Rose JM, Greene WH. Applied choice analysis: a primer. Cambridge: Cambridge University Press, 2005.

16. Ryan M, Skåtun D. Modelling non-demanders in choice experiments. Health Econ 2004;13:397-402.

17. Van der Pol M, Cairns J. Estimating time preferences for health using discrete choice experiments. Soc Sci Med 2001;52:1459-1470.

18. McIntosh E, Ryan M. Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences. J Econ Psychol 2002;23:367-382.

19. Bech M, Gyrd-Hansen D. Effects coding in discrete choice experiments. Health Econ. 2005;14:1079-1083.

20. Boxall PC, Adamowicz WT. Understanding heterogeneous preferences in random utility models: a latent class approach. Environmental and Resource Economics 2002;23:421-446.

21. Hole AR. Modelling heterogeneity in patients' preferences for the attribute of a general practitioner appointment. J Health Econ 2008;27:1078-1094.

22. Greene WH, Hensher DA. A latent class model for discrete choices analysis: contrasts with mixed logit. Transportation Research Part B 2003;37:681-698.

23. Ryan M, Gerard K, Amaya-Amaya M (eds.). Using discrete choice experiments to value health and health care. Chapter one: discrete choice experiments in a nutshell. pp:13-46. Dordrecht : Springer, 2008.

24. Urama KC, Hodge ID. Are stated preferences convergent with revealed preferences? Empirical evidence from Nigeria. Ecological Economics. 2006;59:24-37.

25. Van der Pol M, Shiell A, Au F, Johnston D, Tough S. Convergent validity between a discrete choice experiment and a direct, open-ended method: comparison of preferred attribute levels and willingness to pay estimates. Soc Sci Med 2008;67:2043-2050.

26. Ryan M, Netten A, Skåtun D, Smith P. Using discrete choice experiments to estimate a preference-based measure of outcome--an application to social care for older people. J Health Econ 2006;25:927-944.

CHAPTER 6 GENERAL DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Patient-reported outcomes (PROs) have increasingly been used to assess the impact of illness and its treatment(s) in clinical research and also been utilized as the measure of benefits gained from health interventions in economic evaluations to facilitate policy making. Given the value of PRO measurements and the lack of PRO data in tuberculosis (TB) population in the literature, this thesis intended to: (1) measure and compare the health-related quality of life (HRQL) in active TB patients and latent TB infection (LTBI) patients; (2) generate health state utility values (HSUVs) that could be utilized in future economic evaluations; (3) investigate the relationship between HRQL measurements and treatment associated adverse drug reactions (ADRs) in active TB patients; and (4) quantify patients' preferences toward the preventive treatment of LTBI. The results of this work added value to the growing realization of the importance of PROs in the management of TB and other diseases as well.

6.1 SUMMARY AND DISCUSSION OF STUDY FINDINGS

6.1.1 Health-related Quality of Life in Tuberculosis

The HRQL experienced by TB patients was measured using one of the most utilized generic instruments, the Short Form 36 (SF-36) [1]. All the eight subscale scores (i.e. physical functioning (PF), role-physical (RP), bodily pain (BP), general health (GH),

vitality (VT), social functioning (SF), role-emotional (RE), and mental health (MH)) and the two summary scores (i.e. physical component summary (PCS) and mental component summary (MCS)) of SF-36 in active TB patients have been shown to be significantly lower compared to those of LTBI patients, after adjusting for the socio-demographic differences between the two groups of patients. The average SF-36 scores of LTBI patients were observed to be comparable to the U.S. general population level before or at the early stage of the preventive treatment [1,2]. These findings suggested that active TB disease had a significant impact on patients' HRQL status and the impact was encompassing from physical health, psychological well-being to social functioning. The health utility scores, Health Utility Index (HUI) -2 and -3 and Short From 6D (SF-6D), also suggested poorer health status among active TB patients compared to LTBI patients.

To examine the construct validity of SF-36 and the three utility instruments, active TB patients were categorized into five groups according to the level of self-reported disease severity (i.e. very mild, mild, moderate, severe, and very severe). Group average SF-36 scores clearly demonstrated a monotonically decreasing trend with the increasing severity level. In addition, PCS and MCS scores showed moderate to good correlations with HUI-2, HUI-3, SF-6D and Visual Analog Scale (VAS) scores, which provided further evidence to support the validity of SF-36. Other investigators also found that SF-36 displayed acceptable construct validity during the applications among TB population (see Chapter 2 for details).

During the first 3 months of the anti-TB treatment, 24% of our active TB patients developed at least one major ADR event where the treatment had to be discontinued and/or additional treatment for the ADR symptoms was used; and another 33% only had minor ADRs where routine treatment was not disrupted. Common ADRs included skin rash and/or pruritus, gastrointestinal symptoms, liver damage (including mild liver enzyme elevation and hepatitis), paresthesia, fatigue/weakness, visual disturbance, and joint/muscle pain.

We categorized patients into three groups according to the occurrence and severity of ADRs within the 3 months: *No ADR*, *Minor ADR*, and *Major ADR*. Among the 3 groups, the '*No ADR*' group had the highest SF-36 scores and the '*Major ADR*' group had the lowest scores. After adjusting for baseline differences on SF-36 scores and socio-demographic features in regression analyses, patients who developed major ADRs scored significantly lower on two mental health subscales (i.e. VT and MH), the mental health summary (i.e. MCS) and one physical subscale (i.e. PF) of SF-36, compared to patients who had no ADRs. On the other hand, patients who developed only minor ADRs did not score significantly different from those who had no ADRs. The results suggested that major ADRs had significant impact on patients' HRQL and the impact seemed to be more of a mental well-being burden than a physical one to the patients. Minor ADRs had some negative impacts on patients' HRQL but these were not found to be statistically significant in our study. It has to be mentioned that the relatively small sample size in this study (n=89; Chapter 4) might have limited us finding more significant results.

Another important observation was that active TB patients who developed major ADRs within the 3 months of treatment tended to have lower baseline SF-36 scores before the treatment. After adjusting for socio-demographic factors using ordinal logistic regressions, we found that the lower baseline SF-36 score was significantly associated with a higher risk of developing ADRs during the treatment. Poor HRQL status might reflect individuals' known or unknown physical and mental health problems, involvement of risk behaviors, unhealthy lifestyles and their personalities and psychosocial coping styles, which could predispose them to a higher risk of unfavorable health outcomes. The predictive value of self-rated health status has long been one focus of research interests, due to its potential clinical implication [3-7]. In Idler et al. (1990), individuals' responses to a single item "*Would you say your health in general is excellent, very good, good, fair, or poor*?" were found to be significantly associated with the mortality over a 12-year follow-up period among middle-aged males [4]. These consistent findings in the literature strongly suggest that HRQL measurements may have the potential to predict individuals' future health outcomes.

Measuring HRQL and health utility in TB population is a relatively new yet growing research area, given that a wide range of methodologies have been used in the literature, including focus group discussion, individual interview, various generic instruments, respiratory disease-specific instrument and a combination of several condition-specific instruments. Until now, the generic questionnaire, SF-36, was the most frequently used standardized HRQL instrument and displayed acceptable psychometric properties in TB population. There is a lack of one well-established TB-specific HRQL instrument,

although we found that a new TB-specific instrument, DR-12, has been recently developed in India. DR-12 needs further application and systematic psychometric evaluation.

HRQL is a complex and multi-dimensional concept and can be a reflection of patients' demographic characteristics and socioeconomic status besides the illness and its treatment. In this study, we found that older subjects tended to report lower HRQL status than those who were younger. No gender difference was observed in SF-36 scores in our study. Other studies investigating this topic reported some more interesting results (see Chapter 2 for details). These findings also stress the importance of adjusting for potential socio-demographic confounders when analyzing HRQL data.

6.1.2 Health State Utility Values in Tuberculosis

In this study, three indirect utility instruments (SF-6D, HUI-2, and HUI-3) were used to generate HSUV among a sample of active TB and LTBI patients. Overall, the three instruments were able to distinguish patient groups with different levels of disease severity, which supported their construct validity. The three instruments displayed moderate to good agreement with each other and with SF-36 summary scales. However, they yielded quite different utility scores for the same individual. HUI-2 often gave the highest score and SF-6D generated the lowest score. These discrepancies are likely due to the different health dimensions and levels covered in each instrument and the different techniques and population samples they used to obtain the preference weights. A large

body of research has been done in various patient populations that consistently displayed the differences between these utility instruments [8-14]. The key concern is that the difference in these instruments would have an impact on the estimated Quality-adjusted Life Years (QALYs) and incremental cost-utility ratios and would eventually result in different decisions.

We also observed different degrees of ceiling and floor effect problems in the three utility instruments and each instrument appeared to have its own strengths and weaknesses. The distribution of HUI-2 and HUI-3 scores from active TB patients was significantly skewed toward the best utility scores (closer to 1.0), while SF-6D scores appeared to have a relatively normal distribution around a mean value of 0.68. This suggested that the global HUI-2 and HUI-3 scores suffered from a serious ceiling effect problem, with about 25% of active TB patients reporting a HUI-2 score of 1.0 (indicating perfect health) and 21% reporting a HUI-3 score of 1.0. The ceiling effect problem of HUI could be better characterized by looking at the attribute level. For example, 97.6% and 91.7% of our active TB patients chose the best level of "hearing" and "speech", respectively, in HUI-3; and 94% of patients reported the best level of "self-care" in HUI-2 (TABLE 3.4). On the other hand, the floor effect problems of HUI-2 and HUI-3 among TB patients were very limited. SF-6D global scores did not show much evidence of ceiling or floor effects. However, some attributes of SF-6D might have floor effect problems. For example, 63.1% of active TB patients reported the worst level of "role limitation" and 28.6% reported the worst "vitality". Another notable problem of SF-6D was that its global

scores were limited from 0.30 to 1.0, while HUI-2 global scores distributed from 0.13 to 1.00 and HUI-3 scores spanned approximately twice the range of SF-6D scores.

6.1.3 Stated Preferences for the Latent Tuberculosis Infection Treatment

The acceptance rate of the preventive treatment among patients with LTBI has been well below the optimal level in developed countries. To understand patients' preferences when making the decision whether to accept the preventive treatment, we developed and administered a discrete choice experiment (DCE) survey among LTBI patients. The preference estimates derived from both conditional logistic regression and latent class analysis were consistent, reasonable and generally in line with our hypotheses, which supported the theoretical validity of our DCE methodology. All the six studied attributes (i.e. length of treatment, frequency of clinic visit, risk of developing active TB after treatment, risk of developing liver damage, risk of developing skin rash, and risk of developing fatigue) were found to significantly influence respondents' treatment choices. Among the three treatment associated ADRs, respondents were more concerned about the serious liver damage over the mild skin rash and fatigue. Among the six attributes, the length of treatment and the risk of developing skin rash and fatigue were relatively unimportant, compared to the frequency of clinic visit, risk of developing active TB and risk of developing liver damage. A number of previous studies have found that the length of the regimen was a critical predictor to treatment adherence and completion. But our findings showed that when patients made the initial decision on whether to start the treatment, its potential benefits and risks were considered more important than its length.

In British Columbia in 2007, among patients who agreed to initiate the LTBI preventive treatment, only 50% successfully completed the prescribed doses. As we observed, a fair proportion of patients simply stopped the treatment after being on the treatment for a while, even without experiencing side effects. It suggests that treatment decision and adherence are different behaviors and may have different determinants. Understanding patients' preferences and decision-making process might help health care professionals convince more people to initiate the preventive treatment. However, subsequent treatment adherence needs to be studied and targeted with different strategies in order to improve treatment completion rate.

Respondents' preferences were observed to vary significantly by their socio-demographic characteristics and other TB-relevant factors. To examine the preference heterogeneity, we performed latent class analysis. A 3-class latent class choice model best fitted our choice data where the respondents were assigned into three groups according to their preferences: 47% in class-1, 32% in class-2, and 21% in class-3. Significant preference differences across the three groups existed in two attributes, the frequency of clinic visit and the risk of developing liver damage, as indicated by the different sign and magnitude of preference estimates. Five socio-demographic factors were significantly associated with the class membership, including the origin of birth, education level, employment status, having children or not, and use of over-the-counter (OTC) medications.

Important characteristics of the three groups were summarized as following: (1) Individuals in class-1 had the lowest probability of choosing 'Neither' option in the DCE

survey and they were most likely to have children. The risk of developing active TB and frequency of clinic visit were the most important factors influencing their treatment decisions. In terms of the frequency of clinic visit, they preferred monthly clinic visit, which is the current practice at the BC Center for Disease Control (BCCDC) TB clinics. (2) Class-2 respondents had the highest employment rate among the three groups. The risk of liver damage was most concerned factor when class-2 respondents made choices about treatment. They preferred a less frequent clinic visit, every 2 months, compared to class-1 respondents. (3) Class-3 respondents were more likely to be foreign-born, have higher education, be currently unemployed, and take OTC medications. They had the highest probability of choosing the 'Neither' option in the DCE choice tasks. The risk of developing active TB, chance of developing liver damage and frequency of clinic visit had similar importance in their treatment choices. In terms of the frequency of clinic visit, they preferred no clinic visits during the treatment.

We incorporated two identical fixed questions in the survey to check the internal consistency of the DCE survey [15,16]. In total, 4.7% of our respondents were inconsistent on both fixed questions and another 10.3% were inconsistent on one fixed question. This suggested that some respondents might not fully understand the DCE tasks. They might be learning how to make the trade-offs during the survey or they were constructing their preferences in the process because, as some researchers suggested, patients are not used to making treatment decisions without a doctor's guidance [17]. Another potential reason might be the complexity of DCE exercises. Completing the DCE tasks requires heavy information processing in a relatively short time and coherent

judgment involving many outcomes and probabilities. It is well known that people are not very efficient at processing probabilities in a survey. Some researchers are also concerned that people may have problems in weighting and making trade-offs across as many as six attributes at once. Consequently, the respondents might ignore some attributes and only focus on a few or a single attribute(s) or they might make choices in a random and/or inconsistent way [18]. However, also considering the fact that English was the second language to the majority of our respondents, our DCE survey was overall well accepted and understood.

6.2 STUDY STRENGTHS, METHODOLOGICAL CONCERNS, AND LIMITATIONS

As with any study, limitations are inevitable. However, with the great efforts we took at every study stage to minimize their impacts, none of these limitations is expected to significantly affect our findings.

6.2.1 Health-related Quality of Life and Health State Utility Values in Tuberculosis

In this study, we generated the health utility values among TB patients that are much needed to conduct economic evaluations for TB diagnostic tests and new therapeutic modalities as they become available. To our knowledge, till 2007, only one published study measured health utility scores in TB patients [19]. However, the sample of this

study was small and was composed of a mixture of patients with latent TB infection (n=25), current active TB (n=17) and previous TB (n=8), which limited the generalizability of the study findings. Our results were based on a relatively large and homogeneous sample (78 LTBI patients and 84 active TB patients before they started their treatment). The socio-demographic characteristics of our sample were representative to the TB population in North American context, which increases the applicability of our results.

The cross-sectional comparison between three common health utility instruments (HUI-2, HUI-3 and SF-6D) also adds great value to the literature and stresses the importance of standardizing the approaches and instruments used to derive HSUVs. Unfortunately, one common utility instrument, Euro Quality of Life 5D (EQ-5D), was not included in our study, so we did not have the chance to examine its performance and compare it with other utility instruments in TB population.

In this study, we also measured HRQL and health utility among people with LTBI, which has rarely been done in the literature. Although assumed to be otherwise healthy and comparable to the general population, people with LTBI are expected to experience some degree of HRQL decrement, especially after their knowledge of diagnosis and taking the preventive treatment.

6.2.2 Impact of Adverse Drug Reactions on Health-related Quality of Life

This study was the first to investigate the adverse impact of anti-TB treatment associated ADRs on HRQL. More importantly, our findings showed that patients' HRQL status might have the ability to predict their future health outcomes. However, we should be cautious in the interpretation of the results due to some limitations of the study. Firstly, the information of ADRs was obtained retrospectively from a public health database that was not originally maintained for research purpose. Secondly, the sample size was relatively small. However, the fact that we could obtain such significant findings even in a small sample increases our confidence in the results and conclusions.

6.2.3 Stated Preferences toward the Latent Tuberculosis Infection Treatment

The hypothetical nature of DCE raises some methodological concerns, such as its experimental design, sample size and selection, and choice data collection and analysis approaches. Therefore, efforts to improve and evaluate the validity of the methodology have been critical to increase our confidence in the study findings and to establish the creditability of the DCE technique in health economics.

In this study, we ensured the validity of our DCE survey from several aspects. Firstly, at the survey design stage, we carefully chose the treatment attributes primarily based on in-depth patient interviews, along with experts' opinions and literature review findings. The levels of each attribute were selected to reflect the reality and also to cover a wide range

of possibilities. Secondly, we piloted the DCE survey on a small sample of LTBI patients (n=60) before final administration and modified the survey according to patients' performance and feedback. The pre-testing also provided valuable information in terms of how to effectively communicate and explain the DCE task to the potential respondents. Thirdly, during the data collection stage, each respondent completed a 'warm-up' DCE choice with the researcher before they went on completing the survey on their own. Fourthly, we incorporated two fixed consistency-check questions in the survey to check the validity of the DCE. 4.7% of our respondents were inconsistent on both fixed questions. These respondents were likely to not fully understand the DCE choice tasks and might not have made meaningful choices. Therefore, their responses were excluded from the data analysis to ensure the quality of our choice data. Fifthly, in choice data modeling, we performed conditional logistic regression, mixed-effect logistic regression and latent class analysis and decided that a 3-class latent class model best fit our data. Originally used in market research, latent class modeling allows the determination of the number of segments of customer preferences, which provides insights into effective product targeting and strategic positioning [20-23].

One issue in latent class modeling is the determination of the optimal number of latent classes. Since there have been no established statistical tests available for this purpose, common information theoretic criteria such as the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC) are generally recommended as a statistical guide [22,23]. We also considered the significance of parameter estimates, the parsimony of the model and, more importantly, the interpretability of class membership.

Determination of optimal sample size is an important issue in DCE studies, but there is no definitive answer to this question [24-27]. The sample size used in the literature varies significantly from study to study. It is generally accepted that a relatively large sample size is needed to obtain a satisfied level of precision and Louviere et al. (2000) proposed a formula to calculate the minimum sample size needed to achieve a desired level of accuracy using a random sample [24,25]. Amaya-amaya [24] discussed a list of factors that need to be considered before determining the sample size and sampling approaches, e.g. the level of accuracy desired, whether estimates for subgroups or just the overall population needed, data collection method to be used (e.g. self-completed or interviewer-administered), and also research budget constraints.

In our sample of LTBI patients, patients who tended to refuse the preventive treatment might be under represented for two reasons. First, among people who were diagnosed with LTBI, a large proportion of them did not come to the TB clinic for their appointment of discussing the preventive treatment. These individuals likely had already made up their minds to refuse the treatment without discussing it with a TB physician. Second, we observed that patients who eventually refused the treatment were less likely to agree to participate in our study. We were unable to investigate the differences between patients who agreed to participate and those who refused to participate due to ethics constraints.

6.3 CONTRIBUTIONS, IMPACTS AND IMPLICATIONS

The work presented in this thesis focuses on understanding the impact of TB and its treatment from patients' perspectives and offers valuable insights into the clinical

management of TB. In addition, the findings of this work contribute significantly to the current body of literature at a methodological level in outcomes research and help identify some areas for improvement in regard to promoting a wider and more appropriate application of PRO measures in clinical practice and research.

The clinical management of patients with TB has been exclusively focused on microbiological 'cure'. Our work demonstrated the importance and usefulness of measuring HRQL in TB population. Firstly, HRQL provides more insights into human concerns that could add supplementary values to traditional physical and laboratory indicators. HRQL measures areas that are more familiar and important to patients. Considering HRQL measurements in clinical practice could adequately reflect the impact of illnesses, promote the communication between health care professionals and patients, encourage patients' involvement, and improve treatment adherence. Secondly, we found that HRQL measurements might carry a lot more clinically relevant information that appears to be able to predict patients' future health outcomes. The predictive potential of HRQL measurements would help health care professionals with disease management and health care resource allocation. For example, improving the HRQL of those high-risk patients through lifestyle modification consultation and social-psychological support might be able to improve their outcomes to medical interventions.

HSUVs were generated for both active TB patients and LTBI patients. These utility values could be utilized to calculate QALYs in future economic evaluations to justify the strategies and resources used in TB control and prevention. The SF-6D, HUI-2 and HUI-3 all appeared to be appropriate health utility measures for TB patients in that they were

able to discriminate between patients with different levels of severity and were well correlated with other PRO measurements. However, HUI-2 and HUI-3 had significant ceiling effect problems both at global and attribute levels, while some attributes of SF-6D had floor effect problems. Overall, it is hard to conclude that any one instrument is the best and should be chosen over the others to measure health utility in TB patients. Each instrument is designed to bring unique strengths to different situations and is deemed to have its limitations at the same time. Researchers should weigh each unique research situation and accordingly choose the right instrument or the right combination of instruments.

As a measure of effectiveness combining both quality and quantity of health, one advantage of QALY is that it allows the comparison across health interventions and patient populations. Our findings suggested that the three commonly used instruments would not generate comparable utility scores for the same individual within a study. The choice of instrument would have considerable impact on the calculation of QALY and the resulting incremental cost–utility ratio [28,29]. For example, McDonough et al. found that EQ-5D was more likely to provide favorable incremental cost–utility ratios than SF-6D [29]. This incomparability would also make the comparisons across disease populations or studies more difficult. It has been suggested that sensitivity analysis should be conducted that covers a wide range of health utility scores that could be obtained from different utility instruments [30,31].

To our knowledge, the study presented in Chapter 5 was the first to evaluate patients' preferences for the LTBI preventive treatment using a DCE technique. Our findings were generally plausible and were congruent with our hypotheses, which greatly supported the theoretical validity of our DCE survey. We observed significant preference heterogeneity among respondents which underlines the importance of choosing proper statistical approaches when analyzing choice data. The use of latent class analysis allowed us to categorize and quantify respondents' preferences while taking into account their socio-demographic information. The findings provide informative insights into understanding how respondents perceive differently on risks and benefits and weight treatment attributes differently in decision-making. Our results could possibly help health care professionals in effective communication with patients when advocating the preventive treatment, which would lead to improved treatment acceptance. Methodologically, our study showed that the DCE technique could be an internally consistent and valid tool to understand patients' preferences toward health products and interventions and to inform health care practice. Adding respondents' socio-demographic information into choice data modeling would provide additional valuable insights.

6.4 RECOMMENDATIONS AND FUTURE RESEARCH

This work yielded many novel and interesting results that could help better understand the role of PRO measurements in clinical management of TB patients. Methodologically, the findings greatly contributed to the current body of literature in outcomes research and health economics and identified some potential areas for future improvement.

The finding that HRQL measurements might have the potential ability to predict patients' future health outcomes has significant clinical implication. This is an area worthy of continued research efforts. The mechanism underlying the association between patient-reported health status and future clinical outcomes is not well understood, but deserves closer investigation. Future prospective studies with larger sample sizes are needed to better understand the predictive value of patients' self-reported health status and to investigate whether interventions could be launched to improve outcomes.

The incomparability of the commonly used HSUV instruments would significantly compromise the creditability of using QALYs to facilitate health care decision-making. Therefore, it is important to further improve our understanding of the impact of using different utility instruments on the final decisions, by obtaining more comparative data with more than one utility instrument applied in various patient populations [28]. Greater reporting transparency is also helpful and strongly recommended [29].

The assessment of psychometric properties is an important component in the application of PRO measurements. In our study, we focused on assessing the construct validity of SF-36, HUI-2, HUI-3, and SF-6D. Other researchers also have reported satisfying internal consistency reliability and test-retest reliability of SF-36 in various TB populations. Another critical property of HRQL and HSUV instruments is the responsiveness. We did not have a chance to assess the responsiveness of the instruments in this study. Future research should be directed to assess the responsiveness of these instruments when applied to the TB population. Despite the fact that some consider

responsiveness to be an essential property of an evaluative instrument, there is an evident lack of clarity in the literature about its definition, what it purports to measure, and how it should be quantified and evaluated [32]. In addition, some researchers argue that responsiveness is just a longitudinal aspect of validity; while others believe it should be a separate property of an instrument.

One barrier to the wide and meaningful use of HRQL measures in clinical practice is the lack of information necessary to interpret the scores. It is important to establish the relevance and interpretability of HRQL measurements anchored to clinical criteria that are familiar to health care professionals. One approach is to document minimally clinically important differences (MCID) for commonly used HRQL instruments and standardize the approaches of estimating MCID. MCID can be defined as the minimal change in measurements that signifies an important improvement or worsening in health, typically according to patients' perception [33-35]. For example, Samsa et al. (1999) suggested that a 3-5 point in SF-36 scores represented a MCID [36]. It was also reported that, for a number of similarly structured disease-specific HRQL instruments, a change of 0.5 point on the 7-point scale approximates the MCID in the measurement scores [37,38]. This kind of information will help to distinguish patient groups, assess treatment effect over time, and also enable sample size calculation for future outcomes research studies.

Measurement of health status is contentious due to the complex and abstract nature of health itself. Generic HRQL instruments measure common and important health constructs, and they allow for comparisons across different patient populations. A

disease/condition-specific HRQL instrument is designed to focus on health constructs that are most relevant to a given patient population. They are assumed to be superior to generic measures to detect a small but important difference or change when evaluating HRQL in a specific patient population. Because generic and disease-specific HRQL instruments both have strengths and are conceptually distinct, it is often recommended that generic and disease-specific instruments be administered together to achieve a complete assessment [39-41]. However, there is no established TB-specific instrument that has been described in the literature. Therefore, a valid and reliable specific instrument is needed in measuring HRQL among TB population.

There might be some challenges worth noting in developing a TB-specific instrument. TB can affect almost any part of the human body and could present as pulmonary or extra-pulmonary disease or merely latent infection, which would have very different clinical presentations and impose various impacts on patients' well-being and functioning. Most existing HRQL instruments are developed using classical test theory (CTT). Under CTT, in order to obtain higher measurement precision, a large number of items or questions are needed to cover a wide range of problems experienced by the targeted population [42]. This strategy would increase patients' response burden and it is very likely that many items would become irrelevant and uninformative for a particular patient. The application of item response theory (IRT) and item banking may provide a promising future for HRQL measurements in TB and other health conditions as well [42,43].

Combining the modern IRT and advanced computer technologies, computerized adaptive testing (CAT) might have the potential to optimize measurement precision without increasing response burden [44-47]. The success of CAT relies on developing an operational item bank, which is a large collection of items that are measuring the same health construct and are calibrated onto a common scale using IRT-based approaches. For example, a health construct of "depression" can be represented by a large number of items that range from the minimum to a maximum level of depression. Through the application of CAT, unique items can be tailored and administered to each individual respondent and scores are still comparable across different subsets of items. The ceiling and floor effect problems that are common with most current HRQL/heath utility instruments can also be minimized.

Despite most HRQL instruments have been developed in the English language, they are often used in a culturally diverse patient population or internationally. In our study, the majority of TB patients were immigrants originally from Asian countries. Although some of them have been living in Canada for years and were comfortable with the English language, we did not know whether the cultural difference would have an impact on the validity of the instruments. This underlines the need for proper linguistic translation and cultural adaptation in order to retain the validity and reliability when applying HRQL and health utility instruments cross-culturally. There have been guidelines in the literature outlining the comprehensive process of cross-cultural adaptation [48-49].

In DCE choice exercises, respondents have to weigh and trade-off a number of attributes at a time which increases information load and respondent cognitive burden and may consequently jeopardize the study validity. The need to limit the number of attributes in a DCE has been well recognized and accepted in health economics [24,50]. In some other stated preference techniques, e.g. standard gamble and trade-off, respondents only focus on changes in one attribute at a time. These tasks are relatively easier to understand and complete. For example, Kopec et al. (2007) used a trade-off approach to measure patients' preferences for osteoarthritis medications in terms of the maximum acceptable risk increments for a few medication-associated adverse events [51]. As the authors discussed, the application of these techniques could be limited due to the focus on a single attribute at a time, because decision-makings in reality often involves multiple attributes.

Another alternative approach to reduce the impact of multiple attributes is the hierarchical information integration (HII) [18,52,53]. It assumes that, when facing complex decision-making tasks, individuals use hierarchical strategies, where they first categorize all involved attributes into meaningful subsets. Therefore, separate experimental designs for each of these subsets can be constructed. Then, a bridging design would integrate these part-worth utilities into an overall utility function [53]. This approach has been used in marketing and transportation and recently has been proposed to be applied in health economics by van Helvoort-Postulart et al. (2005) [18]. The evaluation of health care interventions often involves many attributes and HII may

provide a promising approach of handling the complexity of health decision-making to reduce information overload and respondent burden in DCE studies [18].

Our results revealed that patients' preferences varied significantly when making decisions on LTBI preventive treatment. There has been research attempting to identify factors that may affect or predict individual judgment of their preferences, e.g. socio-demographic features, individuals' risk attitudes (risk seeking vs. risk averse) and individuals' perceived control over their own health. In our study, we found a number of patients' socio-demographic factors were significantly associated with the preference heterogeneity that we observed. Previous studies have yielded few results on this topic. One study showed that osteoarthritis patients' preferences varied widely in therapeutic decisions, but their clinical, socio-demographic and psychological characteristics were not found to explain the heterogeneity [51]. This finding was inconsistent with our results. However, this study applied a probabilistic threshold technique, different from our DCE methodology.

When evaluating preferences, it is important to distinguish between consumers as patients and consumers as the general public (potential but not yet patients). The preference on treatment attributes and the relative importance of these attributes may be very different depending on whether the respondents are patients or the general public [54]. Future research should explore how patients' preferences are different from the general public and even health care professionals in a quantifiable way.

Validation of DCE results is an important research area and also a critical step to establish the creditability of DCE in health economics. The internal consistency of DCE is often assessed through incorporating consistency-check questions in the survey. The theoretical validity is often judged by examining the sign, significance and plausibility of preference estimates derived from choice modeling. Another important property of DCE is test-retest reliability or stability. The classical utility preference theory assumes that human preferences stay stable. However, it is recognized that individuals' preferences might change over time due to unknown factors, such as individual interactions in a social context [55,56]. Few health economic studies have assessed the test-retest reliability of the DCE responses [18,56,57]. Future prospective studies are needed to investigate whether preferences change over time and how the change is related to individuals' socio-demographic characteristics and other factors.

The most expressed concern over the validity of DCE and other stated preference techniques is that the measured "stated" preference may not reflect respondents' actual behavior in a real market situation (revealed preferences) [55,58-61]. DCE choices are made in a controlled experimental context, which may simplify the real-life decision-making situation. In addition, the respondents may be not willing to or not able to express their "true" preferences. The most effective solution to this concern is to compare stated choices with actual decision-making behavior, which would provide the strongest evidence to support the validity of the DCE technique [55,60,61]. This is also an important future avenue for research in applying DCE techniques in health care.

Lastly, since various PRO measurements have been increasingly advocated to be used in health care practice and research, there is a great need of international standardization of terminology and definitions in outcomes research, PROs instruments and their applications, data analysis, and result interpretation.

6.5 CONCLUSIONS

PRO measurements appear to be a promising line of research in order to provide evidence-based and patient-centered care in clinical research and practice. This work illustrates the value and importance of incorporating PRO measurements in TB population and underlines the need of international standardization in outcomes research community.

Active TB disease and its treatment associated ADRs had substantial and encompassing impacts on patients' HRQL. Patients' HRQL and other self-reported health status measures might have the ability to predict future health outcomes. Given the absence of a valid and reliable TB-specific instrument, the generic instrument, SF-36, was most often used to measure HRQL in TB population and displayed acceptable construct validity in our study.

Three commonly used indirect utility instruments (HUI-2, HUI-3 and SF-6D) appeared to be valid tools to be used in TB population but they generated different scores for the same individual, which may have significant impact on economic evaluation and final

decision-making. However, there has been little evidence of which instrument is overall the best. Given the increasing role of economic evaluation, it is important to further improve our understanding of the impact of using different utility instruments in future research.

This works also shows how a DCE technique could be a promising tool to understand patients' preferences for health care products and services and to possibly inform health care practice. Significant treatment preference heterogeneity was observed among our respondents, which highlights the importance of choosing the proper statistical approach in choice modeling. Our latent class analysis findings showed how the consideration of socio-demographic variables in choice data analysis could provide valuable insights into understanding the heterogeneity of respondents' decision-making behavior. To extend the usefulness of DCE in informing health care practice, future efforts should be directed to validation of this technique.

6.6 REFERENCES

1. Ware JE Jr., Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. User's manual for the SF-36v2™ health survey (2nd ed). Lincoln, RI: QualityMetric Incorporated, 2007.

2. Marra CA, Marra F, Colley L, Moadebi S, Elwood RK, Fitzgerald JM. Health-related quality of life trajectories among adults with tuberculosis: differences between latent and active infection. Chest 2008;133:396-403.

3. Mossey JM, Shapiro E. Self-rated health: a predictor of mortality among the elderly. Am J Public Health 1982;72:800-808.

4. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. Am J Public Health. 1990;80:446-452.

5. Gold M, Franks P, Erickson P. Assessing the health of the nation. The predictive validity of a preference-based measure and self-rated health. Med Care 1996;34:163–177.

6. Franks P, Gold MR, Fiscella K. Sociodemographics, self-rated health, and mortality in the US. Soc Sci Med 2003;56:2505–2514.

7. Sehlen S, Lenk M, Hollenhorst H, Schymura B, Aydemir U, Herschbach P, Dühmke E. Quality of life (QoL) as predictive mediator variable for survival in patients with intracerebral neoplasma during radiotherapy. Onkologie 2003;26:38–43.

8. Longworth L, Bryan S. An empirical comparison of the EQ-5D and SF-6D in liver transplant patients. Health Econ 2003;12:1061–1067.

9. Marra CA, Esdaile JM, Guh D, Kopec JA, Brazier JE, Koehler BE, Chalmers A, Anis AH. A comparison of four indirect methods of assessing utility values in rheumatoid arthritis. Med Care 2004;42:1125-1131.

10. Szende A, Svensson K, Ståhl E, Mészáros A, Berta GY. Psychometric and utility based measures of health status of asthmatic patients with different disease control level. Pharmacoeconomics 2004;22:537–547.

11. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures. Int J Audiol 2005;44:157–163.

12. Bryan S, Longworth L. Measuring health related utility: why the disparity between EQ-5D and SF-6D? Eur J Health Econ 2005;6:253–260.

13. McDonough CM, Grove MR, Tosteson TD, Lurie JD, Hilibrand AS, Tosteson AN. Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among Spine Patient Outcomes Research Trial (SPORT) participants. Qual Life Res 2005;14:1321–1332.

14. Stavem K, Froland SS, Hellum KB. Comparison of preferencebased utilities of the 15D, EQ-5D and SF-6D in patients with HIV/AIDS. Qual Life Res 2005;14:971–980.

15. Van der Pol M, Cairns J. Estimating time preferences for health using discrete choice experiments. Soc Sci Med 2001;52:1459-1470.

16. McIntosh E, Ryan M. Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences. J Econ Psychol 2002;23:367-382.

17. Ryan M, San Miguel F. Revisiting the axiom of completeness in health care. Health Economics 2003;12:295–307.

18. van Helvoort-Postulart D, Dellaert BG, van der Weijden T, von Meyenfeldt MF, Dirksen CD. Discrete choice experiments for complex health-care decisions: does hierarchical information integration offer a solution? Health Econ 2009;18:903-920.

19. Dion MJ, Tousignant P, Bourbeau J, Menzies D, Schwartzman K. Measurement of health preferences among patients with tuberculous infection and disease. Medical Decision Making 2002;22:S102-114.

20. Boxall PC, Adamowicz WT. Understanding heterogeneous preferences in random utility models: a latent class approach. Environmental and Resource Economics 2002;23:421-446.

21. Greene WH, Hensher DA. A latent class model for discrete choices analysis: contrasts with mixed logit. Transportation Research Part B 2003;37:681-698.

22. Hole AR. Modelling heterogeneity in patients' preferences for the attribute of a general practitioner appointment. J Health Econ 2008;27:1078-1094.

23. Ruto E, Garrod G, Scarpa R. Valuing animal genetic resources: a choice modeling application to indigenous cattle in Kenya. Agricultural Economics, International Association of Agricultural Economists 2008;38:89-98.

24. Amaya-Amaya M. Chapter one: discrete choice experiments in a nutshell. In: Ryan M, Gerard K, Amaya-Amaya M (eds.). Using discrete choice experiments to value health and health care. pp:13-46. Dordrecht: Springer, 2008.

25. Louviere JJ. Why stated preference discrete choice modelling is NOT conjoint analysis (and what SPDCM IS). 2000. Memetrics white paper.

26. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26:661-677.

27. Hensher D, Rose J, Greene W. Applied Choice Analysis: A Primer. Cambridge: Cambridge University Press, 2005.

28. Joore M, Brunenberg D, Nelemans P, Wouters E, Kuijpers P, Honig A, Willems D, de Leeuw P, Severens J, Boonen A. The impact of differences in EQ-5D and SF-6D utility scores on the acceptability of cost-utility ratios: results across five trial-based cost-utility studies. Value Health 2009;13:222-229.

29. McDonough CM, Tosteson ANA. Measuring preferences for cost-utility analysis. How choice of method may influence decision making. Pharmacoeconomics 2007;25:93–106.

30. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI®): concepts, measurement properties and applications. Health Qual Life Outcomes 2003;1:54.

31. Marra CA, Marion SA, Guh DP, Najafzadeh M, Wolfe F, Esdaile JM, Clarke AE, Gignac MA, Anis AH. Not all "quality-adjusted life years" are equal. J Clin Epidemiol 2007;60:616-624.

32. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349-362.

33. Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, Brooks P, Tugwell P. .Minimal clinically important differences: review of methods. J Rheumatol 2001;28:406–412.

34. Tubach F, Wells GA, Ravaud P, Dougados M. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. J Rheumatol 2005;32:2025–2029.

35. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis 2007;66(Suppl 3):iii40-iii41.

36. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics

1999;15:141-155.

37. Younossi ZM, Guyatt G, Kiwi M, Boparai N, King D. Development of a disease specific questionnaire to measure health related quality of life in patients with chronic liver disease. Gut 1999;45:295-300.

38. Rowan PJ, Al-Jurdi R, Tavakoli-Tabasi S, Kunik ME, Satrom SL, El-Serag HB. Physical and psychosocial contributors to quality of life in veterans with hepatitis C not on antiviral therapy. Journal of Clinical Gastroenterology 2005;39:731-736.

39. Leong KP, Yeak SC, Saurajen AS, Mok PK, Earnest A, Siow JK, Chee NW, Yeo SB, Khoo ML, Lee JC, Seshadri R, Chan SP, Tang CY, Chng HH. Why generic and disease-specific quality-of-life instruments should be used together for the evaluation of patients with persistent allergic rhinitis. Clin Exp Allergy 2005;35:288–298.

40. Engstrom CP, Persson LO, Larsson S, Sullivan M. Health-related quality of life in COPD: why both disease-specific and generic measures should be used. Eur Respir J 2001;18:69–76.

41. Brommels M, Sintonen H. Be generic and specific: quality of life measurement in clinical studies. Ann Med 2001;33:319–322.

42. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of item response theory. Newbury Park, Calif: Sage Publications, 1991.

43. Guo N, Marra CA, Marra F. Responses to comments of Weis. Health Qual Life Outcomes 2010;8:6.

44. Bjorner JB, Kosinski M, Ware JE Jr. Calibration of an item pool for assessing the burden of headaches: an application of item response theory to the headache impact test (HIT). Qual Life Res 2003;12:913-933.

45. Lai JS, Cella D, Dineen K, Bode R, Von Roenn J, Gershon RC, Shevrin D. An item bank was created to improve the measurement of cancer-related fatigue. J Clin Epidemiol 2005;58:190-197.

46. Kopec JA, Sayre EC, Davis AM, Badley EM, Abrahamowicz M, Sherlock L, Williams JI, Anis AH, Esdaile JM. Assessment of health-related quality of life in arthritis: conceptualization and development of five item banks using item response theory. Health.Qual.Life.Outcomes 2006;4:33.

47. Walter OB, Becker J, Bjorner JB, Fliege H, Klapp BF, Rose M. Development and evaluation of a computer adaptive test for 'Anxiety' (Anxiety-CAT). Qual.Life Res 2007;16(Suppl 1):143-155.

48. Beaton DE, Bombardier C, Guillemin F, Ferraz MB: Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186-3191.

49. Bullinger M, Alonso J, Apolone G, Leplege A, Sullivan M, Wood-Dauphinee S, Gandek B, Wagner A, Aaronson N, Bech P, et al.: Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. J Clin Epidemiol 1998;51:913-923.

50. Lancsar E, Donaldson C. Discrete choice experiments in health economics: distinguishing between the method and its application. European Journal of Health Economics 2005;6:314–316.

51. Kopec JA, Richardson CG, Llewellyn-Thomasc H, Klinkhoffd A, Carswelle A, Chalmersd A. Probabilistic threshold technique showed that patients' preferences for specific trade-offs between pain relief and each side effect of treatment in osteoarthritis varied. J Clin Epidemiol 2007;60:929-938.

52. Louviere J J. Hierarchical Information Integration: a new method for the design and analysis of complex multi-attribute judgment problems. Advances in Consumer Research 1984;11:148–155.

53. Molin EJE, Timmermans HJP. Hierarchical information integration experiments and integrated choice experiments. Transport Reviews 2009;29:635 – 655.

54. Ryan M, Shackley P. Assessing the benefits of health care: how far should we go? Quality in Health Care 1995;4:207–213.

55. Urama KC, Hodge ID. Are stated preferences convergent with revealed preferences? Empirical evidence from Nigeria. Ecological Economics. 2006;59:24-37.

56. San Miguel F, Ryan M, Scott A. Are preferences stable? The case of health care. Journal of Economic Behavior and Organization 2002; 48: 1–14.

57. Phillips KA, Maddala T, Johnson FR. Measuring preferences for health care interventions using conjoint analysis: an application to HIV testing. Health Services Research . 2002;37:1681–1705.

58. Loureiro ML, McCluskey JJ, Mittelhammer MC. Are stated preferences good predictors of market behavior? Land Economics 2003;79:44-55.

59. Van der Pol M, Shiell A, Au F, Johnston D, Tough S. Convergent validity between a discrete choice experiment and a direct, open-ended method: comparison of preferred attribute levels and willingness to pay estimates. Soc Sci Med 2008;67:2043-2050.

60. Delaney L, O'Toole F. Decomposing demand for public expenditure in Ireland. Applied Economics Letters 2007;14:1091–1095.

61. Lampi E, Orth M. Who visits the museums? A comparison between stated preferences and observed effects of entrance fees. KYKLOS 2009;62:85–102.

APPENDIX 1: ETHICS APPROVAL CERTIFICATE FOR CHAPTER 3 AND 4 STUDY

Project Title: Development of a Quality of Life Instrument for Patients Living with Tuberculosis

The University of British Columbia
Office of Research Services and Administration
Behavioural Research Ethics Board

Certificate of Approval

PRINCIPAL INVESTIGATOR	DEPARTMENT	NUMBER
Marra, F.	Pharmaceutical Sciences	**B04-0048**

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT

UBC Centre for Disease Control ,

CO-INVESTIGATORS:

Fitzgerald, Mark, Medicine; Marra, Carlo, Medicine; Moadebi, Susanne,

SPONSORING AGENCIES

TITLE:

Development of a Quality of Life Instrument for Patients Living with Tuberculosis

APPROVAL RENEWED DATE	TERM (YEARS)	
MAR 17 2005	1	

CERTIFICATION:

The protocol describing the above-named project has been reviewed by the Committee and the experimental procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval of the Behavioural Research Ethics Board by one of the following:
Dr. James Frankish, Chair,
Dr. Cay Holbrook, Associate Chair,
Dr. Susan Rowley, Associate Chair
Dr. Anita Hubley, Associate Chair

This Certificate of Approval is valid for the above term provided there is no change in the experimental procedures

APPENDIX 2: ETHICS APPROVAL CERTIFICATE FOR CHAPTER 5 STUDY (PART ONE: PATIENT INTERVIEW)

Project Title: A Focus Group Study to Understand People's Perspectives on Treatment of Latent Tuberculosis Infection

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - FULL BOARD

PRINCIPAL INVESTIGATOR:	INSTITUTION / DEPARTMENT:	UBC BREB NUMBER:
Fawziah Marra	UBC/Pharmaceutical Sciences	H08-03057

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
BC Centre for Disease Control	BC Centre for Disease Control

Other locations where the research will be conducted:
N/A

CO-INVESTIGATOR(S):
J. Mark FitzGerald
Richard K. Elwood
Carlo Marra
Na Guo

SPONSORING AGENCIES:
N/A

PROJECT TITLE:
A Focus Group Study to Understand People's Perspectives on Treatment of Latent Tuberculosis Infection

REB MEETING DATE:	CERTIFICATE EXPIRY DATE:
January 22, 2009	January 22, 2010

DOCUMENTS INCLUDED IN THIS APPROVAL:	DATE APPROVED:
	February 11, 2009

Document Name	Version	Date
Protocol:		
Research Proposal_Using Discrete Choice Experiment to Evaluate People's Preferences for Treatment of Latent Tuberculosis Infection	2	February 3, 2009
Research Proposal_Using Discrete Choice Experiment to Evaluate People's Preferences for Treatment of Latent Tuberculosis Infection	N/A	December 18, 2008
Consent Forms:		
Consent Form_Focus Group Study on LTBI treatment (Part one for DCE study)	N/A	December 18, 2008
Consent Form_Focus Group Study on LTBI treatment (Part one for DCE study)_Mandarin	1	February 3, 2009
Consent Form_Focus Group Study on LTBI treatment (Part one for DCE study)_English	2	February 3, 2009
Advertisements:		
Recruitment Advertisement_Focus Group Study on LTBI Treatment	N/A	December 18, 2008
Recruitment Advertisement_Focus Group Study on LTBI Treatment	2	February 3, 2009
Questionnaire, Questionnaire Cover Letter, Tests:		
Subject Information Form_Focus Group Study on LTBI Treatment	2	February 3, 2009
Interview Guide_Focus Group Study on LTBI Treatment	N/A	December 18, 2008
Subject Information Form_Focus Group Study on LTBI Treatment	N/A	December 18, 2008
Letter of Initial Contact:		
Scripts of Initial Contact_Focus Group Study	N/A	December 18, 2008

The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair
Dr. Ken Craig, Chair
Dr. Jim Rupert, Associate Chair
Dr. Laurie Ford, Associate Chair
Dr. Daniel Salhani, Associate Chair
Dr. Anita Ho, Associate Chair

APPENDIX 3: ETHICS APPROVAL CERTIFICATE FOR CHAPTER 5 STUDY (PART TWO: DISCRETE CHOICE EXPERIMENT SURVEY)

Project Title: Using Discrete Choice Experiment to Evaluate People's Preferences for Treatment of Latent Tuberculosis Infection

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - MINIMAL RISK AMENDMENT

PRINCIPAL INVESTIGATOR:	DEPARTMENT:	UBC BREB NUMBER:
Fawziah Marra	UBC/Pharmaceutical Sciences	H09-01442

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
BC Centre for Disease Control	BC Centre for Disease Control

Other locations where the research will be conducted:
N/A

CO-INVESTIGATOR(S):
J. Mark FitzGerald
Richard K. Elwood
Carlo Marra
Na Guo

SPONSORING AGENCIES:
N/A

PROJECT TITLE:
Using Discrete Choice Experiment to Evaluate People's Preferences for Treatment of Latent Tuberculosis Infection

Expiry Date - Approval of an amendment does not change the expiry date on the current UBC BREB approval of this study. An application for renewal is required on or before: July 28, 2010

AMENDMENT(S):	AMENDMENT APPROVAL DATE:
	September 24, 2009

Document Name	Version	Date
Questionnaire, Questionnaire Cover Letter, Tests:		
DCE survey and demographic survey_updated	September 12, 2009	September 12, 2009

The amendment(s) and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair
Dr. Ken Craig, Chair
Dr. Jim Rupert, Associate Chair
Dr. Laurie Ford, Associate Chair
Dr. Anita Ho, Associate Chair

APPENDIX 4: SUPPLEMENT TO CHAPTER 4

The *integrated* Public Health Information System (*i*PHIS) is "*a health record and reporting system that supports public health interventions, tracking, follow-up, case management and reporting*" [1]. In our study, the definition of ADR adopted was "*noxious or unintended response to a drug, which occurs at doses normally used in human for prophylaxis, diagnosis or treatment*" [2]. By retrospectively reviewing physician narratives and nurse notes kept in *i*PHIS, ADRs were identified using the following criteria: (1) recognized adverse drug reaction to the suspected medication; (2) temporal sequence after the suspected medication; (3) the adverse reaction had been relieved or resolved after dose reduction or withdrawal of the suspected medication; (4) the present adverse reaction could not be explained by any other known conditions or predispositions of the patient; (5) the adverse reaction reappeared on rechallenge, or laboratory tests showed toxic drug levels or metabolic disturbances, which explained the present symptoms. An ADR was classified as *definite* if all the five criteria are satisfied; *possible* if the first four criteria were satisfied; *probable* if the first three criteria were satisfied. An ADR was defined as *unlikely* if the relevant information could not be obtained, or if a reasonable temporal sequence could not be established, or if it was more likely to be related to other health conditions or predispositions other than the treatment.

Among the patients enrolled, 15 were excluded from the study due to various reasons: 3 had abnormal liver function at the beginning of treatment; 1 had severe HIV/AIDS; 1 was on anti-TB treatment for less than 3 months; 10 failed to complete baseline and/or 3-month HRQL assessments. TABLE A1 presents the baseline socio-demographic characteristics of the included 89 patients. Patients were categorized into three groups

according to whether they had *No ADR*, *Minor ADR* or *Major ADR* within the 3 months of treatment. TABLE A2 compares the socio-demographic differences across the 3 groups. We performed ordinal logistic regression to examine the association between the occurrence of ADRs and baseline SF-36 scores. The results are shown in TABLE A3.

TABLE A1: Baseline Socio-demographic Characteristics

	All subjects (N=89)
Age (mean in yrs, SD)	49.2 (18.9)
Sex (n, % female)	50 (56.2%)
Foreign-born (n, % yes)	78 (87.6%)
Race (n, %)	
Asian or Pacific Islander	53 (59.6%)
East Indian (South Asian)	10 (11.2%)
Caucasian	14 (15.7%)
Black	3 (3.4%)
Aboriginal	3 (3.4%)
Others	6 (6.7%)
Marital Status (n, %)	
Married	49 (55.1%)
Single (never married)	27 (30.3%)
Others	13 (14.6%)
Smoking (n, % yes)	21 (23.6%)
Alcohol use (n, % yes)	12 (13.5%)
Co-morbidity (n, % yes)	41 (46.1%)

TABLE A2: Socio-demographic Comparison by ADR Occurrence and Severity

	Patients who had No ADR (N=39)	Patients who had Minor ADR (N=29)	Patients who had Major ADR (N=21)	P-Value
Age (mean in yrs, SD)	51.1 (20.8)	48.4 (17.8)	47.0 (17.2)	0.70
Sex (n, % female)	18 (46.2%)	17 (58.6%)	15 (71.4%)	0.16
Foreign-born (n, % yes)	35 (89.7%)	27 (93.1%)	16 (76.2%)	0.24
Race (n, %)				0.21
Asian	26 (66.7%)	24 (82.8%)	13 (61.9%)	
Others	13 (33.3%)	5 (17.2%)	8 (38.1%)	
Marital Status (n, %)				0.28
Married	18 (46.2%)	17 (58.6%)	14 (66.7%)	
Others	21 (53.8%)	12 (41.4%)	7 (33.3%)	
Smoking (n, % yes)	9 (23.1%)	6 (20.7%)	6 (28.6%)	0.81
Alcohol use (n, % yes)	4 (10.3%)	4 (13.8%)	4 (19.1%)	0.57
Co-morbidity (n, % yes)	17 (43.6%)	11 (37.9%)	13 (61.9%)	0.22

TABLE A3: Parameter Estimates in Ordinal Logistic Regression Models [+]

Predictor variable [§] (Baseline SF-36 scores)	Estimate (β)	Standard Error	Odds Ratio (OR)	95% C.I .for OR		P-value
				Lower	Upper	
Physical Functioning (PF)	-0.06	0.02	0.94	0.91	0.97	<0.01 *
Role-physical (RP)	-0.07	0.02	0.94	0.91	0.97	<0.01 *
Bodily pain (BP)	-0.07	0.02	0.93	0.90	0.97	<0.01 *
General Health (GH)	-0.05	0.02	0.95	0.90	0.99	0.03 *
Vitality (VT)	-0.05	0.02	0.95	0.92	0.99	0.01 *
Social Functioning (SF)	-0.06	0.02	0.94	0.91	0.97	<0.01 *
Role-emotional (RE)	-0.05	0.01	0.95	0.93	0.98	<0.01 *
Mental Health (MH)	-0.06	0.02	0.94	0.90	0.98	<0.01 *
Physical summary (PCS)	-0.09	0.02	0.92	0.88	0.96	<0.01 *
Mental summary (MCS)	-0.06	0.02	0.94	0.91	0.97	<0.01 *

[+] The dependent variable is occurrence of ADR with 3 levels: No ADR, Minor ADR, Major ADR

[§] Adjusted variables include age, sex, race, marital status, smoking, comorbidity

* Statistical significance

REFERENCES

1. Integrated Public Health Information System. British Columbia Centre for Disease Control. http://www.bccdc.org/content.php?item=10 (accessed at June 20, 2010).

2. World Health Organization (WHO). International drug monitoring: the role of national centres. Report of a WHO meeting. World Health Organ Tech Rep Ser 1972;498:1-25.

APPENDIX 5: IDENTIFICATION OF TREATMENT ATTRIBUTES AND DEVELOPMENT OF DISCRETE CHOICE EXPERIMENT SURVEY

Chapter 5 was a study to quantify patients' preference toward the latent tuberculosis infection (LTBI) treatment using a discrete choice experiment (DCE) survey. Two sub-studies were conducted. At the first stage, we conducted a qualitative study which was to identify key attributes to describe the LTBI preventive treatment. Primarily based on the findings from this qualitative study, we designed the DCE survey. The content in APPENDIX 5 consists of two parts: (1) the methods and results of the qualitative study conducted to identify treatment attributes and levels; and (2) the steps taken and issues considered during the development of the DCE survey.

PART ONE: Patient Interviews to Identify Treatment Attributes

1. Objectives

Semi-structured individual interviews were conducted among people diagnosed with LTBI. The objectives were (1) to explore patients' attitudes towards LTBI and its treatment; (2) to probe key factors that influence patients' decisions on whether to accept the preventive treatment (i.e. 9 months of isoniazid).

2. Methods

Study participants were recruited from the Vancouver TB clinic at the British Columbia Center for Disease Control (BCCDC). Individuals would be eligible for this study, if they were: (1) at least 19 years old at the time of entry of this study; (2) newly diagnosed with LTBI as per clinical practice; (3) able to communicate in English or Mandarin; (4) willing to sign an informed consent form. Participants were purposefully selected to include both patients who accepted the preventive treatment and those who declined it. Ethics approval (APPENDIX 2) was obtained from the University of British Columbia's Behavioural Research Ethics Board and all participants provided the informed consent.

Although we intended to use a focus group discussion format, [1,2] we eventually decided to conduct individual patient interviews for the following reasons. Due to the cultural background, most of the participants did not feel comfortable to talk about their

infection and express their feelings in front of a group of people. In addition, it was often hard to schedule a time that would accommodate 6-7 people for a group discussion. The ethics application and interview guide were applied to both focus group discussions and individual interviews.

An initial interview guide (attached at the end) was developed based on literature review and opinions from physicians experienced at treating and managing people with TB. The purpose of the guide was to direct the interview discussion and ensure the same questions would be asked to all participants. The interview started with simple, open-ended questions, e.g. "*Why did you get the skin test done? What do you know about tuberculosis?*", and then moved to more probing questions. Each participant was invited to talk about their knowledge and feelings about LTBI and the 9-month isoniazid regimen and their perceived motivating factors and barriers of accepting the regimen. At the end of each interview, the interviewer summarized the points that arose from the discussion, confirmed with the participant and invited further comments and discussion around these points. Participants' basic socio-demographic information was also collected at the end of the interview.

Preliminary data analysis was performed concurrently with the interview progressing. As the interviews went along, emerging themes would inform the focus of the subsequent interview questioning and the interview guide would be modified. We continued to conduct interviews until the information obtained was saturated. To analyze the data,

content analysis were used to identify and extract common themes and group them into logical categories that are relevant to people's preferences for the LTBI treatment.

3. Results

In-depth individual interviews were conducted among 20 people diagnosed with LTBI, among which 10 accepted the preventive treatment and 10 declined it. Their mean age was 41.9 years and 9 were male. Out of the 20 participants, 17 were ethnically Asian or Pacific Islanders, 2 Canadian-born Aboriginal and 1 black. The average year of education was 14.5.

Many intertwined factors influenced people's willingness of accepting the 9-month isoniazid regimen. The commonly mentioned factors that motivated the participants to initiate the preventive treatment included: (1) Concerns about children and family, friends and other people around (N=9 mentioned); (2) Positive impression on the effectiveness of the treatment (N=11); (3) Perceived high risk of progressing to active TB disease if not taking preventive treatment (N=3); (4) Perceived seriousness of active TB disease, e.g. from knowing someone who had active TB disease (N=3); (5) Established trust in health care professionals, e.g. the monthly monitor of liver enzyme (N=5); (6) Work or school obligation, mostly in health care or early education area (N=2); (7) The medications are free of charge (N=2).

On the other hand, most people resent the need to taking medications while they feel healthy and normal. The important perceived barriers of initiating the treatment included: (1) 9 months is too long (N=11); (2) Concerns about treatment associated side effects, especially the liver damage (N=11); (3) Convenience to come for the monthly clinic visit (N=6); (4) Self-perceived treatment compliance problem (N=5); (5) The risk of developing active TB disease from LTBI is considered small (4~5%) (N=3); (6) Negative impression on the effectiveness of the treatment, e.g. it does not offer 100% protection (N=5); (7) Doubt about the diagnosis of LTBI, i.e. the skin test result is confounded by the prior BCG vaccination (N=3); (8) Personal health issues, e.g. already on medications for other underlying conditions (N=2).

PART TWO: Development of Discrete Choice Experiment Survey

1. Selection of Attributes and Levels

Selection of attributes and levels is the most important step in designing a DCE survey [3,4]. The following factors were considered in our study. Firstly, the selected attributes had to be important in describing the LTBI treatment. We chose the attributes primarily based on the results from the in-depth patient interviews, along with TB experts' opinions and literature review findings. Secondly, the attributes and levels should be relevant to reflect the reality that the respondents might expect to experience. Since the treatment of LTBI in Canada is publicly funded and there is no charge to the individual patients, therefore we did not include 'cost of treatment' as an attribute. The attribute levels were

chosen to reflect the current practice or values that were observed in previous studies and also to allow accommodating a wide range of possibility. Thirdly, the attributes and their presentation format were not to introduce unnecessary information bias. The treatment option in the DCE survey was only described using the 6 attributes, without mentioning any name or brand of the medications. Fourthly, respondents' cognitive burden of making trade-offs among multiple attributes was also considered. Following the suggestions in the literature [5], we decided to include a maximum of 6 attributes in our DCE survey: (1) length of treatment; (2) frequency of clinic visit; (3) risk of developing active TB disease after treatment, which is an indicator of the treatment effectiveness; (4) chance of developing liver damage; (5) chance of developing skin rash; (6) chance of developing fatigue.

2. Experimental Design

The next step in the DCE design was to generate the hypothetical LTBI treatment alternatives and combine them to create choice sets [3,4]. Although a full factorial design can cover all possible combinations of the attribute levels, it is too cost-prohibitive and tedious to make respondents complete all possible combined choices. Therefore, a more feasible and applicable design is a fractional factorial experimental one. In addition, orthogonality (minimal correlation between attribute levels), balanced level (each attribute level occurs at an equal frequency), and minimal overlap (each attribute level is not the same across alternatives within a choice set) were taken into account in the experimental design.

3. Number of Choice Sets

We used Sawtooth® to generate 12 versions of DCE questionnaires, with 10 random choice sets in each version. The number of choice sets included in a DCE survey varies from study to study in the literature. A large number of choice sets may increase the complexity of the survey and respondent burden, which is likely to compromise the validity of the survey. In addition, in our study, most of the potential respondents were foreign-born and spoke English as a second language and they often had other work, school or family commitment and were pressed for time. Therefore, we decided to include 10 random choice sets in each version of the survey.

4. 'Neither' Option in Each Choice Set

In each choice set, two hypothetical treatment options were presented to respondents for comparison. In the real world situation, the preventive treatment is optional to individuals with LTBI. Therefore, a 'Neither' option [6] was also available in each choice set if respondents decided not to take both treatment alternatives. The 'Neither' option was explicitly described in each choice set.

5. Consistency-check Question

To check the internal consistency of DCE responses, we included two identical consistency-check choice questions in each version of the survey, located at question 1

and 11. In the fixed choice question, one treatment alternative was clearly 'dominant' or 'better' than the other one [7]. It was assumed that rational respondents would choose the treatment with shorter length, higher effectiveness, and fewer side effects, if they wanted to take a preventive treatment. Respondents were defined '*consistent*' if they chose the 'better' treatment alternative or the 'Neither' option; otherwise they were considered as '*inconsistent*'. On the other hand, the treatment options in random DCE choice sets were manually checked to ensure no 'dominant' scenario was present.

6. Administration of Survey

Survey usually can be administered by mail, telephone or face-to-face and it can be done in paper-pencil-based format or through the internet. Considering the cognitive demands of DCE tasks and the English language level of our respondents, we chose to use paper-based self-administration, where the researcher could explain the questionnaire and the DCE task in detail and the respondent would ask questions if necessary. Each consenting subject practiced a 'warm-up' DCE choice question with the researcher before completing the survey on their own. To make sure the respondents have enough time to work on the survey, they could choose to bring the questionnaire back home to complete it. The 12 versions of DCE questionnaires were randomly administered among the respondents. Socio-demographic and other relevant information was also collected along with the DCE responses.

7. Pre-test of Survey

Prior to final administration, the DCE survey was pre-tested among 60 LTBI patients. Pre-testing was a critical stage in survey development for that it helped to ensure: (1) the background information provided was easy to understand and enough to help respondents make informed and meaningful choices; (2) respondents understood the DCE choice task in a way they were expected to; (3) the attributes and levels were important and relevant; (4) the number of choice sets in each questionnaire was manageable and could be completed within a reasonable amount of time; and (5) the proportion of '*inconsistent*' responses was acceptable. Modifications were made accordingly after the pre-testing.

Patient Interview Guide

This guide was initially developed to be used for focus group discussion. The guide still applies for individual patient interviews.

Part One: Welcome and Thanks

1. Good morning/afternoon, everyone! Thank you for taking time to attend today's discussion. My name is Na, a Ph.D. student at the University of British Columbia. I will be the moderator of today's discussion.

Part Two: Informed Consent Form

Before we start, I would like to go through the informed consent form with you.

1. First page, the 'background' section gives you the information about latent tuberculosis (TB) infection and its treatment and why we are doing this research study.

2. The next paragraph, 'purpose' section talks about the objective of this study, which is to assess people's preferences for latent TB infection treatment. This objective will be achieved by a large survey in the future. To develop the survey, we are doing a pre-study, which is today's focus group discussion. We want to know your perspectives on latent TB infection and its treatment.

3. The 'study procedures' section talks about what you will be expected to do if you decide to participate in this study. All we ask you to do is to volunteer your time and share your opinions during this discussion.

4. The last page is the 'subject information form'. If you decide to participate in this study, please complete this page for record. All the information you provide will be kept strictly confidential and can only be accessed by investigators. Your name will not appear in any reports of the completed study. No information that discloses your identity will be released or published without your specific consent to the disclosure.

4. Finally, I would like to assure you that your participation is totally voluntary. As a thank-you for your time and contribution to our research study, you will each receive $30 at the end of the discussion. So now, feel free to look over the consent form carefully before you sign and do not hesitate to ask me any questions.

Part Three: Introduction to Discussion

1. Thank you for your participation. You are invited to this discussion because you were diagnosed with latent TB infection by doctors. To prevent latent TB infection from progressing to active TB disease, your doctor offered you preventive treatment. Some of

you took the treatment, and some of you did not. We would like you to share with us what you think of the latent TB treatment and why you accepted the treatment or why you did not. I will ask you several questions during the discussion, please feel free to make comments. Keep in mind, we are interested in knowing your opinions, so there are no right or wrong answers. Everyone's inputs will be helpful to us.

2. During the discussion, I will make notes on what we are talking. Also, I will also tape recording the discussion with this digital voice recorder, because I don't want to miss any of your comments. People often say very helpful things and I can't write fast enough to get them all down. But you can be assured of complete confidentiality. After files are transcribed, all the contents will be erased.

3. This discussion will last for maximum one and a half hours. Now it is _____ am/pm, and we will finish by _____ am/pm. Do you have any questions or comments at this point?

Part Four: Discussion

1. Opening

Would you like to tell us about yourself? For example, what is your name? Where are you from? When did you get diagnosed with latent TB infection by a doctor?

2. Example Questions (from general to specific)

(**Note**: These questions act only as a general guide. During the discussion, the moderator may ask other probing questions as necessary to stimulate discussion.)

(1) What do you know about latent tuberculosis infection?

(2) To prevent latent tuberculosis infection from progressing to active tuberculosis disease, people who are diagnosed with latent tuberculosis infection will be offered 6-month of isoniazid in British Columbia.

Did you take the treatment?

(3) What made you take the latent TB treatment?

(4) What stopped or prevented you from taking the treatment?

(5) Once people are infected with TB, one out of 10 would develop active TB disease in their life time. Active TB disease is infectious and can be lethal, and requires more drugs to treat. What do you think of this information?

(6) Clinical trials showed that 6-month of isoniazid treatment would reduce your chance of developing active TB disease by 70%. What do you think of this information?

(7) How much does the length of the treatment matter to you? Would you consider taking the treatment if it is shorter, say, 4 months, 2 months?

(8) How much does it matter that you have to face the adverse side-effects of the treatment?

(9) Does the support from the doctors or nurses matter to you when you decide to take or not take the treatment? Why?

(10) Does the support from your family or friends matter to you? Why?

Part Five: Summarization and Closing

1. The facilitator summarizes major points of the discussion.

2. Did I get everything right? Is there anything else you would like to add? Do you have any suggestions for us to improve our future interviews?

3. Thank you for participating in this discussion.

REFERENCES

1. McLafferty I. Focus group interviews as a data collecting strategy. J Adv Nurs 2004;48:187-194.

2. Krueger RA, Casey MA. Focus groups: a practical guide for applied research. 3rd ed. Thousand Oaks, CA: Sage, 2000.

3. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26:661-677.

4. Hensher D, Rose J, Greene W. Applied choice analysis: a primer. Cambridge: Cambridge University Press, 2005.

5. Lancsar E, Donaldson C. Discrete choice experiments in health economics: distinguishing between the method and its application. European Journal of Health Economics 2005;6:314–316.

6. Ryan M, Skåtun D. Modelling non-demanders in choice experiments. Health Econ 2004;13:397-402.

7. Van der Pol M, Cairns J. Estimating time preferences for health using discrete choice experiments. Soc Sci Med 2001;52:1459-1470.

APPENDIX 6: DISCRETE CHOICE EXPERIMENT QUESTIONNAIRE

This appendix presents the questionnaire used to collect data for the Chapter 5 study and it consists of two parts: a version of discrete choice experiment survey and a socio-demographic & medical history survey. As discussed in Chapter 5, twelve versions of discrete choice experiment surveys (paper-based) were randomly administered to the respondents. Each version has 10 different random DCE choice questions, plus two identical fixed questions (Question 1 and 11) in order to check the validity and consistency of the responses. Only Version 1 is included in this appendix. The other versions share the same format.

Excellence Program
2000 Silver Level Recipient

BC Centre for Disease Control
AN AGENCY OF THE PROVINCIAL HEALTH SERVICES AUTHORITY

Using Discrete Choice Experiment to Evaluate People's Preferences for Treatment of Latent Tuberculosis Infection (LTBI)

Part II – DCE Survey

Background Information

Tuberculosis (TB) is a bacterial infection that is transmitted from person to person primarily through the air. TB usually affects the lungs but it can affect any part of human body such as lymph nodes, kidneys, and bones. Approximately 9 million people get TB disease and 2 million people die of TB each year worldwide. The World Health Organization estimated that 1/3 of the world population is latently infected with TB.

Latent tuberculosis infection (LTBI) means that the person has TB germs in his/her body, but the germs are not active and cannot be spread to other people. It is usually diagnosed through positive tuberculin skin test. People with LTBI alone have no symptoms and feel well. However, when the TB germs become active and grow, it is called "**active TB disease**". This can happen at any time during the life time of people with LTBI. People with active TB disease can get really sick and may die if they don't get proper treatment. People with active lung TB disease may spread the disease to others who have contact with them.

Currently, **preventive treatment** is available to significantly reduce the risk of developing active TB disease from LTBI. The medications can kill the TB germs before they have a chance to become active and make people sick. In Canada, the preventive treatment is publicly funded, so the medications are free of charge for people who are treated. It is the individual's own decision to accept or refuse the treatment.

We would like to know what factors influence your decision whether or not to take the preventive treatment and to determine what characteristics of the treatment are most important to you.

You are invited to complete a discrete choice experiment (DCE) survey as well as a socio-demographic and medical history survey. They should take about 15-20 minutes to complete.

Section I – DCE Survey

The DCE survey will consist of **12 choice questions**. Each question will look like:

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**6 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**1 out of 100 (1%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**5 out of 100 (5%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

In each question, you are given **2 hypothetical TB preventive treatments**: **A** and **B**. They are different in **6 features**:

1. Length of treatment
2. Frequency of clinic visit
3. Risk of developing active TB after treatment (benefit)
4. Chance of developing liver damage (side effect)
5. Chance of developing skin rash (side effect)
6. Chance of developing fatigue (side effect)

Next page will explain what these features mean.
Please read carefully and make yourself familiar with them.

Definition of Treatment Features

Having **latent TB infection** means you have TB germs sleeping in your body. When the germs wake up and grow, **active TB disease** will develop. **Out of 100 people with the infection, 10 people (10%) will develop active TB disease** at anytime during their lifetime.

People with active lung TB disease can spread the disease by coughing, sneezing, or talking. Their close contacts (eg, family members or co-workers) are at the highest risk of getting TB infection. Given that you have TB infection now, the **overall risk of your close contacts getting TB infection from you is 3%,** meaning out of 100 people who have close contacts with you, 3 of them may get TB infection in the future.

The preventive treatment has been shown to be able to significantly reduce your risk of developing into active TB disease, and consequently, to reduce the risk of your close contacts being infected.

1. Length of treatment

It takes months for the medications to kill all the TB germs in your body. This feature refers to the length of TB preventive treatment that is you will be taken. It has 4 levels.

- 4 months of 1 pill daily
- 6 months of 1 pill daily
- 9 months of 1 pill daily
- 12 months of 1 pill daily

For example, '4 months of 1 pill daily' means you have to take 1 pill once a day for 4 months.

2. Frequency of clinic visit

To make sure your treatment is going well and to pick up your next prescription, you will see a doctor at the TB clinic regularly until the treatment is completed.

This feature refers to the frequency of visiting the TB clinic to see a doctor and pick up medications during your treatment. There are 3 different levels.

- Once every 2 weeks
- Once every 1 month
- Once every 2 months

3. Risk of developing active TB after treatment (benefit**)**

Once infected with TB, you have a 10% lifetime risk of developing active TB disease. People with active TB disease can be infectious. Their close contacts (eg, family members or co-workers) are at the highest risk of getting infected. If you complete the preventive treatment, the 10% lifetime risk of developing active TB disease will be reduced, and consequently, the risk your close contacts getting infected will also be reduced.

This feature describes the effectiveness of the treatment, and it literally refers to the risk of you developing active TB disease after you complete preventive treatment. It has 4 levels.

- 4 out of 100 (4%)
- 2 out of 100 (2%)
- 1 out of 100 (1%)
- 0 out of 100 (0%)

For example, '1 out of 100 (1%)' means, out of 100 people who have TB infection and complete the preventive treatment, only 1 person will develop active TB disease. Please note, if you don't take any treatment, your risk of developing active TB disease will be 10%.

4. Chance of developing liver damage (side effect**)**

Like all medications, there may be side effects. A small number of people will develop severe liver damage that may RARELY lead to liver failure and death. Therefore, while on treatment, you will be required to do monthly blood test or more frequently. If evidences of hepatitis appear, your medications can be stopped immediately to avoid serious conditions.

This feature refers to your chance of developing severe liver damage during the treatment. There are 4 different levels.

- 0 out of 100 (0%)
- 1 out of 100 (1%)
- 3 out of 100 (3%)
- 5 out of 100 (5%)
- 10 out of 100 (10%)

For example, '3 out of 100 (3%)' means, out of 100 people on treatment, 3 will develop severe liver damage. The larger the percentage, the greater the risk.

5. Chance of developing skin rash (side effect**)**

Skin rash and/or itching may occur during the preventive treatment. In some people, the symptoms will go away without discontinuing the medication. For severe symptoms, the medications should stop and the symptoms will disappear. For mild reactions, antihistamine treatment will help.

This feature refers to your risk of developing skin rash and/or itching while on preventive treatment.

- 0 out of 100 (0%)
- 5 out of 100 (5%)
- 10 out of 100 (10%)

For example, '5 out of 100 (5%)' means, out of 100 people on treatment, 5 will develop skin rash. The larger the percentage, the greater the risk.

6. Chance of developing fatigue (side effect**)**

Fatigue or drowsiness may occur during the preventive treatment. In some people, the symptoms will go away without discontinuing the medication. In some people, taking the medication during evening time may reduce the impacts.

This feature refers to your risk of developing drowsiness or fatigue while on preventive treatment.

- 0 out of 100 (0%)
- 5 out of 100 (5%)
- 10 out of 100 (10%)

For example, '5 out of 100 (5%)' means, out of 100 people on treatment, 5 will develop drowsiness or fatigue. The larger the percentage, the greater the risk.

Below is an example question (read carefully and understand it):

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**6 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**1 out of 100 (1%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**5 out of 100 (5%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

The survey will start from next page, and you will be given **12 choice questions.**
Please remember that **all the treatment options are not real**,
and there are no right or wrong answers.

Question 1 of 12 (FIX 01)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**6 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 1 month**	**Once every 2 months**	None
Risk of developing active TB after treatment (benefit)	**4 out of 100 (4%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**3 out of 100 (3%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**10 out of 100 (10%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 2 of 12 (RAN 01)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**4 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 1 month**	**Once every 2 weeks**	None
Risk of developing active TB after treatment (benefit)	**4 out of 100 (4%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**3 out of 100 (3%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**10 out of 100 (10%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 3 of 12 (RAN 02)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**6 months** of 1 pill daily	**9 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**0 out of 100 (0%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**0 out of 100 (0%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**10 out of 100 (10%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 4 of 12 (RAN 03)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**9 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 weeks**	**Once every 2 months**	None
Risk of developing active TB after treatment (benefit)	**0 out of 100 (0%)**	**4 out of 100 (4%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**1 out of 100 (1%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**0 out of 100 (0%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 5 of 12 (RAN 04)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**4 months** of 1 pill daily	**6 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 1 month**	**Once every 2 weeks**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**1 out of 100 (1%)**	**3 out of 100 (3%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**10 out of 100 (10%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 6 of 12 (RAN 05)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**9 months** of 1 pill daily	**4 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 2 weeks**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**4 out of 100 (4%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**0 out of 100 (0%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**10 out of 100 (10%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 7 of 12 (RAN 06)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**6 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**1 out of 100 (1%)**	**0 out of 100 (0%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**0 out of 100 (0%)**	**1 out of 100 (1%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**0 out of 100 (0%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**10 out of 100 (10%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 8 of 12 (RAN 07)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**4 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**0 out of 100 (0%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**3 out of 100 (3%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**10 out of 100 (10%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**0 out of 100 (0%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 9 of 12 (RAN 08)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**6 months** of 1 pill daily	**9 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 weeks**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**4 out of 100 (4%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**1 out of 100 (1%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**0 out of 100 (0%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 10 of 12 (RAN 09)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**9 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 2 weeks**	None
Risk of developing active TB after treatment (benefit)	**1 out of 100 (1%)**	**4 out of 100 (4%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**5 out of 100 (5%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 11 of 12 (FIX 02)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**12 months** of 1 pill daily	**6 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 1 month**	**Once every 2 months**	None
Risk of developing active TB after treatment (benefit)	**4 out of 100 (4%)**	**1 out of 100 (1%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**3 out of 100 (3%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**5 out of 100 (5%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**10 out of 100 (10%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Question 12 of 12 (RAN 10)

Treatment **A** and **B** are 2 **hypothetical** preventive treatment. They are different in **6 features**. Compare **A** to **B**, considering the 6 features together. If these 2 treatments were your only treatment options, which one would you choose?

If you don't want choose either A or B, you can also choose "**Neither**" option, meaning you would not have any treatment. Please remember, if you do not get any treatment, you will:

• have a lifetime risk of 10% to develop active TB disease
• not experience any side effects

Treatment Features	**Treatment A**	**Treatment B**	**Neither**
Length of treatment	**6 months** of 1 pill daily	**4 months** of 1 pill daily	No Treatment
Frequency of clinic visit	**Once every 2 months**	**Once every 1 month**	None
Risk of developing active TB after treatment (benefit)	**2 out of 100 (2%)**	**0 out of 100 (0%)**	**10 out of 100 (10%)**
Chance of developing liver damage (side effect)	**3 out of 100 (3%)**	**10 out of 100 (10%)**	0 out of 100 (0%)
Chance of developing skin rash (side effect)	**10 out of 100 (10%)**	**0 out of 100 (0%)**	0 out of 100 (0%)
Chance of developing fatigue (side effect)	**0 out of 100 (0%)**	**5 out of 100 (5%)**	0 out of 100 (0%)
Which would you choose? (**tick only one box**)	Prefer Treatment A ☐	Prefer Treatment B ☐	Prefer No Treatment ☐

Section II – Socio-demographic & Medical History Survey

1. Date of birth: _____ _____ _____
Day Month Year

2. Gender: Male Female

3. Ethnic origin (*check only one*):

Asian or Pacific Islander East Indian

Black Canadian Aboriginal

Caucasian Others, please specify _____

4. Were you born in Canada?

Yes
No → **How long have you been in Canada? _____ years**

5. Current marital status (*check only one*):

Single, never married Married

Separated Divorced

Widowed Others, please specify _____

6. How many people live with you in the same house (except yourself)? _____

7. How many children do you have? _____

8. How many *years of school* have you completed (*circle one number*)?

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
(*primary*) (*high school*) (*college/university*) (*graduate school*)

9. Are you currently employed (or self-employed)?

Yes No

10. Yearly *family income*, including all earners in your household (*check only one*):

$ 0 - $ 19, 999 (CAD)
$ 20,000 - $ 39,999 (CAD)
$ 40,000 - $ 59,999 (CAD)
$ 60,000 - $ 99,999 (CAD)
$ 100,000 or greater (CAD)

11. Do you have any other diagnosed health conditions, such as diabetes, asthma, arthritis, high blood pressure, HIV/AIDS, and cancers, etc?

No Yes, please specify _____

12. Are you currently taking any *Prescription Medications* (besides TB medications)?

No Yes, please specify _____

13. Are you currently taking any *Over-the-counter Medications*?

No Yes, please specify _____

14. Are you currently using any *Herbal or Natural Remedies*?

No Yes, please specify _____

15. Why did you do the TB skin test?

School enrollment Employment

Contact of a TB patient Immigration

Other reasons, please specify _____

16. Do you know the size of your skin test?

No**,** I don't know

Yes, I know → **How big was your skin test? _____ mm (millimeters)**

17. Do you know someone who had active TB disease?

No Yes

18. Have you had BCG vaccine (against TB) before?

No Yes I am not sure

19. Did you decide to take the TB prevention treatment?

Yes. **Why**? Please specify the most important reason _____

No. **Why**? Please specify the most important reason _____

20. Have you started the TB prevention treatment now?

No Yes

→ If yes, have you experienced any side effects from the treatment so far?

No Yes, please specify _____

21. If you can choose FREELY (eg, free of other medical reasons, school or work obligations), would you take the 9-month prevention treatment offered at this clinic?

No Yes

You are now finished the survey. Thank you for your participation!