QUALITY OF LIFE OF CHINESE TYPE 2 DIABETIC PATIENTS

MEASUREMENT OF QUALITY OF LIFE OF PATIENTS WITH TYPE 2 DIABETES IN CHINA

By XUEJING JIN, B.Econ., M.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2017) Hamilton, Ontario
(Health Research Methodology)

TITLE: Measurement of Quality of Life of Patients with Type 2 Diabetes in China

AUTHOR: Xuejing Jin, B.Econ. (China Pharmaceutical University), M.Sc.
(China Pharmaceutical University)

SUPERVISOR: Dr. Feng Xie

NUMBER OF PAGES: xix; 145

ABSTRACT

Background: As one of the major global chronic diseases, diabetes mellitus (DM) places substantial humanistic and economic burdens on patients and their families, healthcare systems and society. Type 2 DM (T2DM) accounts for about 90% of cases of diabetes, and it causes heavier impairment on patients' quality of life (QoL) compared to type 1 DM. China has the world's largest type 2 diabetic patient (T2DP)population. However, little is known about the determinants of long-term QoL of Chinese T2DP. The Chinese Diabetes Quality-of-life (DQOL) measure is one of the most commonly used diabetes-specific QoL measures in Chinese clinical research. Patients have identified the long length of the DQOL a challenge for its implementation. In addition, since it is not a preference-based measure, the Chinese DQOL measured QoL data also cannot be employed in cost-utility analysis, which can achieve a broader economical comparison of interventions across different disease areas in the medical decision-making process. Therefore, this thesis aimed to 1) identify both statistically significant and clinically relevant determinants of long-term QoL of Chinese T2DPs, 2) develop a short version for the Chinese DQOL, and 3) map the Chinese DQOL onto the generic preference-based QoL instrument, EuroOoL5 dimensions 3-level (EQ-5D-3L).

Methods: Data from a longitudinal observational study which recruited clinically diagnosed T2DPs (n=2,886) from community health centers in five Chinese cities was used. The Chinese DQOL and EQ-5D-3L were administered, and demographics, diabetes-related, and other health-related information was collected at baseline and at the end of 12 months. We used anchor-based approach to estimate the minimally important difference (MID) of the Chinese EQ-5D-3L for T2DP. The MID was used to identify the clinically relevant change in QoL over a one-year period for this patient population. Then logistic and ordered logistic regression models were fitted to identify statistically significant factors that explain these clinically relevant changes. Both the classical test theory and item response theory, each combined with exploratory factor analysis, were applied to reduce the number of items of the Chinese DQOL. Using the same data set, we also estimated the mapping algorithm between the Chinese DQOL (and its short version) and the Chinese EQ-5D-3L index by exploring different estimators and model specifications.

Results and conclusions: This thesis estimated the MID of the EQ-5D-3L index was 0.0262 for Chinese T2DP and found that age, gender, education, income, exercise, and glycosylated hemoglobin level were the statistically significant and clinically relevant factors predicting a change in QoL for ChineseT2DP over a one

year period. We provided a 24-item short version for the 46-item Chinese DQOL to reduce the patients' burden in future clinical research. Mapping algorithms using age, gender, and domain scores were also established to map the original Chinese DQOL and the short version onto the EQ-5D-3L index.

ACKNOWLEDGEMENTS

This thesis work would not have been possible without the support and help of many extraordinary people. I would like to express the deepest gratitude and appreciation to them.

First and foremost, I am deeply indebted to my supervisor and mentor, Dr. Feng Xie. I want to thank him for his supervision, encouragement, and patience during my PhD study. The experience he shared helped me to quickly adapt to a new language and learning environment. His insight, generosity, and strictness assisted me to grow into a researcher and young professional. He provided me opportunities to collaborate with many excellent people and to broaden my horizons. I look forward to a long future cooperation with Dr. Feng Xie.

I want to warmly thank Dr. Hertzel Gerstein, my PhD supervisory committee member, for his constant availability and gentle support during the completion of my thesis. I would like to sincerely thank my PhD supervisory committee member and comprehensive examination supervisor, Dr. Mitch Levine, for his guidance and support during my graduate training. I have been blessed to have such two eminent clinicians on my committee. I do not have any medical background, and their advice from the clinician's point of view gave me a lot of help.

I would like to thank Dr. Gordon Liu, who was my supervisor during my research fellowship at Peking University China Center for Health Economics Research (CCHER). Gordon has been supporting my research for almost 6 years. It was he who encouraged me to apply for a doctoral program. The data I used in my thesis was also provided by Gordon and CCHER. I cannot even begin to describe how thankful I am for his generosity and support.

I want to thank all the amazing faculty members and staff of the most wonderful department: Department of Health Research Methods, Evidence and Impact (formerly "Department of Clinical Epidemiology and Biostatistics"). Special thanks to Lorraine Carroll and Kristina Vukelic for your understanding, for answering all those questions I had, and for helping me solve my problems.

I would also like to thank my friends and colleagues from McMaster University, Peking University, and China Pharmaceutical University. Special thanks to Haijing Guan, Hongchao Li, Yankun Sun, Bruno Kovic, Thuva Vanniyasingam, Dena Zeraatkar, Kathleen Steeves, Guowei Li, Yuan Zhang, Gian Paolo Morgano, Lin Jin, and Yaping Chang for the support over the years. Those idea exchanges, wonderful conversations, and “companionship” make my graduate life full of color.

Finally, a million thanks to my dearest parents, little sister and grandmother. Without their love, understanding, and unconditional emotional and financial support, I cannot finish the largest project of my life thus far.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF ABBREVIATIONS

15-D	15-dimension
ADDQOL	Audit of Diabetes - Dependent Quality of Life
BMI	Body mass index
CCHER	China Center for Health Economics Research
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CI	Confidence interval
CLAD	Censored least absolute deviations
Coef .	Coefficient
Cont.	Continued
COPD	Chronic obstructive pulmonary disease
CTT	Classical test theory
CUA	Cost-utility analysis
DCCT	Diabetes Control and Complications Trial
DDS	Diabetes Distress Scale
DM	Diabetes mellitus
DMQLS	Quality of Life Scale for Patients with Type2 Diabetes Mellitus
DQOL	Diabetes Quality-of-life measure
DSQL	Diabetes Specific Quality of Life
EFA	Exploratory factor analysis
EQ-5D	EuroQol 5 dimensions
EQ-5D-3L	EuroQol 5 dimensions 3-level
GLM	Generalized linear models

GRM	Graded response model
HbA1c	Glycated hemoglobin
HS	Health state
IDF	International Diabetes Federation
IIF	Item information function
IRF	Item response function
IRT	Item response theory
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ITR-QOL	Insulin Therapy Related Quality of Life
MAE	Mean absolute error
MID	Minimally important difference
OLS	Ordinary least squares
OR	Odds ratio
PRO	Patient-reported outcome
QALY	Quality-adjusted life-year
QLICD-DM	Quality of Life Instrument for Chronic Diseases-Diabetes Mellitus
QoL	Quality of life
RL	Reference level
RMSE	Root mean square error
RMSEA	Root mean square error of approximation
SD	Standard deviation
SE	Standard error
SF-12	12-Item Short Form Health Survey
SF-12 PCS	12-Item Short Form Health Survey Physical component score
SF-6D	Short-Form Six-Dimension

SRMR	Root mean squared residual
T2DM	Type 2 diabetes mellitus
T2DP	Type 2 diabetic patients
TTO	Time trade-off
VAS	Visual analogue scale
WHO	World Health Origination

DECLARATION OF ACADEMIC ACHIEVEMENT

This thesis is a ‘sandwich’ thesis, which combines three individual projects prepared for publication in peer-reviewed journals. At the time of writing this thesis, all three manuscripts have been submitted for publication. With guidance from my supervisor, Dr. Feng Xie, my contributions to all manuscripts included developing the research question, writing analysis plan, conducting data analysis, and writing the manuscript. In addition, I was served as an interviewer and data coordinator during the data collection period. My co-authors contributed in leading the survey and preparing the manuscripts for publication. The work of this thesis was conducted between July 2015 and March 2017.

Chapter 1

Introduction

1.1 A brief overview of diabetes in the world and in China

Globally, more than 415 million people are living with diabetes mellitus (DM) and annually more than 5 million people die from DM [1]. As one of the global major chronic diseases, DM causes huge clinical, economic, and humanistic burdens on patients and their families, health systems, and societies. In 2015, the estimated cost of diabetes treatments and complication preventions ranged between 673 billion USD and 1,197 billion USD [1]. Numerous studies have reported that, generally, quality of life (QoL) of patients with DM is significantly poorer than that of the non-diabetic population in physical and psychological domains [2-5]. The health utility of diabetic patients without diabetes-related complications ranges from 0.63 to 0.94 [6, 7], and the disutility (utility decrement) for diabetic patients due to diabetes-related complications ranges from 0.014 for minor hypoglycemia to 0.28 for amputation [7].

There are three main types of DM, i.e., type 1 DM, type 2 DM (T2DM), and gestational diabetes [8]. T2DM accounts for more than 90% of all diagnosed cases of diabetes. Compared to type 1 DM, T2DM causes greater impairment on patients' quality of life [9].

As the world's largest population, China also has the world's largest population of patients with DM. In 2015, the estimated diabetic patient population in China was 109.6 million, which made up more than a quarter of all cases of diabetes worldwide [1]. The estimated diabetes-related health expenditure in China was around 51 billion USD in 2015 [1]. The Chinese government has launched a long-term management strategy in community healthcare settings to manage the large numbers of T2DM cases. One of the purposes of this strategy is to improve the QoL of patients with T2DM [9].

1.2 A brief review of diabetes-related quality of life research in China

Diabetes-related QoL research in China has been mainly focused on three areas: 1) translating and validating diabetes-specific quality of life measures developed in non-Chinese settings among the Chinese patients population; 2) measuring the QoL of Chinese diabetic patients using diabetes-specific and generic quality of life (QoL) measures in clinical trials and observational studies; and 3) exploring the factors that affect the QoL of patients with diabetes.

Currently, four commonly used diabetes-specific quality of life measures have been translated into Chinese and validated among the Chinese T2DP, including the Audit of Diabetes - Dependent Quality of Life (ADDQOL, 19 items) [10], Insulin Therapy Related Quality of Life (ITR-QOL, 23 items) [11], Diabetes Distress Scale (DDS, 28 items) [12], and Diabetes Quality of Life (DQOL, 46 items) [13, 14]. On the basis of these diabetes-specific measures, Chinese researchers have developed several measures in the Chinese setting, such as Diabetes Specific Quality of Life (DSQL, 27 items)[15], Quality of Life Instrument for Chronic Diseases-Diabetes Mellitus (QLICD-DM, 48 items) [16], and Quality of Life Scale for Patients with Type2 Diabetes Mellitus (DMQLS, 87 items) [17]. At present, there is no generally preferred diabetes-specific measure in the Chinese research setting, and the selection of a measurement instrument is mainly based on the actual needs of research questions. During application, researchers have identified that the successful implementation of a QoL measures is impacted by its length [18-20]. Thus, in recent years, a trend is reducing the length of the diabetes-specific measures. The 28-item DDS and the 87-item DMQLS have been reduced into 17-item [21] and 39-item measures [22], respectively. The Chinese DQOL has been commonly used in Chinese diabetes-related research, but during its application, patients complained that it was too long; however, there has been no short version of the DQOL available until now.

As patients' quality of life becomes one of the most important outcomes in diabetes-related research and medical decision-making [23], both the diabetes-specific measures mentioned above and the generic preference-based quality of life measures have been increasingly used in Chinese clinical trials and observational studies. Since these above mentioned diabetes-specific measures are all non-preference based, the results of the studies which only used diabetes-specific measure(s) cannot be employed in the cost-utility analysis (CUA), which can achieve a broader economical comparison of interventions across different disease areas in the medical decision-making process [24]. Considering a large proportion of the existing diabetes studies only used diabetes-specific measures [25-28], building a bridge between the diabetes-specific measures, such as the DQOL, and the CUA can help to extract more valuable information from these studies to support medical decision-making.

Exploring the determinants of the QoL of type 2 diabetic patients (T2DP) can provide evidence for improving diabetes management and the goal of improving T2DP's QoL can be eventually arrived at [1]. Currently, numerous cross-sectional studies have identified that QoL of Chinese T2DP is determined by gender, social support, depressive symptoms, diabetic complications, medications, etc [29-31]. However, little is known about the determinants of the long-term QoL of Chinese T2DP.

To solve the above mentioned issues in existing diabetes-related QoL research in the Chinese setting, using data from an observational longitudinal national survey, this thesis aimed to identify both statistically significant and clinically relevant determinants of long-term QoL of Chinese T2DP, develop a short version of the Chinese DQOL, and map the Chinese DQOL onto the generic preference-based QoL measure, Euroqol 5 dimensions 3-level(EQ-5D-3L).

1.3 Overview of the survey and data used in this thesis

Data used in this thesis was from an observational longitudinal survey of T2DP which funded by Guangzhou Zhongyi Pharmaceutical and conducted by Peking University China Center for Health Economics Research between December 2010 and December 2012. The Primary objective of this study was to conduct a real-world economic evaluation of using different types of oral hypoglycemic agents to treat T2DM in the community setting. This was an observational study, and no intervention or any medical advice was provided to participants. This thesis only used the quality of life data and demographics and health-related information. The Economic evaluation is still ongoing.

1.3.1 Sampling and participants

Patients were recruited from 66 community health centres (24 in Beijing, 5 in Chengdu, 11 in Guangzhou, 8 in Nanjing, and 18 in Shenyang) using quota sampling in terms of age and sex according to the Fifth National Population Census [32] between December 2010 and October 2011. A convenience sampling strategy was used in each health centre. The recruitment was facilitated by one endocrinologist (or a general practitioner if the health centre did not have an endocrine department) from each health centre, and was assisted by trained interviewers. All clinically diagnosed type 2 diabetic patients who visited the health care centres due to diabetes in the year 2010 and provided phone number were contacted over the phone. Patients were interviewed at baseline and followed every three months over the one year study period.

The participant inclusion criteria were: 1) aged 16 years or older; 2) clinically diagnosed with T2DM; 3) taking oral hypoglycemic agents; 4) without any cognitive impairment and serious vision and hearing problems; 5) able to read and communicate in Mandarin; and 6) consent to participate in the study. Informed consent was obtained from all patients included in the study.

1.3.2Data collection

Patients were invited to the health centre for face-to-face paper-and-pencil interviews. At the baseline interview and at the last follow-up, patients received a medical examination including blood pressure, height, and weight. A fasting blood sample was collected to test the blood lipids, glycated hemoglobin (HbA1c) level and fasting blood glucose level. Each participant was also asked to complete a long form questionnaire, which consisted of: 1) basic demographic information such as age, gender, employment, marital status, and health insurance; 2) personal health information and health-related behavior, including comorbidities, smoking, drinking, frequency of blood glucose monitoring, eating habits, and physical exercise frequency; 3) medication usage, including both oral hypoglycemic agents and insulin usage; 4) costs of the most recent hospitalization, outpatient visit and purchasing medications from pharmacies, respectively; 5) QoL, which was measured by both the EQ-5D-3L and the Chinese version of Diabetes Quality-of-Life (DQOL); and 6) family economic status. A question about the change in general health status over a one-year period with response options "about the same," "better now," "worse now" and "it is hard to say" was included in the questionnaire. The other three quarterly follow-ups collected the same information but without the medication examination, family economic status, and DQOL.

The EQ-5D-3L measures health status in five dimensions, including mobility, self-care, usual activities, pain/discomfort and anxiety/depression with 3-level response options for "no," "some," and "extreme problems" [33]. The DQOL measure is a 46 item instrument with responses recorded on a 5- point Likert scale. It comprises four subscales: satisfaction, impact, diabetes related worry and social/vocational worry [34] and has been translated into a Chinese version and validated in the Chinese diabetic population [13, 14].

This thesis only used baseline and one-year follow-up data. An English translation of the long form questionnaire (originally in Chinese) is provided in Appendix 1.1.

1.3.3 Training and quality control

A total of 159 interviewers attended a one-day training session which included an introduction of the study, explanations for possible questions, and mock interviews. Throughout the data collection process, every filled questionnaire was checked by two other interviewers independently. A double-entry method was adopted to ensure the accuracy of data entry.

1.4 Scope of chapters

This is a "sandwich" thesis, which combines three individual projects, i.e., Chapters 2-4. As an introduction of the whole thesis, Chapter 1 provided general background and rationale for conducting these three projects. Since data used in the three projects was from the same longitudinal study, Chapter 1 also provided a detailed description of the sampling methods, patients, data collection, interviewer training, and quality control of this longitudinal study.

Chapter 2 estimated the minimally important difference (MID) of the EQ-5D-3L in Chinese T2DP using an anchor-based approach, identified the clinically relevant change in QoL of this patient population by using the MID as the cut-point, and explored the determinants of this change over a one-year study time period using econometric models.

Chapter 3 developed short versions of the commonly used diabetes-specific QoL measure, DQOL, using both the classical test theory and item response theory, each combined with exploratory factor analysis. The short versions were validated using confirmatory factor analysis and criterion validity tests.

Chapter 4 mapped the diabetes-specific measure, DQOL (and its short version developed in Chapter 3), onto the generic preference-based QoL measure, EQ-5D-3L, by exploring different estimators and model specifications.

Chapter 5 summarized the main findings from Chapters 2-4.The three projects' common data-related limitations and implications for further research were also discussed. More detailed explanation regarding methodological and other non-data related issues were provided in the discussion sections of Chapters 2-4.

References

[1] International Diabetes Federation. (2015). IDF Diabetes Atlas Seventh Edition.From http://www.diabetesatlas.org/ (Accessed on Mar 20, 2017).

[2] Eljedi A, Mikolajczyk RT, Kraemer A, Laaser U. (2006). Health-related quality of life in diabetic patients and controls without diabetes in refugee camps in the Gaza strip: a cross-sectional study. BMC Public Health. 6:268-273.

[3] Kiadaliri AA, Najafi B, Mirmalek-Sani M. (2013). Quality of life in people with diabetes: a systematic review of studies in Iran.J Diabetes MetabDisord. 12(1):54-63.

[4] Javanbakht M, Abolhasani F, Mashayekhi A, Baradaran HR, Jahangirinoudeh Y. (2012). Health related quality of life in patients with type 2 diabetes mellitus in Iran: a national survey.PLoS One.7 (8): e44526.

[5] Murillo M, Bel J, Pérez J, Corripio R, Carreras G, Herrero X, Mengibar JM, Rodriguez-Arjona D, Ravens-Sieberer U, Raat H, Rajmil L. (2017). Health-related quality of life (HRQOL) and its associated factors in children with Type 1 Diabetes Mellitus (T1DM).BMC Pediatr.17(1):16-24.

[6] Lee JM, Rhee K, O'grady MJ, Basu A, Winn A, John P, Meltzer DO, Kollman C, Laffel LM, Lawrence JM, Tamborlane WV, Wysocki T, Xing D, Huang ES; JDRF Continuous Glucose Monitoring Study Group.(2011). Health Utilities for Children and Adults with Type 1 Diabetes.Med Care. 2011 October ; 49(10): 924–931.

[7]Beaudet A, Clegg J, Thuresson PO, Lloyd A, McEwan P. (2014). Review of utility values for economic modeling in type 2 diabetes.Value Health.17(4):462-470.

[8]DiabetesInfo.Types of diabetes.From http://www.diabetes-info.co.uk/what-is-diabetes/types-of-diabetes.html (Accessed on Mar 20, 2017).

[9] Chinese Diabetes Society. China's prevention and treatment guideline for type 2 diabetes Mellitus (2013 edition). Chin J Diab Mellitus 2014;6:447–98.

[10] Kong D, Ding Y, Zuo X, Su W, Xiu L, Lin M, Rao S, Yu S. (2011). Adaptation of the Audit of Diabetes-Dependent Quality of Life questionnaire to people with diabetes in China.Diabetes Res ClinPract. 94(1):45-52.

[11] Liu W, Zhao Q. (2009). Research of reliability and validity of Insulin Therapy Related Quality of Life Measure Questionnaire. Chin J Nurs. 44(9): 849-851.

[12] Ting RZW, Nan H, Yu MWM, Kong APS, Ma RCW, Wong RYM, Loo K, So WY, Chow CC, Ko GTC, Wing YK, Chan JCN. (2011). Diabetes-Related Distress and Physical and Psychological Health in Chinese Type 2 Diabetic Patients. Diabetes Care. 34(5): 1094–1096.

[13] Ding Y, Ni Z, Zhang J, Chen G, Feng H. (2000). The Assessment on Reliability and Validity of Adjusted Diabetes Quality of Life (A-DQOL) Scale. Chin J PrevContrChron Non-commun Dis. 8(4): 160-162.

[14] Ding Y, Kong D, Ni Z, Deng H. (2004). Culture adaption and revision of diabetes-specific quality of life scale (QDOL). Chin J Behav Med Sci. 13(1): 102-103.

[15] Fang J. (2000). Quality of Life: Measurement methods and applications. Beijing: Beijing Medical University Press.

[16] Luo N, Li H. (2012). Evaluation on quality of life instrument for patients with diabetes mellitus. Chin J Public Health. 28 (5): 588-560.

[17] Wang LS, Sun ZQ, Cai TS, Zhou ZG. (2005). Development and evaluation of quality of life scale for patients with type 2 diabetes mellitus.Zhong Nan Da XueXueBao Yi Xue Ban. 30(1):21-27.

[18] Psycho-oncology Co-operative Research Group. How do I choose which HRQoL questionnaire to use in my study?From http://www.pocog.org.au/qolfaq.aspx?question=3.(Accessed on Mar 21, 2017).

[19] Schmitt N,Stuits DM. (1985).Factors Defined by NegativelyKeyed Items: The Result of Careless Respondents? ApplPsychol Meas. 9(4):367-373.

[20] Rolstad S, Adler J, Rydén A. (2011). Response burden and questionnaire length: is shorter better? A review and meta-analysis.Value Health. 14(8):1101-1108.

[21] Polonsky WH, Fisher L, Earles J, Dudl RJ, Lees J, Mullan J, Jackson RA. (2005). Assessing psychosocial distress in diabetes: development of the diabetes distress scale.Diabetes Care. 28(3):626-631.

[22]Fu C. (2010). Revision and validation of the Quality of life Scale for

Chinese patients with type 2 diabetes mellitus. Chin JMod Med. 20(1): 60-65.

[23]Diapedia. Patient reported outcomes among patients with diabetes and cancer: an overview and research agenda. From: http://www.diapedia.org/associated-disorders/6104721115/patient-reported-outcomes-among-patients-with-diabetes-and-cancer-an-overview-and-research-agenda. (Accessed on January 25, 2017).

[24] Thorat T, Lin P, Neumann P. (2015). The State of Cost-Utility Analyses in Asia: A Systematic Review. Value Health Reg Issues. 6: 7-13.

[25] Chen A, Su A, Bai H. (2009). Impact factors of quality of life type 2 diabetic patients. Chin Community Doctors. 11(220):38.

[26]Shiu AT, Thompson DR, Wong RY. (2008). Quality of life and its predictors among Hong Kong Chinese patients with diabetes.J ClinNurs. 17(5A):125-132.

[27] Ye S, Chen M, Hu H, He Y, Deng D, Dai F, Jia J, Wang C, Zhang Q, Wang Y. (2012). Clinical follow-up observation for 34 type 2 diabetic patients with six-year intensive insulin therapy. Chin J EndocrinolMetab. 28(04): 293-294.

[28] Borel AL, Pépin JL, Nasse L, Baguet JP, Netter S, Benhamou PY. (2013). Short sleep duration measured by wrist actimetry is associated with deteriorated glycemic control in type 1 diabetes.Diabetes Care. 36(10):2902-2908.

[29]Liu Y, Maier M, Hao Y, Chen Y, Qin Y, Huo R. (2012). Factors related to QoL for patients with type 2 diabetes withor without depressive symptoms–results from a community-basedstudy in China. J ClinNurs. 22: 80-88.

[30] Wang X, Shen B, Zhuang X, Wang X, Weng W. (2017).Investigating Factors Associated with Depressive Symptoms of Chronic Kidney Diseases in China with Type 2 Diabetes.J Diabetes Res. doi: 10.1155/2017/1769897. [Epub ahead of print].

[31]Luk AOY, Zhang Y, Ko GTC, Brown N, Ozaki R, et al. (2014). Health-Related Quality of Life in Chinese Patients with Type 2 Diabetes: An Analysis of the Joint Asia Diabetes Evaluation (JADE) Program. J Diabetes Metab 5(2):333.

[32]National Bureau of Statistics of China. (2000). Fifth National Population Census.From http://www.stats.gov.cn/tjsj/ndsj/renkoupucha/2000pucha/pucha.htm (Accessed

on Dec 1, 2016).

[33] EuroQol Group. How to use EQ-5D.From http://www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html (Accessed on Dec 1, 2016).

[34] The DCCT Research Group. (1988). Reliability and Validity of a Diabetes Quality-of-Life Measure for the Diabetes Control and Complications Trial (DCCT). Diabetes Care. 11(9): 725-732.

Appendix

Appendix 1.1Chinese Diabetic Patients' Health Status Survey Long Form Questionnaire

Form A. Basic Information		
A1	Gender	(1) Male; (2) Female.
A2	Date of birth	YYYY-MM
A3	Chinese Zodiac (Sheng Xiao)	(1) Rat; (2) Ox; (3) Tiger; (4) Rabbit; (5) Dragon; (6) Snake;(7) Horse; (8) Goat; (9) Monkey; (10)Rooster; (11) Dog; (12) Pig.
A4	Ethnic groups	(1) Han; (2) Other please specify___
A5	Marital status	(1)Single; (2) Married and living together; (3)Married but separated; (4) Divorce; (5) Widowed; (6) Other please specify____
A6	Education	(1)Primary or less; (2)Primary School; (3)Junior high school; (4) High school or polytechnic schools; (5)College; (6)Undergraduate; (7)Master or higher.
A7	Location of Hukou	(1) Current residence; (2) Other places.
A8	Category of Hukou	(1) Non-agricultural Hukou; (2) Agricultural Hukou.
A9	Health insurance (Multiple selections allowed)	(1) The basic medical insurance for urban employees; (2) Urban Resident Basic Medical Insurance; (3) New rural cooperative medical care; (4) Free national medical insurance; (5) Commercial health insurance; (6) Student Health Insurance; (7) Other health insurance, please specify_____; (8) Do not have any health insurance.
A10	Employment status	(1) Full-time; (2) Temporary; (3) Part-time; (4) Self-employed or freelancer; (5) Retired; (6) Student [skip to Form B]; (7) Preschool child[skip to Form B]; (8) Farming; (9) Unemployed [skip to Form B]; (10) Other, please specify_______.
A11	Category of your	(1) Administrative organs; (2) Public institutions;

	employer (Retirees please choose the employer category before retired.)	(3) State-owned enterprises; (4) Collective enterprises; (5) Joint-stock company; (6) Private company; (7)Overseas-funded enterprises; (8) Self-employed; (9) Rural; (10) Do not know; (11) Other please specify______.	
A12	Over the past year, how many months were you engaged in the work?		________Months.
A13	How many days on average did you work per week?		________Days / Week.
A14	When engaged in this work, how many hours did you work per day?		________Hours / Day.
A15	Over the past year, how much was your average monthly wage (including various bonuses) or pension?		________CNY / Month.
A16	Over the past year, how much was your average monthly other income in addition to wage (including various bonuses) or pension?		________ CNY / Month.

Form B. Personal Health History and Behavior		
B1	Height	______cm
B2	Weight	______kg
B3	Overall, how you would rate your general health status within the past month?	(1) Excellent; (2) Good; (3) Average; (4) Fair; (5) Poor.
B4	Compared to one year ago, how you would rate your health in general now?	(1) About the same; (2) Better now; (3) Worse now; (4) It is hard to say.
B5	Are You Disabled? (Do you have certification of disability?)	(1) Yes (2) No [Skip to B7]
B6	Degree of your disability	________ ("-8" if don't know)
B7	Which of the following diagnosed chronic disease do you have? (Multiple selections allowed)	(0) None; (1) Heart disease (e.g. myocardial infarction, coronary heart disease, congestive heart failure and other heart diseases); (2) Hypertension; (3) Hyperlipidemia; (4) Stroke or other cerebrovascular diseases; (5) Diabetes; (6) Chronic lung disease (e.g. chronic

		bronchitis or emphysema); (7) Asthma; (8) Arthritis; (9) Osteoporosis; (10)Malignant tumors or cancer; (11) Gastric or duodenal; (12) Parkinson's disease; (13) Cataracts; (14) Hip or femur fracture; (15) Other chronic diseases please specify____.
B8	Do you smoke?	(1) Never [Skip to B11]; (2) Sometimes; (3) Often; (4) Have quit smoking.
B9	How many years have you been smoking (or before you quit smoking)?	_______Year(s)
B10	How many cigarettes do you smoke per day on average (or before you quit smoking)?	_______Cigarettes
B11	Do you drink alcohol?	(1) Never [Skip to B13]; (2) Sometimes; (3) Often; (4) Have quit drinking alcohol [Skip to B13].
B12	How many times did you get drunk within the past month?	_________Times
B13	Do you participate in physical exercise?	(1) Often; (2) Sometimes; (3) Never [Skip to B17]
B14	How many times do you participate in physical exercise per week on average?	_______ Times
B15	How many minutes of physical exercise do you do every time on average?	_______Minutes
B16	What is the most common type of physical exercise that you take?	(1) Walking, running, etc.; (2) Fitness equipment; (3) Dance, aerobics, etc.; (4) Ball Games; (5) Tai Chi (or sword); (6) Other, please specify____________
B17	How many times of physical examination did you take in the past year?	_______Times.
B18	Do you often take the initiative	(1) Yes; (2) No.

	to learn some health-related knowledge?	
B19	How often did you eat food within the past month on average?	(1) More than three meals per day; (2) Three meals per day; (3) Two meals per day; (4) One meal per day; (5)5-6 meals per week; (6)3-4 meals per week; (7) Two meals per week; (8)Other, please specify____________.
B20	Do you control diet due to diabetes?	(1) Often; (2) Sometimes; (3) Never
B21	How often do you test your blood glucose on average?	Every _______ day(s).
B22	Which one of the following options best describes your daily diet?	(1) Meat; (2) Vegetarian-based; (3)Meat and vegetables.
B23	How much water did you drink every day on average in the past month? (In milliliters, a bottle of mineral water is 500 ml.)	_________ml.
B24	How many times did you pee per day on average in the past month?	________Times
B25	How often did you poop on average in the past month?	Every_____day(s).

Form C. Medication history		
C1	When did you be diagnosed with type 2 diabetes?	YYYY-MM
C2	How many diabetic patients in your direct blood relatives (i.e. parents, siblings and children)?	________
C3	How many diabetic patients in your family or relatives and friends except for your direct blood relatives?	________
C4	Did you take insulin therapy in the last six months?	(1) Yes;(2) No [Skip to C12]
C5	The brand name of the insulin.	_________
C6	How many days did you use insulin in the past six	_______days

	months?	
C7	Do you still take insulin therapy now?	(1) Yes [Skip to C9]; (2) No.
C8	What was the main reason that you stop insulin therapy?	(1) Poor efficiency; (2) High price; (3) Adverse effects; (4) Have better options; (5) Other, please specify____.
C9	Which of the following adverse effects happened during the treatment (Multiple selections allowed)?	(1) Hypoglycemia; (2) Gastrointestinal or gastrointestinal discomfort; (3) Skin rash and other allergic reactions; (4) Other adverse reactions please specify____; (5) None.
C10	How many units of insulin did you use in the past six months?	_______IU
C11	Total cost of insulin in the past six months (including the cost covered by health insurance and the cost out-off your own pocket).	________CNY
C12-C19 are questions about your experience of oral hypoglycemic agents.		
C12	What are the brand name and generic name of (one of) the oral hypoglycemic drug(s) you took in the past six months?	Brand name__________; Generic name_________.
C13	When did you start taking this drug?	YYYY-MM
C14	Do you still take this drug now?	(1) Yes [Skip to C17]; (2) No.
C15	When did you stop taking this drug?	YYYY-MM
C16	What was the main reason that you stop this drug?	(1)Poor efficiency; (2) High price; (3) Adverse effects; (4) Have better options; (5) Other, please specify____.
C17	Which of the following adverse effects happened	(1) Hypoglycemia;

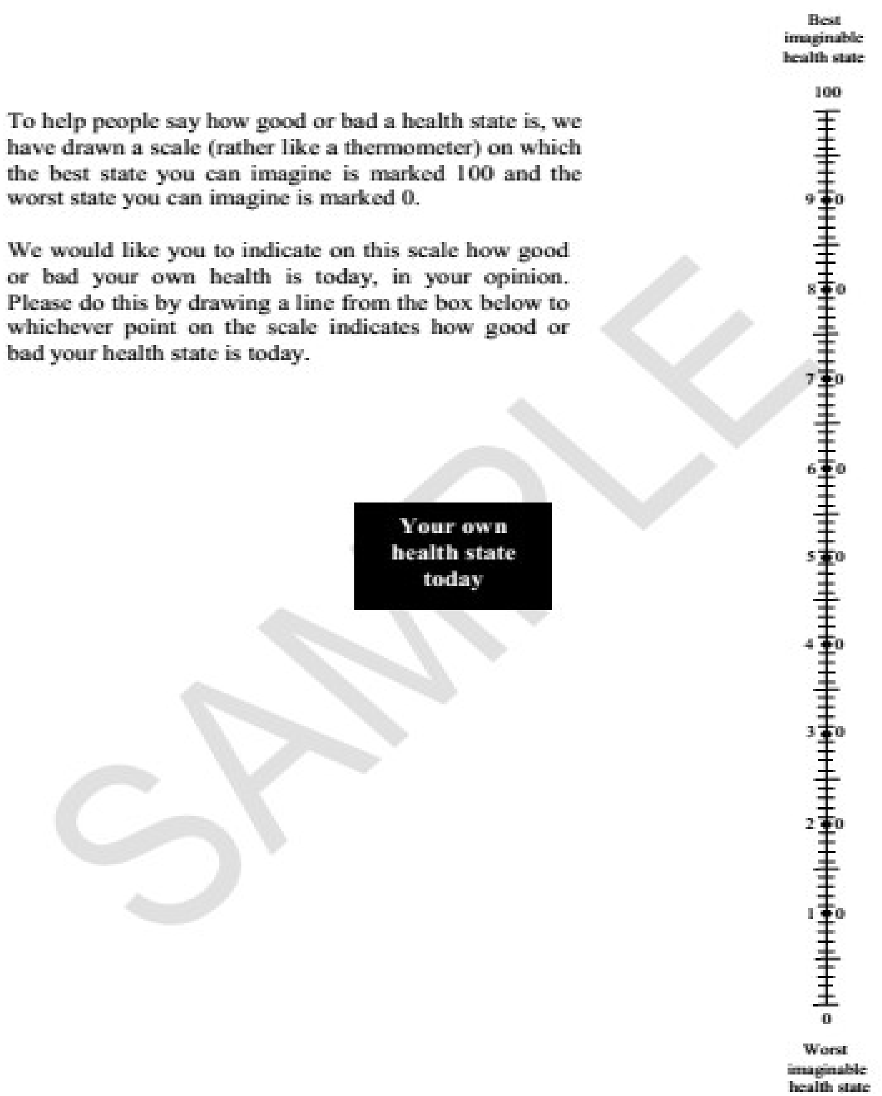
	during the treatment (Multiple selections allowed) ?	(2) Gastrointestinal or gastrointestinal discomfort; (3) Skin rash and other allergic reactions; (4) Other adverse reactions please specify____; (5) None.
C18	How many bottles / boxes of this drug did you use in the past six months?	_______ Bottles / boxes.
C19	Total cost of this drug in the past six months (including the cost covered by health insurance and the cost out-off your own pocket).	________CNY.
Repeat C12-C19 for other oral hypoglycemic agents if you took more than one kind in the past six months.		

Form D. Cost of illness		
About your most recent hospitalization:		
D1	What was the main diagnosis?	_______
D2	What was the name of the hospital?	_______
D3	Hospital admission time	YYYY-MM
D4	Number of days of hospitalization	________days
D5	How many days did you absent from work or school due to this illness?	________days
D6	Total medical cost of this hospitalization (including the cost covered by health insurance and the cost out-off your own pocket).	_______CNY.
D7	Cost of prescriptions (including the cost covered by health insurance and the cost out-off your own pocket).	_______CNY.
D8	Total non-medical cost due to this hospitalization (e.g. accommodation expenses, parking fees).	_______CNY.
D9	How many hospitalizations did you have in the past six months?	_________times
About your most recent outpatient visiting:		
D10	When	YYYY-MM-DD
D11	How many days did you absent from work or school due to this illness?	_____Day

D12	What was the main diagnosis?	__________
D13	What was the name of the medical institution?	__________
D14	Total medical cost of this visiting (including the cost covered by health insurance and the cost out-off your own pocket).	_______CNY.
D15	Cost of prescriptions (including the cost covered by health insurance and the cost out-off your own pocket).	_______CNY.
D16	Total non-medical cost due to this visiting (e.g. accommodation expenses, parking fees).	_______CNY.
D17	How many outpatients visiting did you have in the past month?	_________times
About the most recent experience you purchasing drugs from pharmacies (or online):		
D18	When	YYYY-MM-DD
D19	To treat what kind of disease?	_______
D20	How much did you spent at the drug store this time (including the cost covered by health insurance and the cost out-off your own pocket)?	_______CNY.
D21	Cost of prescriptions (including the cost covered by health insurance and the cost out-off your own pocket).	_______CNY.
D22	How many times did you buy medicine from pharmacies in the past month?	_________times

Form E. Quality of life		
Please indicate your health by select the most appropriate statement in questions E1-E5.		
E1	Mobility	(1) I have no problems in walking about; (2) I have some problems in walking about; (3) I am confined to bed.
E2	Self-Care	(1) I have no problems with self-care; (2) I have some problems washing or dressing myself; (3) I am unable to wash or dress myself.
E3	Usual Activities (e.g. work, study, housework, family or leisure activities)	(1) I have no problems with performing my usual activities; (2) I have some problems with performing my usual activities; (3) I am unable to perform my usual activities.

E4	Pain/Discomfort	(1) I have no pain or discomfort; (2) I have moderate pain or discomfort; (3) I have extreme pain or discomfort.
E5	Anxiety/Depression	(1) I am not anxious or depressed; (2) I am moderately anxious or depressed; (3) I am extremely anxious or depressed.
E6	VAS score	__________



(EQ-VAS image was obtained from: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Products/Sample_UK__English__EQ-5D-3L_Paper_Self_complete_v1.0__ID_23963_.pdf)

E7-E21 Satisfaction

(1 Very satisfied, 2 satisfied, 3 General, 4 Dissatisfied, 5 Very dissatisfied)	
E7	How satisfied are you with the amount of time it takes to manage your diabetes?
E8	How satisfied are you with the amount of time you spend getting checkups?
E9	How satisfied are you with the time it takes to determine your sugar level?
E10	How satisfied are you with your current treatment?
E11	How satisfied are you with the flexibility you have in your diet?
E12	How satisfied are you with the burden your diabetes is placing on your family?
E13	How satisfied are you with your knowledge about your diabetes?
E14	How satisfied are you with your sleep?
E15	How satisfied are you with your social relationships and friendships?
E16	How satisfied are you with your sex life?
E17	How satisfied are you with your work, school, and household activities?
E18	How satisfied are you with the appearance of your body?
E19	How satisfied are you with the time you spend on exercising?
E20	How satisfied are you with your leisure time?
E21	How satisfied are you with life in general?
E22-E41Impact (1 Never, 2 Rarely, 3Occasional, 4 Often, 5 Always)	
E22	How often do you feel pain associated with the treatment for your diabetes?
E23	How often are you embarrassed by having to deal with your diabetes in public?
E24	How often do you have low blood sugar *(e.g. palpitation, sweating, dizziness, trembling)*?
E25	How often do you feel physically ill?
E26	How often does your diabetes interfere with your family life?
E27	How often do you have a bad night's sleep?
E28	How often do you find your diabetes limiting your social relationships and friendships?
E29	How often do you feel good about yourself?
E30	How often do you feel restricted by your diet?
E31	How often does your diabetes interfere with your sex life?
E32	How often does your diabetes keep you from *riding a bike* or being a typist?
E33	How often does your diabetes interfere with your exercising?
E34	How often do you miss household duties because of your diabetes?
E35	How often do you find yourself explaining what it means to have diabetes?
E36	How often do you find that your diabetes interrupts your leisure-time

	activities?
E37	How often do you tell others about your diabetes?
E38	How often are you teased because you have diabetes?
E39	How often do you feel that because of your diabetes you go to the bathroom more than others?
E40	How often do you find that you eat something you shouldn't rather than tell someone that you have diabetes?
E41	How often do you hide from others the fact that you are having an insulin reaction?
E42-E52 Worry (1 Never, 2 Rarely, 3Occasional, 4 Often, 5 Always)	
E42	How often do you worry about your *marriage*?
E43	How often do you worry about your *children's future*?
E44	How often do you worry about whether you will not get a job you want?
E45	How often do you worry about whether you will be denied *pension*?
E46	How often do you worry about whether you will be able to complete your education?
E47	How often do you worry about whether you will lose your job?
E48	How often do you worry about whether you will be able to take a vacation or a trip?
E49	How often do you worry about whether you will pass out?
E50	How often do you worry that your body looks different because you have diabetes?
E51	How often do you worry that you will get complications from your diabetes?
E52	How often do you worry about whether someone will not go out with you because you have diabetes?

Form F. Family economic status		
F1	How many people live in your home?	__________
F2	Among them, how many adults have income?	__________
F3	In the past year, how much is your total household monthly income on average (including government grants).	________CNY/month
F4	Among the total income: (1) Wages (including subsidies, bonuses, pension)	________CNY/month
F5	(2) Business	________CNY/month
F6	(3) Government grants	________CNY/month

F7	(4) Funding from relatives and friends	________CNY/month
F8	(5) property income	________CNY/month
F9	(6) Other	________CNY/month
F10	The house / apartment you are living in is:	(1) Yours; (2) Rented [Skip to F12]; (3) Borrowed [Skip to F12].
F11	Total current value of all houses and apartments you own is:	________CNY
F12	Do you have a car?	(1) Yes; (2) No [Skip to F14].
F13	Total current value of all cars you own is:	________CNY
F14	In the past year, how much is your total household expenditures?	________CNY
F15	Among the total expenditures: (1)Housing costs (e.g. mortgage, rent).	________CNY/year
F16	(2)Healthcare expenditure	________CNY/year
F17	(3)Education expenditure	________CNY/year
F18	(4)Daily living expenses (e.g. food, clothing, daily necessities, water, electricity, gas, and transportation)	________CNY/year
F19	(5) Other (e.g. travel, entertainment)	________CNY/year
F20	Could you tell us how much money your family has?	________CNY
F21	How's the economic burden on your family due to your family members' illness this year?	(1) Very heavy; (2) Heavy; (3) Not heavy; (4)Do not know.
F22	How much is the total debt due to your family members and your illness this year? ("0" for no debt)?	________CNY
F23	Do you have minimum living allowance?	(1)Yes, (2) No.
F24	What is the name of your nearest medical institution?	____________
F25	Is this your medical insurance designated institution?	(1)Yes, (2) No, (3) Do not know, (4) NA (no medical insurance).
F26	This medical institution is:	(1) Public, (2) Private, (3) Do not know.

F27	Does this institution provide Chinese medicine diagnosis and treatment?	(1)Yes, (2) No, (3) Do not know.
F28	Which is your most common form of transportation to this institution?	(1) By car or bus, (2) By bicycle, (3) Walking
F29	How long does it take from your home to this institution by your most transportation?	_________minutes
F30	How many kilometers are from your home to this institution?	_________km

Form G. Interviewer's note		
G1	Could the respondent hear the questions you mentioned clearly?	(1) Could, (2) Basically could, (3) Could not.
G2	Have you checked whether any question was omitted before the respondent leaving?	(1) Yes, (2) No.
G3	Other notes	___________

Chapter 2

Minimally important difference and predictors of change in quality of life as measured using the EQ-5D among Chinese Type 2 Diabetic Patients: A community-based survey in China

Xuejing Jin, MSc [1,2], Gordon G. Liu, PhD [2,3], Hertzel C. Gerstein, M.D., MSc [4],Mitchell A.H. Levine, M.D., MSc [1], Haijing Guan, MSc [2,5], Hongchao Li, MSc [6], Feng Xie, PhD [*1]

1. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
2. China Center for Health Economic Research, Peking University, Beijing, China
3. National School of Development, Peking University, Beijing, China
4. Department of Medicine, McMaster University, Hamilton, Ontario, Canada
5. School of Pharmaceutical Science, Peking University, Beijing, China
6. School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China

* Corresponding author

Abstract

Purpose: To identify the minimally important difference (MID) of the Euroqol 5 dimensions 3-level(EQ-5D-3L) for Chinese type 2 diabetic patients, and identify factors that explain the change in quality of life as measured by the EQ-5D-3L for this patient population over a one year period.

Methods: Clinically diagnosed type 2 diabetic patients were recruited from 66 community health centres in five Chinese cities using a multistage quota sampling method between December 2010 and October 2011. Demographics and socioeconomic status, diabetes duration, comorbidities, treatments (i.e. insulin and oral hypoglycemic medications), and health behaviors were collected via a face-to-face interview at baseline. The EQ-5D-3L was administered at the baseline and at 12months. The China EQ-5D-3L value set was used to calculate the EQ-5D-3L index scores. The change in EQ-5D-3L scores from baseline to the year-end for the patients who answered "worse now" or "better now" to an anchor question "compared to one year ago, how you would rate your health in general now?" were used to calculate the MID. According to their baseline responses to the EQ-5D-3L, patients were categorized into two groups: the impaired HS group (i.e., those reported impaired health states) and the best

possible HS group (i.e., those reported no problems in all five questions of the EQ-5D-3L, which is the best possible EQ-5D-3L health state). For the impaired HS group, the ordered logistic regression with a dependent variable categorizing the change in EQ-5D-3L scores into "worsening," "no change," and "bettering" was performed. For the best possible HS group, the logistic regression with a dependent variable dichotomizing the EQ-5D-3L score change (i.e. "no change" vs. "worsening") was performed. Explanatory variables included age, gender, education, employment, health insurance, per capita disposable income, diabetes duration, comorbidities, treatments, health-related behaviors, and baseline EQ-5D-3L score.

Results: In total 1,958 patients were included in our analysis, with 54.9% female and 67.7% had comorbidities. The mean (standard deviation) age and diabetes duration were 61.2 (11.3) and 7.9 (6.3) years, respectively. The anchor-based MID of the EQ-5D-3L index for Chinese type 2 diabetic patients was 0.0262. For the impaired HS group, age, education, exercise, glycated hemoglobin (HbA1c) level, and baseline EQ-5D-3L scores were significant predictors of the change. For the best possible HS group, age, gender, and income were significant predictors of the change.

Conclusions: Age, gender, education, income, exercise, and HbA1c level were the significant factors predicting a change in quality of life for type 2 diabetic patients over a one year period in China.

Keywords: Quality of life, Type 2 diabetes, minimally important difference, EQ-5D

2.1 Background

Diabetes mellitus (DM) places substantial clinical, humanistic and economic burdens on patients and their families, healthcare systems and society. According to the International Diabetes Federation (IDF) and the World Health Origination (WHO), the global prevalence of DM in adults was 9.1% (415 million people) in 2015. It is expected to rise to 10% (642 million people) by 2040 [1]. The estimated direct annual global cost of DM is more than 827 billion USD [2].In 2010, the prevalence of diabetes among Chinese adults was 11.6% [3]; thus more than 1/4 of all the people with diabetes (109 million people) in the world live in China [1].

Type 2 diabetes mellitus (T2DM) makes up about 90% of cases of diabetes [4]. Studies have reported lower quality of life (QoL) for DM patients, compared to the general population, in terms of both physical functioning and mental health [5-7]. Moreover, people with T2DM generally report lower QoL than those with type 1 diabetes [5, 8].

In order to manage the large numbers of T2DM cases, the Chinese government has launched a long-term management strategy in community healthcare settings [9].Identifying the determinants for long-term QoL of patients with T2DM can help to direct the management of T2DM at the individual level and community level.

Numerous cross-sectional studies have identified that QoL of patients with T2DM is determined by demographic, socioeconomic, health-related behaviors, and diabetes-related factors (e.g., duration, complications, and medication). For example, a Chinese study showed that among Chinese patients with T2DMand depressive symptoms, having more diabetic complications, using hypoglycemic agents or insulin, and having less social support were associated with lower QoL as measured by the Quality of Life Scale for Patients with Type 2 Diabetes Mellitus [10]. However, very few longitudinal studies have explored the longitudinal predictors of QoL of patients with T2DM. A multicountry cohort study found that complications including amputation, stroke, blindness, renal failure, heart failure, and myocardial infarction significantly reduced the EQ-5D-3L index score of patients with T2DM over a five-year period [11]. A German cohort study showed that smoking, having diabetes-related complications and having higher BMI were significantly associated with lower QoL as measured by 12-Item Short Form Health Survey (SF-12) physical component scores over a five-year study period [12]. A Hong Kong study found that over a 2-year observation period, female, unmarried, smoking, no regular exercise, and comorbidities were associated with inferior SF-12 scores [13]. However, the patients included in this study were only recruited from Hong Kong which made the findings not necessarily generalizable to the Chinese type 2 diabetic patient population. Hence, further study with a more representative sample is needed to provide more evidence on the longitudinal predictors of QoL among Chinese patients with T2DM.

In addition, all of these longitudinal studies employed the original QoL scores as continuous outcomes in their regression analyses, and they failed to explore whether those statistically significant factors were clinically relevant [14]. For example, the Hong Kong study used patients' SF-12 physical component score (PCS) as a continuous outcome and found that without regular physical exercise was statistically significantly associated with a decrement of 1.658 points in the SF-12 PCS over the 2-year period. However, they failed to explain whether this decrement in the SF-12 PCS was clinically important for the patients or not. The minimally important difference (MID), which is the smallest change in an outcome that a patient would perceive as important, can be used to solve this problem [15]. A previous study estimated that the MID of EuroQol 5 dimensions (EQ-5D) for the UK diabetic patients was 0.058 [16]; however, the MID of the EuroQol 5 dimensions 3-level (EQ-5D-3L) index in Chinese diabetic patients is unknown. Using the MID derived from the UK population in Chinese setting studies may cause bias since the MID depends on populations' health preferences. Thus, the present study aimed to estimate the MID of the Chinese EQ-5D-3L index for Chinese patients with T2DMand to identify statistically significant and clinically relevant factors that explain the change in QoL for this patient population over a one year period.

2.2 Method

2.2.1 Sample and data

Clinically diagnosed type 2 diabetic patients were recruited from 66 community health centres in five Chinese cities (i.e., Beijing, Chengdu, Guangzhou, Nanjing, and Shenyang) and interviewed face-to-face between December 2010 and October 2011. Patients were followed-up with every three months in a one-year period after the baseline. Demographics and socioeconomic status, diabetes duration, comorbidities, treatments (i.e. insulin and oral hypoglycemic medications), and health behaviors were collected at the baseline. The EQ-5D-3L was administered at the baseline and at 12months. A global question about the change in general health status over the one-year period with response options "about the same," "better now," "worse now," and "it is hard to say" was asked at the last follow-up. A fasting blood sample was collected at the baseline and at 12months to test the blood lipids, glycated hemoglobin (HbA1c) level and fasting blood glucose level. Patients' blood pressure, height, and weight were also examined at the baseline and at 12months. The details of the sampling and data collection were described in Chapter 1.

This chapter focused on changes in QoL as measured by the EQ-5D-3L over one year and was restricted to the baseline and last follow-up data. The observation exclusion criteria were: 1) loss to follow-up; 2) missing data in the EQ-5D-3L questions; 3) age less than the duration of diabetes; 4) HbA1c=0 or HbA1c>20%; and 5) more than one logic errors (a disease recorded at the baseline but not in the last follow-up) in the chronic diseases record (including cardiovascular disease, hypertension, hyperlipideima, stroke or other cerebrovascular diseases, arthritis, osteoporosis, and cataract).

2.2.2 Estimating the minimally important difference of the EQ-5D-3L

The anchor-based approach was employed to estimate the MID of the Chinese EQ-5D-3L for T2DM. The EQ-5D-3L health states were converted to a summary value score using the Chinese EQ-5D-3L value set [17]. The above-mentioned global question "compared to one year ago, how you would rate your health in general now?" together with the four above-mentioned options to it, was used as the anchor. The sign of the change in the EQ-5D-3L score was reversed for those who reported their health became "worse now." Then, the MID was calculated as the mean differences between the baseline EQ-5D-3L scores and the scores at the year-end follow-up among the patients who responded "worse now" (after reversion) and "better now."

2.2.3 Regression analysis

Patients were categorized into two groups according to their baseline responses to the EQ-5D-3L. For those who reported no problems in all questions of the EQ-5D-3L, which is the best possible EQ-5D health state, at the baseline (referred to as the best possible HS group), a logistic regression with a dependent variable dichotomizing the utility change into "no change" and "worsening" was performed. For those who reported impaired health states at the baseline (referred to as the impaired HS group), the ordered logistic regression with a dependent variable categorizing the utility change into "worsening," "no change," and "bettering" was performed. A reduction or an increase in the EQ-5D-3L score by an absolute amount that was larger than the MID over the year was defined as either "worsening" or "bettering," respectively. If the change (absolute amount) was smaller than the MID, the patient was defined as "no change."

Explanatory variables included age, gender, education, employment, health insurance, per capita household income, diabetes duration, comorbidities, baseline insulin usage, baseline oral hypoglycemic medications usage, baseline health-related behaviors (including smoking, drinking, eating habits, exercise, and blood glucose monitoring), baseline body mass index (BMI), and baseline HbA1c level. The baseline EQ-5D-3L scores were controlled in the ordered logistic regression of the impaired HS group. Variables statistically significant in univariate analyses ($P<0.05$) or kept in the backward stepwise regressions (removing criterion: $p\geq0.1$) were included in the final models. The Brant test was employed to check the parallel regression assumption of the ordered logistic model [18]. All statistical analyses were conducted with a two-tailed test at significance level of 0.05 in STATA 12.0 (StataCorp LP, Texas, USA).

2.3 Results

2.3.1 The characteristics of the study participants

A total of 2,886 type 2 diabetic patients were recruited and completed the interview and medical examination at the baseline. Of these, a total of 344 (11.92%) patients were lost to follow-up. As shown in Figure 2.1, among the 2,542 patients who completed the interview and medical examinations, 584 (22.97%) were excluded from the final analysis because of missing data in QoL (n=329), age less than the duration of diabetes (n=11), HbA1c=0 or HbA1c>20% (n=54), or multiple of above-mentioned logic errors (n=295). There was no statistically significant difference in the baseline characteristics of the total patients recruited at the baseline, the patients who completed all rounds of follow-up, and the patients included in the final analysis.

A total of 1,958 patients were included in the final analysis. The mean (standard deviation, SD) age was 61.18 (11.27) years, the mean (SD) duration of diabetes was 7.92 (6.29) years, the mean (SD) HbA1c level was 7.29 (1.88). Out of these patients, 54.9% were female, 64.8% retired, 989 (50.5%) reported no problems in all EQ-5D-3L questions, and 969 (49.5%) reported an impaired health state at the baseline with a mean EQ-5D-3L score of 0.790. The baseline characteristics of patients in the best possible HS group significantly differed from those of patients in the impaired HS group in all variables except for health insurance, using insulin, physical exercise, glucose monitoring, BMI, and HbA1c level (Table 2.1). Compared to the best possible HS group, the impaired HS group had more patients who were older, female, retired/unemployed, less educated, and had a lower income and a longer duration of diabetes. The impaired HS group also had a lower proportion of patients who were married and living together with a partner, drank or smoked, and a higher proportion of patients had comorbidities and a plant-based diet (Table 2.1).

2.3.2 MID and Change in quality of life

Table 2.2 shows the mean change in EQ-5D-3L scores from the baseline to the one-year end for patients with different responses to the anchor question. Very few patients (1.7%) responded “hard to say,” and their responses were incorporated into the category “about the same.” The anchor-based MID for the EQ-5D-3L index in the Chinese type 2 diabetic patients was 0.0262.

Figure 2.2shows the change in EQ-5D-3L scores for each group. Of 989 patients who reported no problem in all five questions of the EQ-5D-3L at the baseline, 283 (28.6%) reported a worse health status (absolute change > MID) at the year end with the mean change of -0.187 (0.115). Of 969 patients who reported impaired health states at the baseline, 652 (67.3%) reported a better health state (absolute change > MID) at the year-end with a mean change of 0.172 (0.088), and 140 (14.5%) reported a worse health state (absolute change > MID) with a mean change of -0.152 (0.095).

2.3.3 Regression analysis

Figures 2.3 and 2.4 show the logistic regression and ordered logistic regression results for the best possible HS group and the impaired HS group, respectively. For the best possible HS group, the QoL (as measured by the EQ-5D-3L) of patients who were older, female, had chronic diseases and had the lowest level income (per capita disposable<10672.0 CNY/year) was more likely to become worse in a one year period than that of those who were younger, male, without chronic disease, and had the highest level income (per capita disposable>35579.2 CNY/year) (after adjusting for city, education, employment, drinking, smoking and number of types of oral hypoglycemic agents). For the impaired HS group, the QoL of patients who were younger, had a college or higher education, who exercised more than 3 hours per week, and who had a lower HbA1c level was more likely to become better in a one year period than that of those who were older, who had high school or lower education, who did not exercise, and who had a higher HbA1c level (after adjusting for the baseline EQ-5D-3L scores, city, education, employment, drinking, smoking and number of types of oral hypoglycemic agents). See Appendixes 2.1-2.4 for the details of the univariate analysis and backward stepwise regression results for each group.

2.4 Discussion

This study aimed to identify the MID of the Chinese EQ-5D-3L for type 2 diabetes, and identify factors that explain the change in QoL for this patient population in a one-year period. The MID was identified as 0.0262 by using the anchor-based approach. We found that, regardless of the baseline responses to the EQ-5D-3L, older patients had a higher probability of reduction (decrease in the EQ-5D-3L index>0.0262) in QoL in the one year period. There were significant correlations between the reduction in QoL in the one year period and age, female, chronic disease and low-income, respectively, for those who reported no problems in all five EQ-5D questions at the baseline. At the same time, levels of education and exercise time were positively associated with an improvement (increase in the EQ-5D-3L index>0.0262) in QoL in one year, respectively, for those who reported an impaired EQ-5D health state at the baseline.

The MID of the EQ-5D-3L for diabetes estimated in our study was lower than that estimated for diabetes in existing study. Mulhern and Meadows reported the anchor-based MID of EQ-5D-3L index for diabetes was 0.058 for UK patients

[16]. Our result was also in the lower end among the MIDs of EQ-5D estimated using anchor-based approaches for chronic disease in general in other countries. Walters and Brazier [19] estimated the MID for chronic diseases ranged from -0.011 for chronic obstructive pulmonary disease (COPD) to 0.139 for leg ulcer among UK patients. Tsiplova et al. estimated the mean MID for chronic conditions (hypertension, heart disease, arthritis, asthma or COPD, cancer, diabetes, chronic back pain, and anxiety or depression) to be 0.044 for Canadian patients [20]. The differences in anchors may lead to the between-study variations. Answer options to anchor questions can vary from simple, such as the "worse now," "about the same," and "better now" employed in our study, to complex, for example, Jaeschke et al. [21] used a 7-point Likert scale to distinguish the degree of worsening and bettering in symptom of shortness of breath for patients with chronic heart and lung disease. In addition to the patient-reported global anchor question, change in other patient-reported [22, 23] and clinician-rated [23] outcome measures also can be used as anchors to calculate the MID.

The determinates of change in QoL for patients who reported no problem in all questions of EQ-5D-3L at the baseline were in line with the that of the general community-dwelling adult and elderly population in China. An interesting point was that females with T2DM were more likely to report worse QoL over one year than male patients with T2DM. The negative association between QoL and female gender was also found in Wan et al's [13] longitudinal study and Luk et al's [24] cross-sectional study among Hong Kong diabetic patients, and Liu et al's [10] cross-sectional study among diabetic patients in Beijing. Similar results were found in patient populations with other chronic diseases in China as well [25]. The reason why women were associated with poorer QoL in Chinese adult population remains unknown. Marital status was not found to significantly affect QoL in patients with T2DM in our study. However, Wan et al's [13] study showed that married patients were more likely to have better QoL over 2 years than unmarried patients.

For the impaired HS group, apart from the demographic factors, more physical exercises were associated with the bettering QoL. This finding is similar to the previous longitudinal study in Hong Kong [13].

Drinking of alcohol was not found associated with poorer QoL either. This is in line with previous studies of diabetic patients in Hong Kong [13] and Germany [12]. Smoking was also not identified as a significant predictor of change in QoL in our study after adjusting for other factors. This is different from previous longitudinal studies of diabetic patients in Hong Kong [13] and Germany [12] and cross-sectional studies in Chinese general population [26], which all found that smoking was associated with poorer QoL. One possible reason is that the proportion of smokers in our sample (19.31%) was lower than that of Chinese general adult population (27.7%) [27]. Patients might tend to conceal the fact that they smoked in the face-to-face interviews.

There was no significant association between the change in QoL and diabetes-related factors (i.e. duration of T2DM, insulin and oral hypoglycemic agent usage, HbA1c level, and blood glucose monitoring frequency) among diabetic patients in our study. Negative associations between QoL with disease duration and HbA1c level were found in Singapore patients [28]. However, our findings were similar to those found in the longitudinal and cross-sectional studies among Hong Kong diabetic patients [13, 24, 29].

Our study used the MID as a cut-point to categorize the change in QoL, which allowed us to identify statistically significant and clinically relevant determinants of change in QoL of Chinese patients with T2DM. However, our study has several limitations. First, the anchor employed in our study was relatively simple. Subdividing the response of “worse now” and “better now” into more detailed categories which describe different degrees of the change (for example, "much better" and "slightly better", "somewhat worse," and "much worse," respectively) can help to more accurately capture the smallest change in the patient-reported outcome that patients perceive as important [30]. Second, our patient population was a relatively healthy diabetic patient population. Patients were asked to go to health care centres for interviews, and the majority of them did not have health problems which seriously limited their daily activities and mobility. We did not recruit patients who have serious complications. So our finding can only be generalized to the community managed diabetic patients with no serious complications. In addition, we had decided not to collect information on diabetes-related complications to shorten the time required for an interview. However, previous studies have shown that diabetes-related complications can affect patients' QoL [10, 31]. Finally, the one-year follow-up was relatively short for a chronic disease to observe enough change in QoL and changeable factors (e.g., income, employment status, usage of insulin and hypoglycemic agents, health-related behaviors, etc.). In addition, we only collected detailed information at the baseline and one-year end. Two data points for each patient did not allow us to use time series methods to explore the effects of changeable factors on patients' QoL over time.

Reference

[1] International Diabetes Federation. (2015). IDF Diabetes Atlas, 7 ed. Brussels, Belgium: International Diabetes Federation.

[2] World Health Organization.(2016). Global Report on Diabetes. Geneva, Switzerland: World Health Organization.

[3] Xu Y, Wang L, He J, et al. (2013). Prevalence and control of diabetes in Chinese adults.Journal of the American Medical Association. 310(9):948-959.

[4]Chinese Diabetes Society.(2012). China guideline for type 2 diabetes (2010).Chin JDiabetes. 20(1): S1–S36.

[5] Rubin RR, Peyrot M.(1999).QoL and diabetes.Diabetes Metab Res Rev. 15(3):205-18.

[6] Kiadaliri AA, NajafiB, Mirmalek-Sani M. (2013).QoL in people with diabetes: a systematic review of studies in Iran. Journal of Diabetes & Metabolic Disorders.;12(1):54-63.

[7] SafitaN, Islam SM, Chow CK6, Niessen L, Lechner A, Holle R, Laxy M. (2016). The impact of type 2 diabetes on health related QoL in Bangladesh: results from a matched study comparing treated cases with non-diabetic controls.Health Qual Life Outcomes.14(1):129-137.

[8]Hariprasad SM, Mieler WF, Grassi M, Green JL, Jager RD, Miller L. (2008). Vision-related QoL in patients with diabetic macular oedema.Br JOphthalmol 2008;92:89-92.

[9] Chinese Diabetes Society. China's prevention and treatment guideline for type 2 diabetes Mellitus (2013 edition). Chin J Diab Mellitus 2014;6:447–98.

[10] Liu Y, Maier M, Hao Y, Chen Y, Qin Y, Huo R. (2012). Factors related to QoL for patients with type 2 diabetes withor without depressive symptoms–results from a community-basedstudy in China. J ClinNurs. 22: 80-88.

[11]Hayes A,Arima H,Woodward M, Chalmers J,Poulter N, Hamet P,Clarke P. (2016). Changes in QoL Associated with Complications ofDiabetes: Results from the ADVANCE Study. Value Health.19:36–41.

[12] Maatouk I, Wild B, Wesche D, Herzog W, Raum E, Muller H,Rothenbacher D, Stegmaier C, Schellberg D, Brenner H. (2012). Temporal Predictors of Health-Related QoL inElderly People with Diabetes: Results of a German

CohortStudy. PLoS ONE 7(1): e31088.

[13] Wan EY, Fung CS, Choi EP, Wong CK, Chan AK, Chan KH, Lam CL. (2016). Main predictors in health-related QoL in Chinese patients with type 2 diabetes mellitus.Qual Life Res. 25(11):2957-2965.

[14] Sedgwick P. (2014). Clinical significance versus statistical significance.BMJ. 348:g2130.

[15] Jaeschke R, Guyatt G H, Cook D, et al. (1994). Effect of increasing doses of beta agonists on spirometric parameters, exercise capacity, and QoL in patients with chronic airflow limitation.Thorax.49(5):479–484.

[16]Mulhern B, Meadows K. (2013). Investigatingthe minimally important difference of the Diabetes Health Profile (DHP-18) and the EQ-5D and SF-6D in a UK diabetes mellitus population.Health. 5(6):1045-1054.

[17] Liu GG, Wu H, Li M, Gao C, Luo N. (2014). Chinese Time Trade-Off Values for EQ-5D Health States.Value Health, 17(5):597-604.

[18] UCLA: Statistical Consulting Group. Stata Data Analysis Examples Ordered Logistic Regression. from http://www.ats.ucla.edu/stat/stata/dae/ologit.htm. (Accessed on November 18, 2015).

[19]Walters SJ, Brazier JE. (2005). Comparison of the minimally important difference for two health state utilitymeasures: EQ-5D and SF-6D. Qual Life Res. 14: 1523–1532.

[20]Tsiplova K, Pullenayegum E,Cooke T, Xie F. (2016). EQ-5D-derived health utilities and minimally importantdifferences for chronic health conditions: 2011 CommonwealthFund Survey of Sicker Adults in Canada. Qual Life Res. 25:3009–3016.

[21] Jaeschke R, Singer J, Guyatt GH.(1989). Measurement of health status.Ascertaining the minimal clinically important difference.Control Clin Trials.10(4):407-415.

[22] Pickard AS, Neary MP, Cella D. (2007). Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer.Health Qual Life Outcomes. 2007 Dec 21;5:70.

[23] McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. (2011). Interpreting patient-reported outcome results: US FDA guidance and emerging methods.Expert Rev Pharmacoecon Outcomes Res. 11(2):163-169.

[24] Luk AOY, Zhang Y, Ko GTC, Brown N, Ozaki R, et al. (2014) Health-

Related QoL in Chinese Patients with Type 2 Diabetes: An Analysis of the Joint Asia Diabetes Evaluation (JADE) Program. J Diabetes Metab 5(2):333-339.

[25] Zhang Y, Zhou Z, Gao J, Wang D, Zhang Q, Zhou Z, Su M, Li D. (2016). Health-related QoL and its influencing factors for patients with hypertension: evidence from the urban and rural areas of Shaanxi Province, China.BMC Health Serv Res. 18(16):277-285.

[26]Sun W, Aodeng S, Tanimoto Y, Watanabe M, Han J, Wang B, Yu L, Kono K. (2015). QoL (QOL) of the community-dwelling elderly and associated factors: a population-based study in urban areas of China.Arch GerontolGeriatr.60(2):311-316.

[27] Chinese center for disease control and prevention. 2015 Chinese adult tobacco survey report.From http://www.chinacdc.cn/zxdt/201512/t20151228_123960.htm (Accessed on November 25, 2016).

[28]ShimYT, Lee J, Toh MPHS, Tang WE, and Ko Y. (2012). Health-related QoL and glycaemic controlin patients with Type 2 diabetes mellitus in Singapore. Diabet.Med. 29, e241–e248.

[29] Shiu AT, Thompson DR, Wong RY.(2008). QoL and its predictors among Hong Kong Chinese patients with diabetes.J ClinNurs.17(5A):125-132.

[30] Johnston BC, Ebrahim S, Carrasco-Labra A, Furukawa TA, Patrick DL, Crawford MW, Hemmelgarn BR, Schunemann HJ, Guyatt GH, Nesrallah G. (2015). Minimally important difference estimates and methods: a protocol.BMJ Open. 5(10):e007953.

[31] Al Hayek AA, Robert AA, Al Saeed A, AlzaidAA, Al Sabaan FS. (2014). FactorsAssociated with Health-Related Quality of Life among Saudi Patients with Type 2Diabetes Mellitus: A Cross-Sectional Survey. Diabetes Metab J. 38:220-229.

Tables

Table 2.1Baseline characteristics of patients

Baseline characteristics	Recruited at baseline (N=2,886)	Completed all follow-ups (N=2,542)	Included in final analysis			p-value[e]
			Total (N=1,958)	Best possible HS group (N=989)	Impaired HS group (N=969)	
Age mean (SD), years	61.15 (11.42)	61.19 (11.29)	61.18 (11.27)	60.13 (11.60)	62.24 (10.83)	<0.001
Female	1,607 (55.68)	1,411 (55.51)	1,074 (54.85)	480 (48.53)	594 (61.30)	0.037
City						<0.001
Beijing	543 (18.81)	477 (18.76)	380 (19.41)	184 (18.60)	196 (20.23)	
Chengdu	657 (22.77)	586 (23.05)	427 (21.81)	268 (27.10)	159 (16.41)	
Guangzhou	558 (19.33)	482 (18.96)	362 (18.49)	188 (19.01)	174 (17.96)	
Nanjing	528 (18.30)	436 (17.15)	340 (17.36)	138 (13.95)	202 (20.85)	
Shenyang	600 (20.79)	561 (22.07)	449 (22.93)	211 (21.33)	238 (24.56)	
Education						0.016
Primary school and lower	780 (27.03)	654 (25.73)	472 (24.11)	216 (21.84)	256 (26.42)	
High school	1719 (59.56)	1,538 (60.50)	1,202 (61.39)	613 (61.98)	589 (60.78)	
College or higher education	387 (13.41)	350 (13.77)	284 (14.50)	160 (16.18)	124 (12.80)	
Employment						<0.001
Full time employee	432 (14.97)	375 (14.75)	309 (15.78)	198 (20.02)	111 (11.46)	
Retired	1,850 (64.10)	1,657 (65.18)	1,269 (64.81)	603 (60.97)	666 (68.73)	
Other [a]	193 (6.69)	168 (6.61)	133 (6.79)	78 (7.89)	55 (5.68)	
Unemployed	411 (14.24)	342 (13.45)	247 (12.61)	110 (11.12)	137 (14.14)	
Health insurance						0.595
Urban employee health insurance	1,847 (64.00)	1,648 (64.83)	1,285 (65.63)	650 (65.72)	635 (65.53)	
Urban residence and new rural insurance	779 (26.99)	665 (26.16)	508 (25.94)	250 (25.28)	285 (26.63)	
Other health insurance [b]	142 (4.92)	130 (5.11)	96 (4.90)	49 (4.95)	47 (4.85)	

No health insurance	118 (4.09)	99 (3.89)	69 (3.52)	40 (4.04)	29 (2.99)	
Comorbidities (chronic disease)						
Cardiovascular disease	623 (21.59)	558 (21.99)	368 (18.79)	144 (14.56)	224 (23.12)	<0.001
Hypertension	1,430 (49.55)	1,257 (49.55)	893 (45.61)	405 (40.95)	488 (50.36)	<0.001
Hyperlipidemia	586 (20.30)	525 (20.69)	293 (14.96)	131 (13.25)	162 (16.72)	0.031
Stroke or other cerebrovascular disease	214 (7.42)	195 (7.69)	108 (5.52)	35 (3.54)	73 (7.53)	<0.001
Other chronic disease [c]	855 (29.62)	776 (30.53)	486 (24.82)	179 (18.10)	307 (31.68)	<0.001
Insulin in last 6 months	467 (16.18)	414 (16.32)	318 (16.24)	158 (15.98)	160 (16.51)	0.748
Exercise time mean (SD), hours per week	5.27 (6.05)	5.30 (6.02)	5.39 (6.04)	5.65 (5.12)	5.12 (5.92)	0.020
Exercise						0.062
0 hour	624 (21.62)	535 (21.05)	398 (20.33)	178 (18.00)	220 (22.70)	
less than 3 hours per week	637 (22.07)	564 (22.19)	425 (21.71)	221 (22.35)	204 (21.05)	
3 to 6 hours per week	582 (20.17)	513 (20.18)	401 (20.48)	202 (20.42)	199 (20.54)	
more than 6 hours per week	1,043 (36.14)	930 (36.59)	734 (37.49)	388 (39.23)	346 (35.71)	
Glucose monitoring mean (SD), days per test	45.30 (82.74)	44.66 (83.96)	45.8 (86.9)	44.15 (71.38)	47.39 (100.33)	0.394
Glucose monitoring						0.302
<=every 3 days	224 (7.76)	204 (8.19)	149 (7.61)	85 (8.59)	64 (6.60)	
every 3-7(=) days	547 (18.95)	479 (19.24)	382 (19.51)	199 (20.12)	183 (18.89)	
every 7 to 30(=) days	1,360 (47.12)	1,215 (48.80)	949 (48.47)	470 (47.52)	479 (49.43)	
> every 30 days	755 (26.16)	592 (23.78)	478 (24.41)	235 (23.76)	243 (25.08)	
Eating habits						0.003
Meat	161 (5.58)	136 (5.35)	98 (5.01)	51 (5.16)	47 (4.85)	
Vegetable	880 (30.49)	770 (30.29)	583 (29.78)	260 (26.29)	323 (33.33)	
Mixed	1,845 (63.93)	1,636 (64.36)	1,277 (65.22)	678 (68.55)	599 (61.82)	
Married and living together	2,432 (84.27)	2,150 (84.58)	1,683 (85.96)	876 (88.57)	807 (83.28)	0.001
Smoke currently	570 (19.75)	497 (19.55)	378 (19.31)	221 (22.35)	157 (16.20)	0.001
Drink currently	724 (25.09)	638 (25.10)	489 (24.97)	267 (27.00)	222 (22.91)	0.037
Per capita disposable income [d]						0.001
Lowest income (<10672.0 CNY/yr)	711 (24.64)	574 (23.47)	456 (23.29)	198 (20.02)	258 (26.63)	
Lower-income to middle level	1,016 (35.20)	882(36.06)	691 (35.29)	352 (35.59)	339 (34.98)	

(10672.0 - 19544.9 CNY/yr)						
Middle level to higher-income (19544.9 - 35579.2CNY/yr)	879 (30.46)	769 (31.44)	620 (31.66)	326 (32.96)	294 (30.34)	
Highest income (>35579.2CNY/yr)	280 (9.70)	221 (9.04)	191 (9.75)	113 (11.43)	78 (8.05)	
BMI mean (SD), kg/m^2	24.93 (3.57)	24.88 (3.50)	24.76 (3.41)	24.66 (3.21)	24.86 (3.61)	0.516
BMI Class						0.482
Normal or thinner	866 (30.00)	769 (30.25)	616 (31.46)	306 (30.94)	310 (31.99)	
Overweight	697 (24.15)	631 (24.82)	484 (24.72)	256 (25.88)	228 (23.53)	
Obesity	1,323 (45.84)	1,142 (44.93)	858 (43.82)	427 (43.17)	431 (44.48)	
Diabetes duration mean (SD), years	7.94 (6.28)	8.01 (6.33)	7.92 (6.29)	7.69 (6.43)	8.16 (6.13)	0.021
HbA1c mean (SD), %	7.35 (2.77)	7.34 (2.87)	7.29 (1.88)	7.28 (1.88)	7.30 (1.87)	0.639

The values presented are numbers (percentage) unless otherwise stated.
a. Includes individual freelancer, hourly worker, student, farmer and other;
b. Includes commercial medical insurance, student medical insurance, and other health insurance;
c. Includes chronic lung diseases, asthma, cancer, gastric ulcer, duodenal ulcer, Parkinson's disease, hip or femur fracture, and other chronic diseases;
d. Per capita disposable income is grouped into the four categories according to the China 2011 per capita disposable income;
e. Chi-square test for frequency, and one-way ANOVA for mean in subgroups (i.e., best possible HS and impaired HS groups).
HS Health state; *SD* Standard deviation; *CNY/yr* Chinese Yuan/year; *BMI* Body Mass Index; *HbA1c*Glycated hemoglobin.

Table 2.2 Mean change in EQ-5D-3L scores for patients with different responses to the anchor question

	Worse now	About the same [a]	Better now	MID calculation [b]	Overall
N (%)	514 (26.25)	1,034 (52.81)	410 (20.94)	924 (47.19)	1,958 (100.00)
Mean change in EQ-5D-3L scores	-0.008	0.021	0.049	0.026 [c]	0.019
First and third quartiles of the change	-0.125, 0.125	0, 0.125	0, 0.131	N/A	0, 0.125

a. A total of 33 (1.69%) patients who responded "hard to say", their responses were incorporated into the category "about the same;"

b. Includes patients who reported "worse now" and "better now", and the sign of the change in the EQ-5D-3L score was reversed for patients who reported "worse now;"

c. Minimally important difference of EQ-5D-3L index scores for the Chinese type 2 diabetic patients.

MID Minimally important difference; *N/A* Not applicable.

Figures

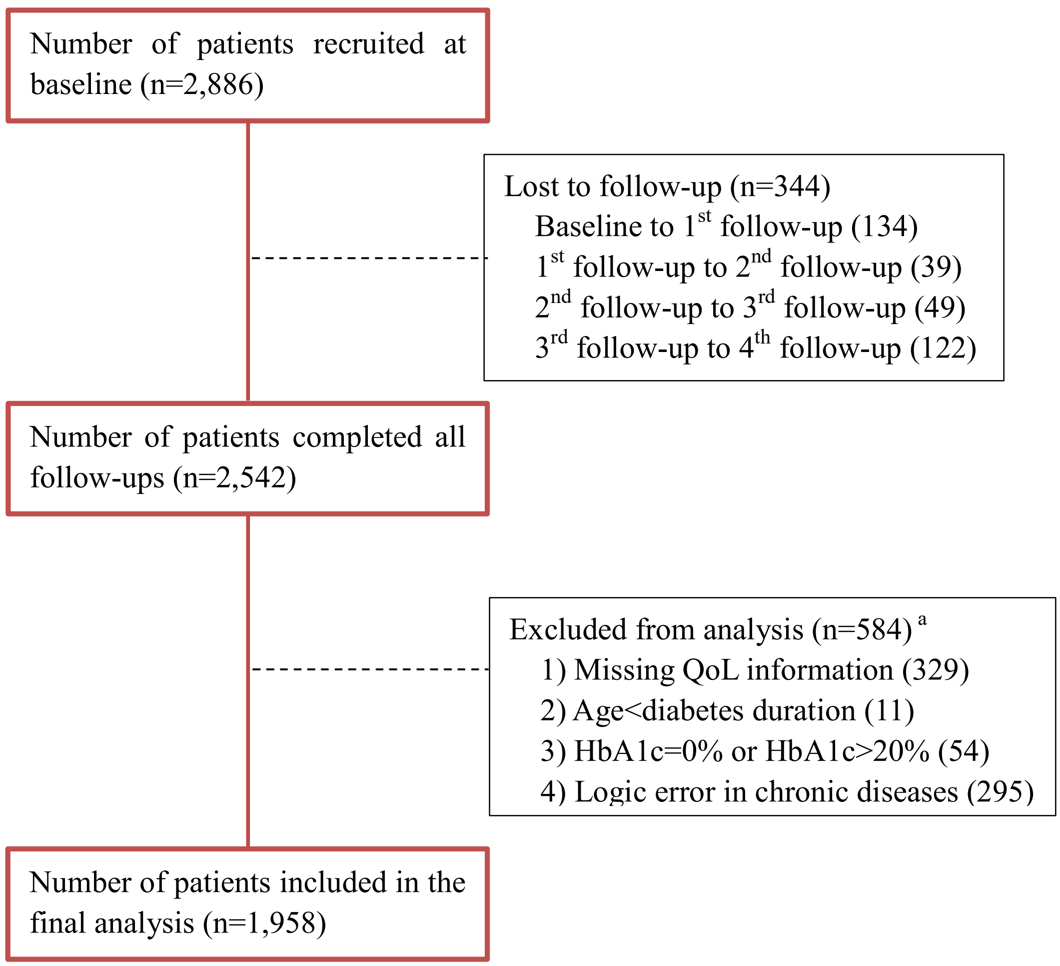


Figure 2.1 Flow diagram

a. The four categories are not mutually exclusive.

QoL Quality of life; *HbA1c* Glycated hemoglobin.

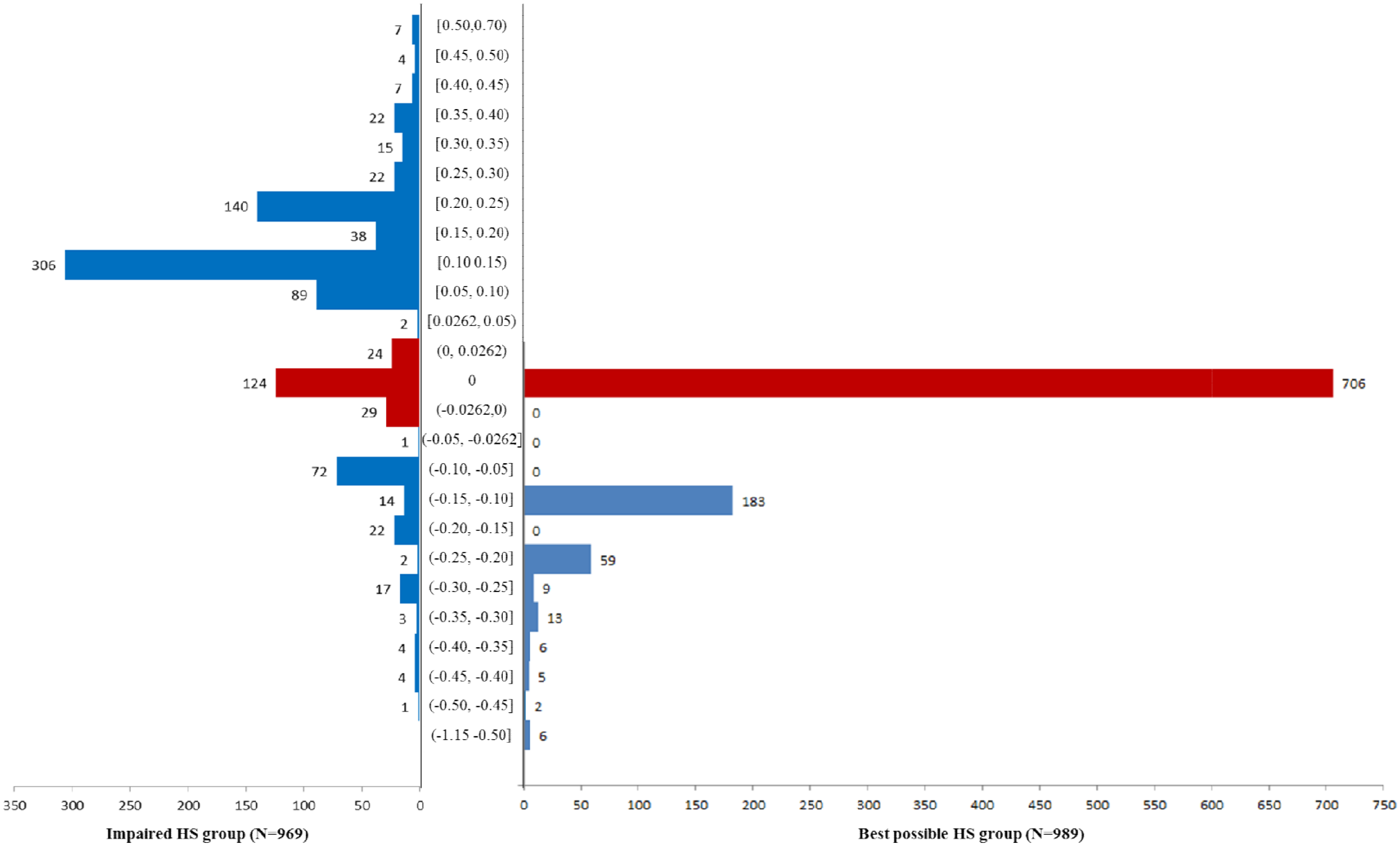


Figure 2.2 Change in EQ-5D-3L scores (1-year end - baseline)

HS Health state

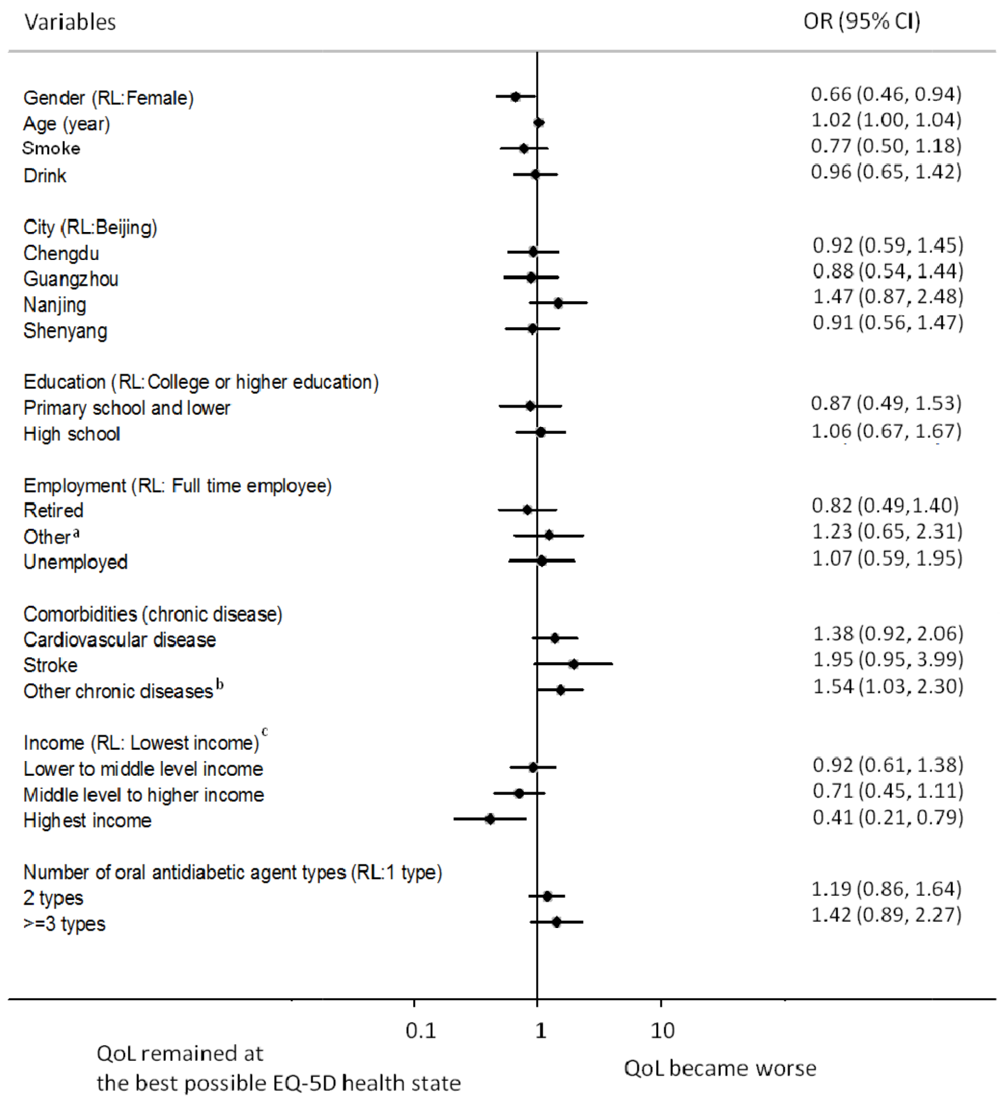


Figure 2.3 Logistic regression results for the best possible HS group

a. Includes individual freelancer, hourly worker, student, farmer and other;
b. Includes chronic lung diseases, asthma, cancer, gastric ulcer, duodenal ulcer, Parkinson's disease, hip or femur fracture, and other chronic diseases;
c. Per capita disposable income is grouped into the four categories according to the China 2011 per capita disposable income.
OR Odds ratio; *CI* Confidence interval; *RL* Reference level; *QoL* Quality of life; *HS* Health state.

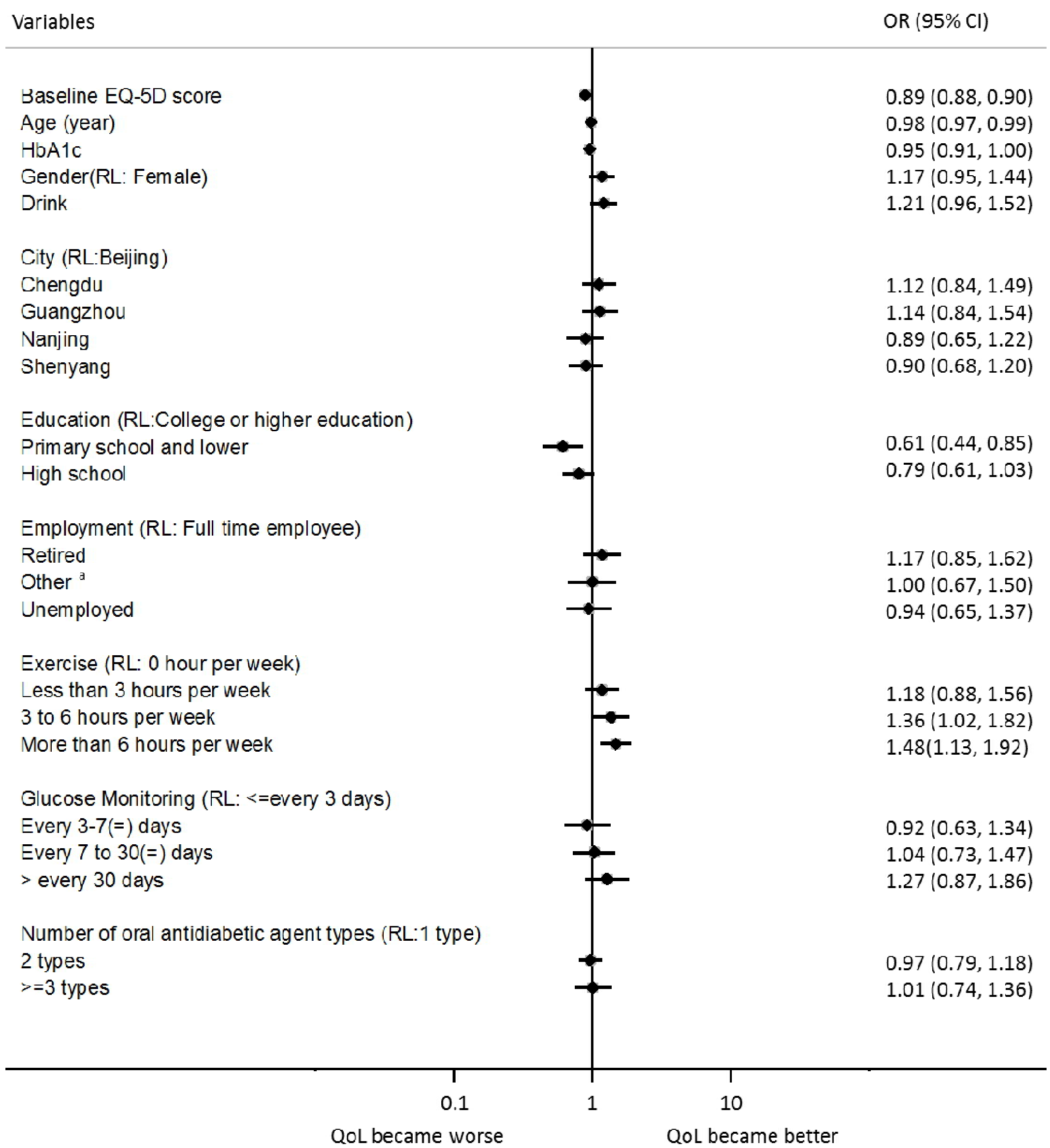


Figure 2.4 Ordered logistic regression results for the impaired HS group

a. Includes individual freelancer, hourly worker, student, farmer and other.

OR Odds ratio; *CI* Confidence interval; *HbA1c* Glycated hemoglobin; *RL* Reference level; *QoL* Quality of life; *HS* Health state.

Appendices

Appendix 2.1 Significant univariate regression results ($p<0.05$) [*] for the best possible HS group

Variable	***Coef***	***SE***	***P-value***
Gender (reference: female)	-0.451	0.142	0.001
Age	0.019	0.006	0.002
Education (reference: College or higher)			
Primary school and lower	0.621	0.239	0.009
High school	0.348	0.211	0.100
Employment (reference: full time employee)			
Retired	0.496	0.197	0.012
Other[a]	0.408	0.307	0.183
Unemployed	0.744	0.266	0.005
Cardiovascular diseases	0.407	0.190	0.032
Stroke or other cerebrovascular disease	0.893	0.346	0.010
Other chronic disease [b]	0.469	0.198	0.018
Smoking	-0.496	0.182	0.006
Drinking	-0.405	0.167	0.015
Income (reference: lowest income (<10672.0 CNY/yr))[c]			
Lower-income to middle level (10672.0 - 19544.9 CNY/yr)	-0.114	0.188	0.545
Middle level to higher-income (19544.9 - 35579.2 CNY/yr)	-0.394	0.196	0.044
Highest income (>35579.2 CNY/yr)	-1.016	0.298	0.001

* Only one (set of dummy) variable was put in the regression model at a time

a. Includes individual freelancer, hourly worker, student, farmer and other;

b. Includes chronic lung diseases, asthma, cancer, gastric ulcer, duodenal ulcer, Parkinson's disease, hip or femur fracture, and other chronic diseases;

c. Per capita disposable income is grouped into the four categories according to the China 2011 per capita disposable income.

HS Health state; *Coef* Coefficient; *SE* Standard error; *CNY/yr* Chinese Yuan/year.

Appendix 2.2Backward stepwise regression results (p<0.1) for the best possible HS group

Variable	*Coef*	*SE*	*P-value*
Gender (reference: female)	-0.438	0.145	0.003
Age	0.018	0.007	0.006
Stroke	0.697	0.360	0.053
Other chronic disease [a]	0.450	0.203	0.027
Income (reference: lowest income (<10672.0 CNY/yr))[b]			
Middle level to higher-income (19544.9 - 35579.2 CNY/yr)	-0.336	0.159	0.035
Highest income (>35579.2 CNY/yr)	-0.874	0.277	0.002
Number of oral anti-diabetic agent types (reference :1 type)			
2 types	0.190	0.158	0.228
>=3 types	0.341	0.225	0.130
_cons	-1.812	0.412	<0.001

a. Includes chronic lung diseases, asthma, cancer, gastric ulcer, duodenal ulcer, Parkinson's disease, hip or femur fracture, and other chronic diseases;

b. Per capita disposable income is grouped into the four categories according to the China 2011 per capita disposable income.

HS Health state; *Coef* Coefficient; *SE* Standard error; *CNY/yr* Chinese Yuan/year.

Appendix 2.3 Significant univariate regression results ($p<0.05$) [*] for the impaired HS group

Variable	*Coef*	*SE*	*P-value*
Gender(reference: female)	0.295	0.140	0.036
Age	-0.024	0.007	<0.001
Education (reference: Primary school and lower)			
High school	0.455	0.149	0.002
College or higher	0.945	0.247	<0.001
Employment (reference: full time employee)			
Retired	-0.468	0.237	0.049
Other [a]	-0.223	0.361	0.537
Unemployed	-0.509	0.284	0.073
Drinking	0.367	0.168	0.029
HbA1c level (%)	-0.081	0.035	0.020
Baseline EQ-5D-3L score	-0.025	0.007	<0.001

* Only one (set of dummy) variable was put in the regression model at a time

a. Includes individual freelancer, hourly worker, student, farmer and other.

HS Health state; *Coef* Coefficient; *SE* Standard error; *HbA1c*Glycated hemoglobin.

Appendix 2.4Backward stepwise regression results (p<0.1) for the impaired HS group

Variable	***Coef***	***SE***	***P-value***
Age	-0.030	0.007	<0.001
City (reference: Beijing)			
Chengdu	0.487	0.211	0.021
Education (reference: Primary school and lower)			
High school	0.313	0.162	0.053
College or higher	0.715	0.263	0.006
Exercise (reference: no exercise)			
less than 3 hours per week	0.478	0.215	0.026
3 to 6 hours per week	0.564	0.213	0.008
more than 6 hours per week	0.603	0.190	0.002
Glucose Monitoring			
every 3-7(=) days	-0.499	0.210	0.017
every 7 to 30(=) days	-0.286	0.161	0.076
Drinking	0.351	0.176	0.046
HbA1c level (%)	-0.084	0.037	0.025
Number of oral anti-diabetic agent types (reference :1 type)			
2 types	-0.007	0.149	0.963
>=3 types	0.223	0.225	0.320
Baseline EQ-5D-3L score	-0.039	0.007	<0.001

HS Health state; *Coef* Coefficient; *SE* Standard error; *HbA1c*Glycated hemoglobin.

Chapter 3

Item Reduction and Validation of the Chinese version of Diabetes Quality-of-Life Measure (DQOL)

Xuejing Jin, MSc [1,2], Gordon G. Liu, PhD [2,3], Hertzel C. Gerstein, M.D., MSc [4],Mitchell A.H. Levine, M.D., MSc [1], Kathleen Steeves, M.A. [5], Haijing Guan, MSc [2,6], Hongchao Li, MSc [7], Feng Xie, PhD [*1]

1. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
2. China Center for Health Economic Research, Peking University, Beijing, China
3. National School of Development, Peking University, Beijing, China
4. Department of Medicine, McMaster University, Hamilton, Ontario, Canada
5. Department of Sociology, McMaster University, Hamilton, Ontario, Canada
6. School of Pharmaceutical Science, Peking University, Beijing, China
7. School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China

* Corresponding author

Abstract

Background: Diabetes Quality-of-Life (DQOL) Measure is a 46-item diabetes-specific quality of life instrument. The original English version of the DQOL has been translated into Chinese after cultural adaption, and the Chinese DQOL has been validated in Chinese diabetic patient population and used in diabetes-related studies. There are two recognized problems of the Chinese DQOL: 1) the instrument is too long, and 2) the non-response rate of certain items is relatively high. This study aimed to develop and validate a short version for the Chinese DQOL.

Methods: Item reduction was conducted based on the classical test theory (CTT) and item response theory (IRT), each combined with exploratory factor analysis (EFA). The confirmatory factor analysis (CFA) and criterion validity tests were employed in validating the shortened versions.

Results: Both the training sample (n=2,886) and the validation sample (n=2,286) were from a longitudinal observation study of Chinese type 2 diabetic patients. The CTT retained 32 items, and the IRT retained 24 items from the original 46-item version. The two shortened versions were comparable in the psychometric properties.

Conclusion: The 24-item IRT-based short version of the Chinese DQOL was selected as the preferred short version because it impose lower burden to patients without compromising the psychometric properties.

Keywords: Item response theory, Classical test theory, Diabetes, Quality of life, Factor analysis

3.1Background

Diabetes mellitus (DM) is a chronic, lifelong disease that affects patients by causing increased concentrations of glucose in the blood and damaging the body's tissues, in particular the blood vessels, nerves, eyes, kidneys, and heart [1]. The global prevalence of DM in adults was 9.1% (415 million people) in 2015 which makes DM rated as one of the most common chronic diseases around the world [2]. Living with DM, patients usually need to take medication therapies and control diet to keep their blood glucose at a stable level [3]. Diabetic patients may experience adverse reactions from hypoglycemic agents or insulin, as well as diabetes-related complications (or anxiety about possible future complications). These concerns seriously affect the patients' (and their family members') quality of life (QoL) in both physical and psychological ways [4]. Hence, diabetic patients' QoL outcomes have been increasingly recognized as valuable and essential information in the fields of clinical research and diabetes management.

Diabetic patients' QoL is measured by generic or diabetes-specific instruments [5]. Diabetes-specific instruments are more sensitive to diabetes symptoms and the related impact on life and quality of life than generic instruments [6]. The

Diabetes Quality-of-Life measure (DQOL) is one of the most commonly used diabetes-specific instruments [7, 8]. It was developed and validated to compare two treatment regimens for chronic complications in patients with diabetes in the Diabetes Control and Complications Trial (DCCT) [9, 10]. The DQOL contains a total of 46 items, and all the items were categorized into one of the following four domains: life satisfaction (15 items), diabetes impact (20 items), social/vocational related worries (7 items), and diabetes related worries (4 items). The DQOL adopts a 5-point Likert scale for its response options. The scores range from 1, labeled as "very satisfied," to 5, labeled as "very dissatisfied," for items in the life satisfaction domain; from 1, labeled as "never impacted," to 5, labeled as "always impacted," for items in the diabetes impact domain; and from 1, labeled as "never worried," to 5, labeled as "always worried," for the social/vocational related and diabetes related worries domains.

The DQOL has been translated into five languages, including Chinese (Mandarin, Simplified) [11]. This measure was first translated and adapted for Chinese-Canadians who lived in the Toronto area by Cheng et al. [12, 13]. They removed 10 intimate personal items (e.g. sexual life) from the original DQOL and added six items regarding diet, worrying about death and so on. However, there was not sufficient psychometric evidence to support the cultural adaptation in Cheng et al.'s study [12], and the translation and validation were conducted based on an immigration population, which cannot necessarily be generalized to the entire Chinese diabetic patient population. Ding et al. translated and adapted the DQOL for Chinese population based on a sample of Chinese patients with diabetes who lived in Mainland China [14], and conducted validation of the Chinses DQOL on a separate sample ofChinese patients with type2 DM lived in Mainland China [15]. The wording of seven items was changed in Ding et al's adaptation (Appendix 3.1). Currently, the Chinese DQOL translated and adapted by Ding et al has been used in diabetes-related clinical studies in China [16-18]. During its application in the Chinese diabetic patient population, the Chinese DQOL has exposed some of its own issues [19]. First, the non-response rate of certain privacy-related items was relatively high; and second, interviewees complained that the instrument was too long [20, 21]. In order to solve these issues, developing and validating a short version of the Chinese DQOL is necessary.

Two psychometric theories can be employed in conducting item reduction. One is the Classical Test Theory (CTT), codified by Novick [22], which assumes that each respondent has a true total score, T (latent variable). Each item is a representative of the score T, and all items are of equal importance for measuring the score T. Generally, CTT tests the difficulty and discrimination at the item level and the reliability at the whole measure level [23]. Item Response Theory (IRT), developed by Lord [24, 25] and Rasch [26], refers to a family of latent trait models (logistic models) used to estimate the psychometric properties of items and scales.IRT is based on the relationship between respondents' performances on a single item and their performance in the overall ability that item was designed to measure [27]. This relationship is usually modeled by the item response function (IRF) which can provide estimations of the parameters "discrimination" and "location" at the individual item level. The item information function (IIF) can judge each item's ability to differentiate among respondents at the whole measure level [28]. Currently, researchers have been using the IRT alone [29], the combination of the IRT and factor analyses [30, 31], or the combination of the CTT and factor analysis [32, 33] when selecting or reducing items.

A relatively short QoL measure can be rapidly administered in practice and can reduce response burden on patients. The present study aimed to use both the CTT and IRT combined with factor analyses to derive and validate a short version of the Chinese DQOL.

3.2 Methods

3.2.1 Sample and data

We used the data from a Chinese community-based longitudinal survey of clinically diagnosed type 2 diabetic patients (T2DP) from five cities: Beijing, Chengdu, Guangzhou, Nanjing, and Shenyang. Patients were recruited and interviewed between December 2010 and October 2011, and followed every three months over a one-year study period. The Chinese DQOL and the EuroQol 5 dimensions 3-level (EQ-5D-3L) were administered at the baseline and at 12-months. Demographic, social-economic and diabetic-related information was also collected. Details of the survey have been described in Chapter 1.

We used the baseline data as the training sample for item reduction analysis, and the one-year end follow-up data as the validation sample to test the short versions of the Chinese DQOL reduced by CTT and IRT.

3.2.2 Reduction based on the Classical Test Theory

Three steps were used to reduce the number of items based on the CTT. The first step tested each item at the individual item level, and the second and third steps examined the items at the whole measure or domain level. The following provides the details of the tests in each step and the corresponding item removal criteria.

3.2.2.1 Step 1. Item level tests

We tested three item-level properties for each of the 46 items in this step, i.e., missing rate, item score mean, and item score standard deviation (SD).

Items which are unclear, ambiguous, or potentially embarrassing usually have a higher chance to have high non-response rate issues. This kind of items can provide very limited useful information, and their results are hard to interpret [34]. The exclusion criterion for the missing rate was higher than 5% [34].

In the CTT, item difficulty and discrimination are often evaluated in item level testing; however, most of the item difficulty and discrimination indexes are designed to test dichotomous items and can hardly be applied to test Likert items [23]. Norman has provided compelling evidence on the appropriateness of using descriptive statistics and parametric methods to test Likert items [35, 36]. The mean and SD of an item can provide fundamental information on whether the item can provide useful information or not [27]. For example, if the mean score is 4.7for a 5-point Likert item (score range: 1 to 5), then the item is left-skewed and may not be able to provide information the item was designed to collect. In addition, if the SD of an item is low, then the item has low variability and it may not be useful either. There are no generally accepted criteria for the item level test using mean and standard deviation, and we used the most lenient criteria reported in existing studies. We used the lowest score option plus 20% of the score range and the highest score option minus 20% of the score range to define the cut point of the exclusion criterion in terms of item score mean [37-39]. The lowest and highest score options for each item is 1 and 5, respectively, and the score range for each item is 4. Thus, the exclusion criterion for the item score mean was lower than 1.8 or higher than 4.2. The exclusion criterion for the item score SD was smaller than one-sixth of the score range, i.e., 0.67 (1/6*4) [37-40].

Any item that met any two or more of the three exclusion criteria was removed from the measure. In addition, any item with a missing rate higher than 10%was removed regardless of the results of the other two criteria.

3.2.2.2 Step 2. Exploratory factor analysis

In this step, exploratory factor analysis (EFA) was employed on the remaining items to examine the underlying structure of the measure and remove items with low factor loadings on common factors.

More specifically, the Bartlett's test of sphericity [41] and the Kaiser-Meyer-Olkin measure of sampling adequacy [42] were conducted before conducting the EFA. Scree plot was used to identify the number of factors [43]. Oblique rotation method was used in the EFA since the DQOL items were not completely unrelated to each other [44]. In this step, any item with a factor loading less than 0.3 was removed [45].

3.2.2.3 Step 3. Internal consistency reliability

Internal consistency reliability was tested in terms of the corrected item-total correlation and Cronbach's alpha [23]. Both tests were conducted at the factor level based on the results of the EFA in step 2.

Since there was no standard scoring method for the Chinese DQOL, we used the patients' mean score of the items in each factor as the “factor score” when calculating the corrected item-total correlation. For each item, the corrected item-total correlation was calculated as the Pearson correlation coefficient between the item score and the mean score of the rest of the items in the factor this item belonged to. A larger corrected item-total correlation coefficient indicates better internal consistency reliability. The exclusion criterion was the correlation coefficient smaller than 0.3 [46]. For the Cronbach's alpha, the exclusion criterion was that the Cronbach's alpha of the factor increased after removing an item [47].

In this step, any item that met one or more of these two exclusion criteria was removed from the measure. An additional EFA was used to check if the factor structure changed after this step; if so, the new factor structure would be used as the final structure of the short version developed based on the CTT.

3.2.3 Reduction based on the Item Response Theory

Two steps were used to reduce the number of items based on the IRT.

3.2.3.1 Step 1. Exploratory factor analysis

Because one of the basic assumptions of the IRT is unidimensionality, and all IRFs need to be established at the factor level, we first conducted EFA using the oblique rotation method on all of the 46 items to explore the factors of the Chinese DQOL. Any item with a factor loading of less than 0.3 was removed in this step.

3.2.3.2 Step 2. Item response theory analysis

The graded response model (GRM), which is a type of item response model for items with ordered response options [48], was employed in this step to analyze the remaining items within each factor identified in step 1.

The GRM was first introduced by Samejima [48]. It models each item with its own discrimination parameter and a set of parameters that identify the boundaries between the ordered options using a logistic regression approach. The probability of respondent *j* with latent ability level θ_j (the latent trait for respondent *j*) to choose response option *k* or higher (in our case, *k*=0, 1, 2, 3, 4, 5) for item *i* is [48, 49]:

$$Pr\ (Y_{ij} \geq k|\theta_j) = \frac{exp\ \{a_i(\theta_j - b_{ik})\}}{1 + exp\ \{a_i(\theta_j - b_{ik})\}} \theta_j \sim N(0,\ 1)$$

where, a_i represents the discrimination of item *i*, and b_{ik} is the cut-point of boundaries between the k^{th} and $(k+1)^{th}$ options for item *i*, which can be considered as the difficulty of choosing option *k* or higher for item *i* [48, 49].

The IIFs were built based on the fitted GRMs to evaluate the "information," i.e., reliability, each item contributed to the factor. The information function $I_i(\theta)$ for item *i* is :

$$I_i(\theta) = \sum_{k=1}^{K} I_{ik}(\theta) p_{ik}(\theta)$$

where, $I_{ik}(\theta)$ is the information function, for response option k of item i, which is defined as:

$$I_{ik}(\theta) = -\frac{\partial^2 \log p_{ik}(\theta)}{\partial \theta^2}$$

where, ∂ is the partial derivative symbol, and $p_{ik}(\theta)$ is the probability of a respondent with the latent trait level θ choosing response option k, which depends on the GRM for item i.

In this step, any item that had an estimation of discrimination parameter less than 1.0 [50] and provided item information less than 0.5 was removed from the measure [31]. An additional EFA was also conducted to check the factor structure; and if the structure changed after this step, the new factor structure would be used as the final structure of the short version developed based on the IRT.

3.2.4 Validating and comparing the two short versions of Chinese DQOL

3.2.4.1 Confirmatory factor analysis

Confirmatory factor analysis (CFA) was employed to validate the structure of the two short versions of the Chinese DQOL. Two statistics produced by the CFA were used to compare the performance of the two versions: standardized root mean squared residual (SRMR) and comparative fit index (CFI).

The SRMR is the square root of the difference between the residuals of the sample covariance matrix and the proposed covariance model. It ranges from 0 to 1, and a smaller value indicates a better fit [51]. The CFI compares the sample covariance matrix with this null model based on the assumption that all latent variables (factors) are uncorrelated. The CFI ranges from 0 to 1, and a larger value indicates a better fit [51].

3.2.4.2 Criterion validity

We tested the criterion validity of the two reduced versions of the Chinese DQOL against the EQ-5D-3L index and EQ visual analogue scale (EQ-VAS).

The EQ-5D-3L is a widely used preference-based generic quality of life instrument which has 5 questions that ask about whether there are any problems in: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each question has three response levels, i.e., no problems, some (or moderate) problems, and extreme problems (or unable to). Patients' EQ-5D-3L responses were converted into EQ-5D-3L values by using the Chinese EQ-5D-3L value set [52]. The EQ-VAS records the patient's self-rated health on a vertical, visual analogue scale which ranges from 0 (the worst imaginable health state) to 100 (the best imaginable health state) [53].

Spearman's correlation coefficients between the EQ-5D-3L index and the mean score of each one of the two short versions of the Chinese DQOL were calculated respectively. The correlation coefficients between the EQ-VAS and the two short versions were also calculated individually. A larger correlation coefficient indicates higher criterion validity [23, 34].

In the event of any conflict between the CFA and the criterion validity results, we used the CFA results as our primary evaluation criteria. All statistical analyses were conducted with a two-tailed test at significance level of 0.05 in STATA 14.2 (StataCorp LP, Texas, USA).

3.3 Results

3.3.1 Sample

A total of 2,886 patients were recruited and interviewed at the baseline. The mean age and diabetes duration of the training sample were 61.15 years and 7.94

years, respectively. Among all patients, 55.68% were female, 64.10% were retired, and 16.18% used insulin in the last 6 months. The mean scores of the EQ-5D-3L index, VAS, and the Chinese DQOL (mean score of the 46 items) were 0.89, 72.71, and 2.07, respectively (Table 3.1). In the validation analyses, the CFA and the calculation of the EQ-5D-3L index only employed observations without missing data. Because of this, our validation sample only included patients with no missing values on responses to the 5 questions of the EuroQol 5 dimensions (EQ-5D) and to the DQOL items kept after the item reduction based on the CTT and IRT. Of the 2,542 patients who completed the year-end follow-up, 2,286 were included in the validation sample (Table 3.1). Compared to the training sample, the validation sample had a higher proportion of people who were older, retired, and used insulin (Table 3.1).Details of the patients' characteristics have been described in Chapter 2.

3.3.2 Item reduction results

Tables 3.2 and 3.3 show the item reduction results based on the CTT and IRT, respectively. A total of 14 and a total of 22 items were removed from the Chinese DQOL based on the CTT and IRT, respectively.

In step 1 of the reduction based on the CTT, two items, item #10 (satisfied with sex life) and item #25 (interfere with sex life) were removed from the measure because their missing rates were higher than 10%. Item #32 (being teased because of having diabetes), item #36 (worry about marriage), item #40 (worry about completing education), and item #41 (worry about unemployment) were removed because of their low mean scores (all <1.8) and small SDs (all <0.67). Item #35 (hide having an insulin reaction) was removed because of the high missing rate (8.07%) and low mean score and small SD. In step 2, the EFA identified two factors among the remaining items. Item #7 (satisfied with knowledge about diabetes), item #23 (feel good about yourself), item #26 (interfere with riding a bike or using a machine), item #29 (explain what it means to have diabetes), item #31(tell others about your diabetes), and item #34 (eat something you shouldn't rather than tell someone that you have diabetes) were removed due to the low factor loadings (<0.3). In step 3, item #38 (worry about whether can get a job you want) was removed because of the low correlation with the mean score of the factor it belonged to. The factor structure identified in Step 2 remained the same after removing item #38 in Step 3.

In the reduction based on the IRT, the EFA identified 2 factors of the 46 DQOL items, and removed items #7, #23, #26, #29, #31, and #34 because their factor loading were all smaller than 0.3. In step 2, item #5 (satisfied with the flexibility on the diet), item #8 (satisfied with sleep), item #10, item #12 (satisfied with the appearance of your body), item #13 (satisfied with the time spent on exercising), item #18 (low blood sugar reactions), item #21 (bad night's sleep), item #24 (feel restricted by diet), item #25, item #32, item #33 (feel that because of diabetes you go to the bathroom more than others), item #38, item #39 (worry about the pension), item #40, and item #41 were removed in the IRT analysis due to their item discrimination being smaller than 1 and their item information being lower than 0.5 (Table 3.3). The factor structure identified in the EFA remained the same after the IRT analysis.

3.3.3 Validation results

Table 3.4 shows the validation results of the two short versions of the Chinese DQOL. In the CFA, the two short versions had similar SRMRs (0.078, after rounding, for both short versions) which were also similar to that of the original Chinese DQOL (SRMR=0.077). The short version based on the IRT had a larger CFI (0.726) than that of the version reduced based on the CTT (CFI=0.630). The CFI of both short versions were larger than that of the original Chinese DQOL (CFI=0.616).

In the criterion validity tests, the absolute Spearman's correlation coefficient between the CTT reduced version of the DQOL and the EQ-5D-3L index scores was 0.298, which was higher than that (ρ=0.288) between the IRT reduced version and the EQ-5D-3L index scores. Both reduced versions had a higher correlation with the EQ-5D-3L index scores than the original Chinese DQOL (ρ=0.276). In terms of using the EQ-VAS as the criterion, the CTT-based short version had a higher correlation (ρ=0.288) than the original version (ρ=0.273), and the IRT-based short version had a slightly lower correlation (ρ=0.269) than the original version.

3.4 Discussion

This study shortened the 46-item Chinese version of the DQOL based on two psychometric theories, the CTT and IRT, each combined with the EFA, respectively. The two shortened versions were validated using the CFA and criterion validity tests. The CTT provided a short version of the Chinese DQOL with 32 items retained, and the IRT provided a short version with 24 items retained. Among the 14 items removed based on the CTT, 13 were removed based on the IRT as well.

There were few published studies we can compare our results with. Two items related to sexual life had high missing rates in our study and were removed from the measure in the reduction processes based on both the CTT and IRT. This was consistent with the translation and cultural adaptation study conducted in 1999 among Chinese diabetic patients lived in Canada [13]. The high missing rate of the sexual life items is still in line with the findings in translation and cultural adaptation studies published after 2015 in other disease specific measures among the Chinese population [54]. Chinese people, especially those who are middle-aged and elderly, tend to be hesitant to talk about sex-related topics because of their relatively conservative culture background [55].

Three working and education-related items, i.e., items #38, #40, and #41, had low mean scores (Table 3.2) and low discriminations (Table 3.3), and were removed based on both the CTT and IRT. This was because most patients (64.10%) in our training sample were retired, and they were not worried about working and education-related issues. These items were also removed according to the expert advice in Chen's [17] translation and cultural adaptation study.

The insulin reaction item (item #35) was removed based on both the CTT and IRT. This was because the majority of the patients in the training sample did not use insulin in the last 6 months. Similarly, the diet-related item (item #34) was also removed mainly because the majority of the patients in the training sample controlled their diet by eating healthy food and balancing the amount of food intake due to their diabetes.

In Ding's [14] translation and cultural adaptation analysis, the descriptive of item 26, "How often does your diabetes keep you from driving a car or using a machine (e.g., a typewriter)?" was changed into "How often does your diabetes keep you from riding a bike or being a typist?" This item was removed because of low factor loading in both reduction processes. Ding et al changed the "driving a car" into "riding a bike" because the civilian vehicle ownership in China was relatively low in the 1990's, and bicycles were the main means of transportation for ordinary people. However, civilian vehicle ownership in 2012 increased by 544% from 1999 [56], which may make this change in descriptive out-of-date. In addition, typewriters have long been replaced by laptops and other smart electronics which are indispensable in contemporary Chinese people's daily lives. Therefore, further studies examining the performance of a more up-to-date descriptive, for example, "How often does your diabetes keep you from driving a vehicle or using a computer or smart phone?" are necessary.

There were 9 items that were removed in the IRT-based short version but kept in the CTT-based short version. All of these items were removed due to their low estimated discrimination and item information in the IRT analysis. One of the possible reasons for this different is that the reduction results were impacted by the exclusion criteria we employed. Even though we used the most lenient fail criteria reported in existing studies for each, respectively, the item reduction results may still not be comparable due to the different statistical approaches applied in the two different theories.

Items #1 to #4 (satisfaction level of “the amount of time it takes to manage your diabetes,” “the amount of time you spend getting a checkup,” “the time it takes to determine your sugar level,” and “your current treatment”) were the only four treatment and diabetes management related items in the DQOL. These items loaded onto the same factor in our EFA. The rest of the 28 items in the CTT-based short version and the rest of the 20 items in the IRT-based short version belonged to the other factor, respectively. This was different than the original Chinese DQOL which has four domains. The CFA and criterion validity results showed that the structures of the two short versions were comparable to the original version. In addition, we did not emphasize the name of the factors identified in the short versions since the present study aimed to focus on reducing the number of items for the Chinese DQOL. Content and face validity of the short versions should be examined in further studies to optimize the structure and rename the factors of the short versions.

The often used fit indexes in the CFA are the Chi-square test and the root mean square error of approximation (RMSEA) [51]. In the present study, we employed the SRMR and CFI instead of the Chi-square test and RMSEA. The Chi-square test result is affected by the number of parameters, the complexity of the model, and the sample size [57]. Adding more parameters into the model can improve the RMSEA as well [58]. Our two short versions of the Chinese DQOL had different numbers of items; therefore, the Chi-square test and RMSEA were inappropriate to use for comparing the CFA results of these two short versions. The SRMR is not affected by the model complexity and the number of parameters. The CFI is affected by the number of parameters added into a model, but it is relatively more stable than the Chi-square test and RMSEA.

Considering the two short versions of the Chinese DQOL were comparable in the validation analysis, we selected the short version based on the IRT (24 items) as a preferred short version for two other reasons. First, this shorter version imposes a lower burden on patients without compromising its measurement properties [59]. Second, theoretically, as a modeling statistic approach, the parameters estimated from a set of IRF can be generalized to the entire population where the study sample comes from; however, as a person-based statistic approach, all CTT test results can only be specified to the given study sample [60].

There are some limitations in our study. First, the study and validation samples were not independent. We did not have a truly external validation sample for our study. Second, our training sample only contained community-based patients, and most of them did not use insulin; therefore, our results cannot necessarily be generalized to the entire diabetic patient population. Other psychometric properties such as test-retest reliability of the short version of the Chinese DQOL need to be examined in future studies.

References

[1] World Health Organization. Diabetes mellitus. From http://www.who.int/mediacentre/factsheets/fs138/en/.(Accessed on December 18, 2016).

[2] International Diabetes Federation. (2015). IDF Diabetes Atlas, 7 ed. Brussels, Belgium: International Diabetes Federation.

[3] Isla Pera P. (2011). Living with diabetes: quality of care and quality of life. Patient Prefer Adherence. 5: 65–72.

[4] Rubin RR. (2000). Diabetes and quality of life.Diabetes Spectrum. 13: 21.

[5] Rubin RR,Peyrot M. (1999). Quality of life and diabetes.DiabetMetab Res Rev. 5: 205–218.

[6] Ware JE Jr. Gandek B, Guyer R, Deng N. (2016). Health Qual Life Outcomes. 14: 84-99.

[7] El Achhab Y, Nejjari C, Chikri M, Lyoussi B. (2008). Disease-specific health-related quality of life instruments among adults diabetic: A systematic review. Diabetes Res ClinPract. 80(2):171-184.

[8] Watkins K, Connell CM. (2004). Measurement of Health-Related QOL in Diabetes Mellitus.Pharmacoeconomics. 22(17):1109-1126.

[9] Jacobson AM, Barofsky I, Cleary P, Rand LL. (1988). Reliability and validity of a diabetes quality-of-life measure for the diabetes control and complications trial (DCCT).Diabet Care. 11: 725–732.

[10] Jacobson AM. (1997). Quality of life in patients with diabetes mellitus.SeminClinNeuropsychol. 2: 82–93.

[11] PROQOLID. Diabetes Quality Of Life Measure (DQOL). From: https://eprovide.mapi-trust.org/instruments/diabetes-quality-of-life-measure#languages. (Accessed on December 18, 2016).

[12] Cheng AY, Tsui EY, Hanley AJ, Zinman B. (1999). Cultural adaptation of the diabetes quality-of-life measure for Chinese patients.Diabetes Care. 22(7): 1216-1217.

[13] Cheng AY, Tsui EY, Hanley AJ, Zinman B. (1999). Developing a quality of life measure for Chinese patients with diabetes.Diabetes Res

ClinPract.46(3):259-267.

[14] Ding Y, Kong D, Ni Z, Deng H. (2004). Culture adaption and revision of diabetes-specific quality of life scale (QDOL). Chin J Behav Med Sci. 13(1): 102-103.

[15]Ding Y, Ni Z, Zhang J, Chen G, Feng H. (2000). The Assessment on Reliability and Validity of Adjusted Diabetes Quality of Life (A-DQOL) Scale. Chin J PrevContrChron Non-commun Dis. 8(4): 160-162.

[16]Liang M. (2014) Quality of life and its impact factors of community-based patients with type 2 diabetes in Beijing.Master degree thesis.Beijing university of Chinese medicine.

[17] Chen A, Su A, Bai H. (2009). Factors impact on the quality of life of type 2 diabetic patients. Chin Community Doctors. 11(19): 38.

[18] Hou Y, Yang Q. (2016). Current research status of health-related quality of life in patients with diabetes mellitus. Chin J Clinicians. 10 (3): 433-436.

[19] Ren Z. (2010). Quality of life in the patients of diabetes mellitus with micro-albuminuria but no renal insufficiency.Master degree thesis.Zhejiang University.

[20] Li D, Ma A, Li H. (2012). Systematic review of diabetes-specific quality of life measures in China. Chin J Pharm Econ. 34: 45-52.

[21] Qu L, Pan M. (2007). Research progress of diabetes-specific quality of life measures in China. 16 (8): 765-766.

[22] NovickMR. (1966). The axioms and principal results of classical test theory. J. Math Psychol. 3(1): 1-18.

[23] Kline TJB. (2005). Psychological Testing: A Practical Approach to Design and Evaluation. Thousand Oaks, CA: SAGE PUBLICATIONS INC.

[24] Lord FM. (1952). A Theory of Test Scores. New York: Psychometric Society.

[25] Lord FM. (1980). Applications of Item Response Theory to Practical Testing Problems. Mahwah, New Jersey, United States: Lawrence Erlbaum Associates.

[26] RaschG. (1960). Probabilistic models for some intelligence and attainment tests. Chicago: The University of Chicago Press.

[27] Cappelleri JC, Lundy JJ, Hays RD. (2014). Overview of Classical Test Theory and Item Response Theory for Quantitative Assessment of Items in Developing Patient-Reported Outcome Measures.ClinTher. 36(5): 648–662.

[28] An X, Yung Y, SAS Institute Inc.. Item Response Theory: What It Is and How You Can Use the IRT Procedure to Apply It. From: https://support.sas.com/resources/papers/proceedings14/SAS364-2014.pdf (Accessed on December 26, 2016)

[29] ten Holt JC, van Duijn MAJ,Boomsma A. (2010). Scale construction and evaluation in practice: A review of factor analysis versus item response theory applications. Psychol Test Assess Model. 52 (3): 272-297.

[30] Edelen MO, Reeve BB. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Qual Life Res. 16:5–18.

[31] Weinhardt JM, Morse BJ, Chimeli J, Fisher J. (2012). An item response theory and factor analytic examination of two prominent maximizing tendency scales.JudgmDecisMak. 7(5): 644-658.

[32] The WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment.Psychol Med, 28: 551-558.

[33] Prieto L, Alonso J, Lamarca R. (2003). Classical test theory versus Rasch analysis for quality of life questionnaire reduction.Health and Qual Life Outcomes. 1:27. doi: 10.1186/1477-7525-1-27.

[33] Streiner DL and Norman GR. (2008). Health measurement scales. A practical guide to their development and use (fourth edition). Oxford: Oxford University Press.

[34] Fayers PM and Machin D. (2016). Quality of life: the assessment, analysis and reporting of patient-reported outcomes (third edition). West Sussex, UK: John Wiley & Sons.

[35] Norman G. (2010). Likert scales, levels of measurement and the ‘‘laws’’ of statistics.Adv in Health Sci Educ. 15:625–632.

[36] Sullivan GM and Artino AR. (2013). Analyzing and Interpreting Data From Likert-Type Scales.J Grad Med Educ. 5(4): 541–542.

[37] Streiner DL and Norman GR.(2008). Health measurement scales. A practical guide to their development and use (fourth edition). Oxford: Oxford University Press.

[38] Lester PE, Inman D, Inman Freitas DL, Bishop LK. (2014). Handbook of Tests and Measurement in Education and the Social Sciences.Lanham: Rowman & LittlefieldPublishing Group.

[39] Nunnally JC, Bernstein IH.(1994). Psychometric Theory (third edition). New York: McGrawHill

[40] Qiu H. (2009). Quantitative research and statistical analysis. Chongqing: Chongqing University Press.

[41] Bartlett MS. (1950).Tests of significance in factor analysis.Br J Stat Psychol. 3: 77–85.

[42] Kaiser HF. (1970). A second generation Little Jiffy.Psychometrika, 35:401-415.

[43] Ledesma RD, Valero-Mora P and Macbeth G. (2015).The Scree Test and the Number of Factors: a DynamicGraphics Approach.Span J Psychol. 18(e11): 1–10.

[44] Osborne JW. (2015). What is Rotating in Exploratory Factor Analysis? PARE. 20(2).

[45] Thompson B. (2004). Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications. Washington, DC: American Psychological Association.

[46] Pallant J. (2007). SPSS survival manual: A step by step guide todata analysis using SPSS for windows (3th ed.). New York,USA: Mc Graw Hill

[47]Gliem JA. andGliemR R. (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. from: https://scholarworks.iupui.edu/bitstream/handle/1805/344/Gliem+&+Gliem.pdf?sequence=1.(Accessed on December 26, 2016).

[48] Samejima F. (1969). Estimation of Latent Ability Using a Response Pattern ofGraded Scores(Psychometric Monograph No. 17). Richmond, VA: PsychometricSociety. Fromhttp://www.psychometrika.org/journal/online/MN17.pdf (Accessed on December 26, 2016).

[49] StataCorp. (2015). Stata Item Response Theory Reference Manual (Release 14). College Station, TX: StataCorp LP. From http://www.stata.com/manuals14/irt.pdf (Accessed on December 26, 2016).

[50] Zickar MJ, Russell SS, Smith CS, BohleP,Tilley AJ. (2002). Evaluatingtwo morningness scales with item response theory. Personality and IndividualDifferences, 33, 11–24.

[51] Hooper D, Coughlan J, and Mullen MR. (2008). Structural equation modelling: Guidelines for determining model fit.J Bus Res Meth.6: 53–60.

[52] Liu GG, Wu H, Li M, Gao C, Luo N. (2014). Chinese time trade-off values for EQ-5D health states.Value Health. 17(5):597-604.

[53] EuroQol Group. How to use EQ-5D. From: http://www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html. (Accessed on December 26, 2016).

[54] Chen X, Qiu Z, Gu M, Su Y, Liu L, Liu Y, Mo C, Xu Q, Sun J, Li D. (2015). Translation and validation of the Chinese version of the Quality OF Life Radiation Therapy Instrument and the Head & Neck Module (QOL-RTI/H&N).

[55] Global Affairs Canada. Cultural Information – China. From: https://www.international.gc.ca/cil-cai/country_insights-apercus_pays/ci-ic_cn.aspx?lang=eng. (Accessed on January 16, 2017).

[56] National Bureau of Statistics of China. (2013). China Statistical Yearbook 2013. From: http://www.stats.gov.cn/tjsj/ndsj/2013/indexch.htm .(Accessed on December 26, 2016).

[57] Moss S. Fit indices for structural equation modeling. From: http://www.sicotests.com/psyarticle.asp?id=277. (Accessed on December 26, 2016).

[58] Kenny DA and McCoach DB. (2003) Effect of the Number ofVariables on Measures of Fit in Structural Equation Modeling.StructEqu Modeling.10(3): 333-351.

[59] Schmitt N and Stuits DM. (1985).Factors Defined by NegativelyKeyed Items: The Resultof Careless Respondents?ApplPsychol Meas. 9(4):367-373.

[60] Fan X. (1998). Item response theory and classical test theory: an empirical comparison of their item/person statistics. EducPsychol Meas. 58(3), 357-381.

Tables

Table 3.1 Patients' baseline demographic and diabetes-related information

	Training sample (N=2,886)	Validation sample (N=2,286[a])	P-value [b]
Age mean (SD), years	61.15 (11.42)	61.84 (11.25)	0.030
Female	1,607 (55.68)	1,269 (55.51)	0.910
Diabetes duration mean (SD), years	7.94 (6.75)	7.69 (6.33)	0.174
Retired	1,850 (64.10)	1,544 (67.54)	0.010
Used insulin in the last 6 months	467 (16.18)	476 (20.82)	<0.001
Controlled diet	2580 (89.40)	2,036 (89.06)	0.718
EQ-5D-3L index mean (SD)	0.89 (0.14)	0.89 (0.13)	0.999
EQ-ED VAS mean (SD)	72.71 (15.46)	73.00 (15.48)	0.503
DQOL score mean (SD)[c]	2.07 (0.38)	2.07 (0.39)	0.999

The values presented are numbers (percentage) unless otherwise stated.

a. A total of 2,542 patients finished the last round follow-up at the year end. Only observations with no missing data on the EQ-5D-3L and DQOL questions were included in the validation sample;

b. T-test for mean and Chi-square test for frequency;

c. The DQOL score was calculated from the mean of the 46 items.

SD Standard deviation; *VAS* Visual analogue scale; *DQOL* Diabetes quality-of-life measure.

Table 3.2 Item reduction results based on the CTT

Item No.	Step 1. Item level tests			Step 2. EFA		Step 3. Internal consistency reliability	
	Missing rate (%)	Mean score	SD	Factor 1	Factor 2	Item-total corrected correlation coefficient [b]	Cronbach's alpha [b]
	Missing rate>5% [a]	<1.8 or >4.2 [a]	<0.67 [a]	Factor loading <0.3 [a]		r<0.3 [a]	Increased after remove [a] (Factor 1: 0.884 Factor 2: 0.822)
1	0.312	2.081	0.711	0.384	0.528[※]	0.667	0.764
2	0.277	2.132	0.693	0.337	0.519[※]	0.648	0.772
3	0.243	2.078	***0.649***	0.357	0.550[※]	0.695	0.751
4	0.104	2.174	0.725	0.388	0.448[※]	0.564	0.811
5	0.104	2.653	0.926	0.394[※]	0.172	0.348	0.882
6	0.069	3.129	0.994	0.518[※]	0.015	0.492	0.879
7	0.069	2.563	0.811	***0.297***	***0.170***	Removed	
8	0.035	2.740	1.090	0.398[※]	0.133	0.367	0.882
9	0.139	2.112	0.679	0.420[※]	0.262	0.360	0.882
10	***14.969***	2.449	0.769	Removed			
11	0.485	2.161	0.692	0.485[※]	0.234	0.419	0.880
12	0.035	2.594	0.912	0.406[※]	0.106	0.361	0.882
13	0.312	2.419	0.900	0.398[※]	0.171	0.345	0.882
14	0.277	2.290	0.748	0.492[※]	0.236	0.434	0.880
15	0.035	2.158	***0.656***	0.566[※]	0.266	0.498	0.879
16	0.139	2.466	1.201	0.557[※]	-0.137	0.539	0.878
17	0.069	***1.727***	0.953	0.431[※]	-0.166	0.415	0.881
18	0.035	2.321	1.072	0.381[※]	-0.068	0.379	0.882
19	0.035	2.556	1.075	0.534[※]	-0.018	0.520	0.879
20	0.069	2.580	1.223	0.608[※]	-0.165	0.597	0.877
21	0.035	2.631	1.225	0.423[※]	0.060	0.396	0.882
22	0.069	***1.724***	0.892	0.492[※]	-0.124	0.468	0.880
23	0.035	2.729	1.135	***0.230***	***0.006***	Removed	
24	0.035	2.960	1.239	0.381[※]	-0.132	0.381	0.882

Table 3.2(Cont). Item reduction results based on the CTT

Item No.	Step 1. Item level tests			Step 2. EFA		Step 3. Internal consistency reliability	
	Missing rate (%)	Mean score	SD	Factor 1	Factor 2	Item-total corrected correlation coefficient [b]	Cronbach's alpha [b]
	Missing rate>5% [a]	<1.8 or >4.2 [a]	<0.67 [a]	Factor loading <0.3 [a]		r<0.3 [a]	Increased after remove [a] (Factor 1: 0.884 Factor 2: 0.822)
25	***14.414***	1.875	0.968	Removed			
26	1.802	***1.581***	0.961	***0.245***	***-0.173***	Removed	
27	0.416	***1.790***	1.012	0.487※	-0.144	0.451	0.880
28	0.104	1.843	1.047	0.514※	-0.152	0.490	0.879
29	0.035	2.445	1.176	***0.185***	***-0.228***	Removed	
30	0.208	1.882	0.977	0.574※	-0.192	0.526	0.878
31	0.035	2.343	1.079	***0.222***	***-0.187***	Removed	
32	0.069	***1.297***	***0.594***	Removed			
33	0.104	2.499	1.260	0.355※	-0.089	0.342	0.883
34	0.035	***1.781***	0.945	***0.234***	***-0.029***	Removed	
35	***8.073***	***1.136***	***0.415***	Removed			
36	2.495	***1.229***	***0.593***	Removed			
37	0.728	1.888	1.147	0.429※	-0.166	0.420	0.881
38	3.222	***1.262***	0.685	0.314※	-0.177	***0.294***	0.883
39	2.668	***1.311***	0.762	0.332※	-0.197	0.305	0.883
40	4.089	***1.134***	***0.431***	Removed			
41	3.915	***1.213***	***0.633***	Removed			
42	0.312	***1.641***	1.018	0.424※	-0.263	0.414	0.881
43	0.104	1.998	1.151	0.519※	-0.264	0.508	0.879
44	0.069	***1.551***	0.908	0.441※	-0.225	0.425	0.881
45	0.104	2.632	1.317	0.502※	-0.182	0.487	0.880
46	0.104	***1.353***	0.720	0.427※	-0.250	0.414	0.881

Bold and italic number indicates the item failed the corresponding test.
Dashed box indicates the item(s) was removed from the scale; "※" indicates which factor the item belongs to based on the EFA.
a. Exclusion criteria;
b. Total scores were calculated as the corrected mean score of the factor;
CTT Classical test theory, *EFA* Exploratory factor analysis, *SD* Standard deviation, *N/A* Not applicable.

Table 3.3 Item reduction results based on the IRT

Item No.	Step 1. EFA		Step 2. IRT	
	Factor 1	Factor 2	Discrimination	Item information function [b]
	Factor loading<0.3[a]		<1[a]	<0.5[a]
1	0.356	0.421※	2.830	>2, <2.5
2	0.313	0.405※	2.848	>2, <2.5
3	0.342	0.407※	3.496	>3, <3.5
4	0.357	0.368※	1.987	>1, <1.5
5	0.361※	0.248	***0.778***	***<0.2***
6	0.506※	0.115	1.097	<0.4
7	***0.274***	***0.181***	Removed	
8	0.369※	0.258	***0.769***	***<0.2***
9	0.427※	0.203	1.024	<0.3
10	0.343※	0.153	***0.768***	***<0.2***
11	0.470※	0.160	1.211	<0.5
12	0.365※	0.184	***0.854***	***<0.3***
13	0.373※	0.171	***0.884***	***<0.3***
14	0.466※	0.218	1.147	<0.4
15	0.544※	0.207	1.485	>0.6, <0.7
16	0.544※	0.045	1.328	>0.5, <0.6
17	0.456※	-0.128	1.237	<0.5
18	0.368※	0.077	***0.801***	***<0.2***
19	0.494※	0.186	1.178	<0.5
20	0.592※	0.031	1.567	<0.5
21	0.394※	0.205	***0.850***	***<0.3***
22	0.491※	-0.063	1.506	>0.6, <0.7
23	***0.220***	***-0.023***	Removed	
24	0.352※	0.068	***0.793***	***<0.2***
25	0.404※	-0.076	***0.991***	***<0.3***
26	***0.262***	***-0.150***	Removed	
27	0.485※	-0.052	1.429	>0.6, <0.7
28	0.519※	-0.041	1.517	>0.7, <0.8
29	***0.173***	***-0.059***	Removed	
30	0.575※	-0.087	1.666	>0.8, <0.9
31	***0.212***	***-0.017***	Removed	
32	0.346※	-0.242	***0.998***	***<0.3***
33	0.345※	0.042	***0.762***	***<0.2***
34	***0.271***	***-0.060***	Removed	
35	***0.214***	***-0.197***	Removed	
36	0.383※	-0.351	1.068	<0.4

Table 3.3 (Cont). Item reduction results based on the IRT

Item No.	Step 1. EFA		Step 2. IRT	
	Factor 1	Factor 2	Discrimination	Item information function [b]
	Factor loading<0.3[a]		<1[a]	<0.5[a]
37	0.449※	-0.162	1.099	<0.4
38	0.410	-0.514※	***0.321***	***around 0***
39	0.403	-0.441※	***0.303***	***around 0***
40	0.330	-0.461※	***0.360***	***around 0***
41	0.372	-0.510※	***0.289***	***around 0***
42	0.433※	-0.182	1.224	<0.5
43	0.526※	-0.089	1.357	>0.5, <0.6
44	0.450※	-0.180	1.306	>0.5, <0.6
45	0.479※	0.021	1.048	<0.4
46	0.465※	-0.280	1.515	>0.7, <0.8

Bold and italic number indicates the item failed the corresponding test.
Dashed box indicates the item(s) was removed from the scale.
a. Exclusion criteria;
b. Highest point on the item information function curve;
"※" indicates which factor the item belongs to based on the EFA.
IRT Item response theory, *EFA* Exploratory factor analysis.

Table 3.4 Validation results [a]

	Short version based on the CTT (32 items)	Short version based on the IRT (24 items)	Original Chinese DQOL (46 items)
Confirmatory factor analysis			
SRMR	0.078	0.078	0.077
CFI	0.630	0.726	0.616
Criterion validity			
ρ(EQ-5D-3L)	-0.298	-0.288	-0.276
ρ(EQ-VAS)	-0.288	-0.269	-0.273

a. Calculations based on a total of 1,350 observations without missing values on all the five EQ-5D-3L questions and all the 46 DQOL items.

CTT Classical test theory, *IRT* Item response theory, *DQOL* Diabetes quality-of-life measure, *SRMR* Standardized root mean squared residual, *CFI* Comparative fit index, ρ Spearman's correlation coefficient, *VAS* Visual analogue scale.

Appendix

Appendix 3.1Original Chinese DQOL and short versions based on the CTT and IRT

	Original Chinese DQOL	*Short version based on the CTT (32 items)*	*Short version based on the IRT (24 items)*
	请您对您以下各个方面(1-15)的主观满意程度进行评价： *（满意度：1.非常满意 2.满意 3.一般 4.不满意 5.非常不满意）* *Satisfaction (1 Very satisfied, 2 satisfied, 3 General, 4 Dissatisfied, 5 Very dissatisfied)*		
1	您对医生控制您的病情所花的时间满意吗？ How satisfied are you with the amount of time it takes to manage your diabetes?	√	√
2	您对常规的体格检查所花的时间满意吗？ How satisfied are you with the amount of time you spend getting checkups?	√	√
3	您对医生确定您的血糖水平所花的时间满意吗？ How satisfied are you with the time it takes to determine your sugar level?	√	√
4	您对目前接受的治疗措施满意吗？ How satisfied are you with your current treatment?	√	√
5	您对自己受限制的饮食满意吗？ How satisfied are you with the flexibility you have in your diet?	√	×
6	您对自己患糖尿病后给家庭带来的经济负担满意吗？ How satisfied are you with the burden your diabetes is placing on your family?	√	√
7	您对自己关于糖尿病知识的了解程度满意吗？ How satisfied are you with your knowledge about your diabetes?	×	×

8	您对自己的睡眠状况满意吗？ How satisfied are you with your sleep?	√	×
9	您对自己的社会关系和得到的友爱满意吗？ How satisfied are you with your social relationships and friendships?	√	√
10	您对自己的性生活满意吗？ How satisfied are you with your sex life?	×	×
11	您对自己的工作、学业和家庭生活满意吗？ How satisfied are you with your work, school, and household activities?	√	√
12	您对自己的体型满意吗？ How satisfied are you with the appearance of your body?	√	×
13	您对自己每天能够用于锻炼身体的时间满意吗？ How satisfied are you with the time you spend on exercising?	√	×
14	您对自己的业余生活满意吗？ How satisfied are you with your leisure time?	√	√
15	总的来说，您对自己的生活感到满意吗？ How satisfied are you with life in general?	√	√
请您对糖尿病给您以下各方面(16-35)带来的影响进行评价： *（影响程度：1.从来没有　2.很少有　3.偶尔有　4.经常有　5.一直有）* *Impact (1 Never, 2 Rarely, 3Occasional, 4 Often, 5 Always)*			
16	您患糖尿病后对经常不得不接受治疗感到痛苦吗？ How often do you feel pain associated with the treatment for your diabetes?	√	√
17	您经常对在公共场合下不得不谈及您的病情而感到尴尬吗？ How often are you embarrassed by having to deal with your diabetes in public?	√	√
18	您经常有心慌、出虚汗、头昏、颤抖等低血糖反应吗？ How often do you have low blood sugar*(e.g. palpitation, sweating, dizziness,*	√	×

	trembling)? [a]		
19	您经常感到身体不舒服吗？ How often do you feel physically ill?	√	√
20	您经常觉得自己患糖尿病给您的家庭生活带来麻烦吗？ How often does your diabetes interfere with your family life?	√	√
21	您经常晚上睡眠不好吗？ How often do you have a bad night's sleep?	√	×
22	您经常感到糖尿病限制了您的社会交往和友谊吗？ How often do you find your diabetes limiting your social relationships and friendships?	√	√
23	您经常自我感觉良好吗？ How often do you feel good about yourself?	×	×
24	您经常感到自己的饮食受到限制吗？ How often do you feel restricted by your diet?	√	×
25	您患糖尿病后性生活经常受到影响吗？ How often does your diabetes interfere with your sex life?	×	×
26	您患糖尿病后经常被人劝阻不要骑车或从事打字员之类的工作吗？ How often does your diabetes keep you from *riding a bike* or being a typist? [b]	×	×
27	您患糖尿病后身体锻炼经常受到影响吗？ How often does your diabetes interfere with your exercising?	√	√
28	您患糖尿病后经常无力承担家庭义务吗？ How often do you miss household duties because of your diabetes? [c]	√	√
29	您经常向别人解释糖尿病的危害吗？ How often do you find yourself explaining what it means to have diabetes?	×	×
30	您患糖尿病后业余活动经常受到影响吗？	√	√

	How often do you find that your diabetes interrupts your leisure-time activities?		
31	您患糖尿病后经常向别人诉说自己的病情吗？ How often do you tell others about your diabetes?	×	×
32	您患糖尿病后经常被别人取笑吗？ How often are you teased because you have diabetes?	×	×
33	您患糖尿病后经常感觉自己去洗手间的次数比别人多吗？ How often do you feel that because of your diabetes you go to the bathroom more than others?	√	×
34	经常发现自己隐瞒病情而去吃一些自己不应该吃的东西吗？ How often do you find that you eat something you shouldn't rather than tell someone that you have diabetes?	×	×
35	您经常隐瞒自己一直有胰岛素副反应的事实吗？ How often do you hide from others the fact that you are having an insulin reaction?	×	×
请您对以下方面(36-46)的忧虑程度进行评价： *（忧虑程度 I：1.从不担心　2.很少担心　3.偶尔担心　4.经常担心　5.总是担心）* *Worry (1 Never, 2 Rarely, 3Occasional, 4 Often, 5 Always)*			
36	您患糖尿病后经常为将来的婚姻状况感到忧虑吗？ How often do you worry about your *marriage*? [d]	×	√
37	您患糖尿病后经常为孩子的将来感到忧虑吗？ How often do you worry about your *children's future*? [e]	√	√
38	您患糖尿病后经常为以后可能找不到理想的工作感到忧虑吗？ How often do you worry about whether you will not get a job you want?	×	×
39	您患糖尿病后经常为以后可能得不到养老金或离退休金感到忧虑吗？ How often do you worry about whether you will be denied *pension*? [f]	√	×
40	您患糖尿病后经常为以后能否完成自己的继续教育感到忧虑吗？	×	×

	How often do you worry about whether you will be able to complete your education?		
41	您患糖尿病后经常为将来可能会失业感到忧虑吗？ How often do you worry about whether you will lose your job? [g]	×	×
42	您患糖尿病后经常为将来可能不能外出旅游感到忧虑吗？ How often do you worry about whether you will be able to take avacation or a trip?	√	√
43	您患糖尿病后经常为将来可能会昏厥感到忧虑吗？ How often do you worry about whether you will pass out?	√	√
44	您患糖尿病后经常为自己的体型与别人不同感到忧虑吗？ How often do you worry that your body looks different because you have diabetes?	√	√
45	您患糖尿病后经常为自己可能会发生并发症感到忧虑吗？ How often do you worry that you will get complications from your diabetes?	√	√
46	您患糖尿病后经常为有人不愿意和您一起外出感到忧虑吗？ How often do you worry about whether someone will not go out with you because you have diabetes?	√	√

English translation is provided after each Chinese item.

"√" indicates the item was kept; "×" indicates the item was removed from the scale.

a. The original English version was "How often do you have low blood sugar?"

b. The original English version was "How often does your diabetes keep you from driving a car or using a machine (e.g., a typewriter)?"

c. The original English version was "How often do you miss work, school, or household duties because of your diabetes?"

d. The original English version was "How often do you worry get married?"

e. The original English version was "How often do you worry have children?"

f. The original English version was "How often do you worry be denied insurance?"

g. The original English version was "How often do you worry about whether you will miss work?"

DQOL Diabetes quality-of-life, *CTT* Classical test theory, *IRT* Item response theory.

Chapter 4

Mapping the Chinese Version of the Diabetes Quality of Life (DQOL) measure onto the EQ-5D-3L index

Xuejing Jin, MSc [1, 2], Gordon G. Liu, PhD [2, 3], Hertzel C. Gerstein, M.D., MSc [4],Mitchell A.H. Levine, M.D., MSc [1], Haijing Guan, MSc [2, 5], Hongchao Li, MSc [6], Feng Xie, PhD [*1]

1. Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada
2. China Center for Health Economic Research, Peking University, Beijing, China
3. National School of Development, Peking University, Beijing, China
4. Department of Medicine, McMaster University, Hamilton, Ontario, Canada
5. School of Pharmaceutical Science, Peking University, Beijing, China
6. School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, China

* Corresponding author

Abstract

Objective: To map the Chinese Version of the Diabetes Quality of Life (DQOL) and the short version of DQOL onto the EuroQol 5 dimensions 3-level (EQ-5D-3L) in Chinese type 2 diabetic patients.

Methods: We used data from a community-based longitudinal observation study of Chinese type 2 diabetic patients to establish the mapping models. The baseline data was employed as the estimation sample, and the data from the one-year end follow-up was used as the testing sample. Models with the ordinary least squares (OLS), generalized linear models (GLM) and censored least absolute deviations (CLAD) were fitted. Four types of model specifications were explored for the DQOL and the short version of DQOL, respectively: 1) domain scores; 2) domain scores plus pair-wise interaction terms; 3) domain scores plus age and gender; and 4) domain scores plus pair-wise interaction terms plus age and gender. Backward stepwise selection was applied to the OLS and GLM models. Mapping model selection criteria included two aspects: 1) the association between the EQ-5D-3L scores and the domain scores should be negative, and 2) model performance evaluated in terms of mean absolute error

(MAE), root mean square error (RMSE), and the individual level prediction precision.

Results: For the DQOL, the preferred mapping algorithm was estimated using the GLM estimator and included the life satisfaction domain score, diabetes impact domain score, age, and gender (MAE=0.0977, RMSE=0.1276).For the short version of DQOL, the preferred mapping algorithm was estimated using the OLS estimator and included the score of a domain contained 20 items, age, and gender (MAE=0.0981, RMSE=0.1269). The preferred mapping algorithm for the short version of DQOL had better individual level prediction precision than that for the DQOL.

Conclusion: The mapping algorithm for the DQOL can be used to predict the EQ-5D-3Lscores in Chinese community-based type 2 diabetic patients in the absence of preference-based measures in those existing studies which reported DQOL scores at the domain level. When DQOL item level scores are accessible, the mapping algorithm for the short version of DQOL is recommended.

Keywords: EQ-5D, DQOL, Mapping, Diabetes

4.1 Background

Globally, the prevalence of diabetes has increased rapidly in the recent decades. According to the World Health Organization (WHO), the global prevalence rate of diabetes in the adult population had risen from 4.7% in 1980 to 8.5% (422 million) in 2014 [1]. In the recent decade, the prevalence of diabetes in low- and middle-income countries has increased faster than that in high-income countries [2]. In China, diabetes prevalence has increased from 1% [3] in the 1980s to 10.6% (109 million) in 2015 [4]. When uncontrolled, diabetes and its complications have an adverse impact on the patients' physical and mental well-being and can lead to disability and premature death [1, 7]. As a patient-reported outcome (PRO), quality of life (QoL) is an important outcome measure which can evaluate the effects of interventions on outcomes important to patients [8]. Thus, WHO suggests physicians and other health care providers should better monitor the QoL of diabetic patients and assess the impact of interventions on patients' QoL [1]. The Diabetes Quality-of-Life (DQOL) measure is one of the most commonly used diabetes-specific instruments. It has been translated into Chinese (Mandarin, Simplified) [9, 10], validated among the Chinese diabetic population [11], and is widely used in diabetes clinical research in China [12-14]. A shortened version of the Chinese DQOL based on the item response theory (IRT) (referred to as “the short version of DQOL”) has been provided in the second chapter.

As one of the four most prominent chronic diseases [5], diabetes also loads a large economic burden on the health-care system: the estimated direct annual cost of diabetes is more than US$ 827 billion globally [1]. In China, 13% of national medical expenditures are attributed to by diabetes, and the annual cost of diabetes treatment is more than US$ 25 billion [6].

Due to limited health resources and the significant economic and humanistic burden caused by diabetes, the results of economic evaluations have been used to guide diabetes-related medical decision-making pertaining to health resource allocation [15-17]. Among the five types of economic evaluation analysis (i.e., cost-effectiveness analysis, cost-utility analysis (CUA), cost-benefit analysis, cost-minimization analysis, and cost-consequences analysis), CUA is the analytical technique that is preferred by most countries’ economic evaluation guidelines [18]. This is because the outcome employed in CUA, i.e., quality-adjusted life-year (QALY), is a generic outcome measure including both quality and quantity of life, and it can achieve a broader economical comparison of interventions across different disease areas. Utility and value are the indexes of quality of life included in the QALY calculation, which are defined as patients' preferences toward health states that are measured under uncertainty and certainty, respectively. Utilities (or values) range from negative values (representing health states which are perceived as worse than death) to 1 (full health), with 0 representing death [19]. Utilities and values can be measured using direct approaches like standard gamble and time trade-off (TTO); however, considering the complexity of direct approaches, indirect approaches, i.e., preference-based instruments, are more commonly used in clinical research. Researchers can use a value set (algorithm) that is derived from direct approaches to convert responses from generic preference-based instruments, such as EuroQol 5 dimensions (EQ-5D)and SF-6D, into utility values.

In clinical trials, disease specific instruments are the other type of QoL measures commonly used, since they are more sensitive to disease-related symptoms and disease-related impacts on quality of life than generic instruments. However, since most disease specific instruments are not preference-based, their results cannot be converted into utilities or values directly, and cannot be employed in the CUA. There is a substantial amount of high-quality clinical studies reporting QoL as measured by only disease-specific instruments. Unfortunately, these results cannot contribute to a CUA.

Mapping disease-specific instruments onto the generic preference-based instrument is one way to link the studies only reporting disease-specific instrument measured QoL to the CUA. Currently, researchers have linked several disease-specific instruments (such as the EORTC core quality of life questionnaire [20, 21], the Western Ontario and McMaster Universities Osteoarthritis Index [22], and Parkinson's Disease Questionnaire -39 [23]) to the generic preference-based measures such as EQ-5D and Health Utility Index. To our knowledge, there are very few studies that have mapped diabetes-specific instrument onto generic preference-based instruments. Chen et al mapped the Diabetes-39 onto the EuroQol 5 dimensions 3-level (EQ-5D-3L) among patient populations from 6 countries which did not include China [24]. In light of this limitation, the objective of this study is to develop mapping algorithms to estimateEQ-5D-3L scores from the responses to the DQOL and the short version of DQOL for Chinese diabetic patients.

4.2 Methods

4.2.1 Sample and data

A convenience sample of type 2 diabetic patients was recruited from 66 community health centers in five Chinese cities between December 2010 and

October 2011 and was followed-up with for one year. The DQOL and the EQ-5D-3L were administrated at baseline and the end of 12 months in face-to-face interviews facilitated by a trained interviewer. The details of the survey deign and the patient inclusion and exclusion criteria were introduced in Chapter 1.

The baseline data was employed as the estimation data, and the data from the one-year end follow-up was used to evaluate the model performance. Individual mean imputation, which is one of the most commonly used imputation methods for multi-item quality of life measures [25], was applied to deal with missing data in the DQOL at the domain level. The imputed value for items with missing responses in a certain domain of the DQOL is the mean of the given individual's completed responses to the remaining items in that domain [26]. Previous studies suggest that the individual mean imputation method can provide reliable imputation results when the missing rate is not larger than 20% [26]. Therefore, observations were excluded if they: 1) had more than 20% of the values for any domain of the DQOL missing; and 2) had a missing value on any (one or more) of the five dimensions of the EQ-5D.

4.2.2 Instruments

4.2.2.1 DQOL

The Chinese version of the DQOL, a 46-item diabetes-specific quality of life measure, consists of four domains, i.e., life satisfaction (15 items), diabetes impact (20 items), social/vocational related worries (7 items), and diabetes related worries (4 items). Each of the 46 items is graded on a five-point Likert scale and scored from 1 to 5 (i.e., from "very satisfied"/ "never impacted"/ "never worried" to "very dissatisfied" / "always impacted" / "always worried") [9, 10]. Scores of negative-keyed items were reversed. Domain scores are calculated as the mean score of all items and ranged from 1 to 5. A higher DQOL mean score or domain scores indicate a worse QoL.

4.2.2.2 Short version of the DQOL

In Chapter 3, the 24-item short version of the Chinese DQOL was derived through item reduction based on the IRT, and appears to be a promising alternative to the Chinese DQOL. It consists of two domains, namely, one treatment-related satisfied domain (referred to as "domain-T") which contains 4 items (item #1, #2, #3, and #4), and another domain (referred to as "domain-R") which contains the remaining 20 items (item #6, #9, #11, #14, #15, #16, #17,

#19, #20, #22, #27, #28, #30, #36, #37, #42, #43, #44, #45, and #46, see Appendix 3.1). Mean scores were employed as the domain scores as well, and a higher domain score indicates a worse QoL.

4.2.2.3 EQ-5D-3L

The EQ-5D-3L is a generic preference-based quality of life measure. It contains five dimensions including mobility, self-care, usual activity, pain/discomfort, and anxiety/depression, where each dimension has three response levels, i.e., no problem, some problem, and extreme problems [27]. It has been translated into Chinese (Mandarin, Simplified) [28] and validated in a Chinese patient population with chronic diseases which included diabetes [29]. The responses to the five EQ-5D-3L questions were converted into utility scores using the Chinese EQ-5D-3L value set which was derived from the TTO method [30]. The Chinese EQ-5D-3L scores range from -0.149 to 1, and higher utility scores indicate better quality of life.

4.2.3 Statistical Analysis

4.2.3.1 Explanatory variables

For the DQOL, the following four types of model specification were explored in our regression analysis: 1) DQOL domain scores (i.e., life satisfaction, diabetes impact, social/vocational related worries, and diabetes related worries); 2) DQOL domain scores plus pair-wise interaction terms (i.e., life satisfaction*life satisfaction, life satisfaction* diabetes impact, life satisfaction* social/vocational related worries, life satisfaction* diabetes related worries, diabetes impact* diabetes impact, diabetes impact* social/vocational related worries, diabetes impact* diabetes related worries, social/vocational related worries * social/vocational related worries, social/vocational related worries * diabetes related worries, diabetes related worries* diabetes related worries); 3) DQOL domain scores plus age and gender; and 4) DQOL domain scores plus pair-wise interaction terms plus age and gender. In line with existing studies [22], we included squared DQOL domain scores in the set of pair-wise interaction terms, which allowed us to explore whether the relationship between the individual DQOL domain score and the EQ-5D-3L score index was quadratic.

Similar model specifications were explored for the short version of DQOL: 1) short version of DQOL domain scores (i.e., domain-T and domain-R); 2) short version of DQOL domain scores plus pair-wise interaction terms (i.e., domain-T*domain-R, domain-T*domain-T, and domain-R*domain-R); 3) short version of DQOL domain scores plus age and gender; and 4)short version of

DQOL domain scores plus pair-wise interaction terms plus age and gender.

4.2.3.2 Model estimators and dependent variables

We used the ordinary least squares (OLS), generalized linear models (GLM) and censored least absolute deviations (CLAD) estimators to estimate the models. The dependent variable of the OLS models was the EQ-5D-3L score. For the GLM, the Gamma distribution was assumed for the dependent variable which was defined as 1 minus the EQ-5D-3L score plus 0.0001, and the log link function was applied. For the CLAD models, the dependent variable was 1 minus the EQ-5D-3L score. For models estimated by the OLS and GLM estimators, backward stepwise selection with a variable removal criterion of $p \geq 0.05$ was used to decide which variables were included in each model. Stepwise selection did not apply to CLAD models since the CLAD is not supported by stepwise.

Therefore, 12 models (i.e., 4 types of model specifications multiplied by 3types of estimators) were estimated for both the DQOL and the short version of DQOL.

4.2.3.3 Model selection criteria

A two-stage strategy was used to select the final mapping models. First, the DQOL domain scores and the short version of DQOL domain scores should be negatively associated with the EQ-5D-3L scores. Thus, the coefficients of the DQOL (or the short version of DQOL) domain scores should be negative in the OLS models and positive in the GLM models, respectively. In the CLAD models, the coefficients of domain scores which were statistically significant should be positive. Any model that did not meet the above requirements was no longer considered eligible for the mapping algorithm.

Second, model performance in terms of the mean absolute error (MAE), root mean square error (RMSE), and the individual level prediction precision of all candidate models kept in the first stage was evaluated using the testing sample. The MAE is the mean of absolute difference between predicted and observed values [31]. The RMSE is the square root of the mean of squared differences between predicted and observed values. While independent of each other, lower MAE and RMSE values indicate better model performance [31]. The RMSE attaches a relatively higher weight to large errors than the MAE, since the errors are squared before being averaged [31]. The individual level prediction precision was calculated as the absolute difference between the predicted and observed utility scores, i.e., the prediction error, for each individual in the testing sample. The minimally important difference of the EQ-5D scores was 0.026 among Chinese type 2 diabetic patients (see Chapter 1), was reported as 0.058 for diabetic patients in UK [32], and was estimated as a range of -0.011 to 0.139 in patients with different kinds of chronic diseases [33]. Therefore, we categorized the absolute differences into one of the following four groups, 0 to 0.026, 0.026 to 0.058, 0.058 to 0.139 and larger than 0.139.The Wilcoxon Rank Sum test was applied to compare the individual level prediction precision of different candidate mapping models. In this stage, the MAE and RMSE were applied as primary criteria for model performance assessment; in the event of any inconsistency between the MAE and RMSE results, individual prediction precision was employed to help to decide the final mapping models.

Since using data from existing studies to map the short version of DQOL onto the EQ-5D-3L requests the DQOL item scores; in the event of DQOL item-level scores are not accessible to the users of our potential mapping algorithm, we selected preferred mapping algorithms for the Chinese DQOL and the short version of DQOL.

All statistical analyses were performed using STATA 14.0 (StataCorp LP, Texas, USA) with two-tailed tests at a significance level of 0.05.

4.3 Results

4.3.1 Descriptive summary of the estimation and testing samples

A total of 2,763 baseline observations were included in estimating the mapping models, and 2,293 observations from the last follow-up were included in the testing sample to test the model performance. Table 4.1 shows the baseline age and gender of the two samples, respectively, and the DQOL mean score, DQOL domain scores, short version of DQOL mean score, and short version of DQOL domain scores for the two samples. The mean (standard deviation, SD) age of the estimation sample was 62.01 (11.47) years with 55.34% being female. There is no statistically significant difference between the estimation sample and the testing sample in terms of age and gender. The mean (SD) EQ-5D-3L scores estimation and testing samples were 0.889 (0.139) and 0.915 (0.133), respectively. The mean (SD) DQOL scores of the training sample and the testing sample were 2.068 (0.386) and 2.015 (0.337), respectively, and the DQOL short version scores for the estimation and testing samples were 2.020 (0.512) and 1.998 (0.404), respectively. The testing sample had lower scores than the estimation sample in all DQOL domains, except for the DQOL diabetes related worries domain.

4.3.2 Model selection

In Table 4.2, models estimated negative correlations between EQ-5D-3L score and domain scores are marked with an asterisk (*), which means these models met the first stage selection criteria. Among all the 12 models estimated for the DQOL, only 2 models have negatively estimated correlations between EQ-5D-3L score and DQOL domain scores. Appendix 4.1 summarizes the coefficient estimates of these two models. Table 4.2 also shows that among all the 12 models estimated for the DQOL short version, 8 models have negatively estimated correlations betweenEQ-5D-3L score and the domain scores (marked with a * in Table 4.2). Appendix 4.2- 4.4 show the coefficient estimates of these 8 models.

Model performance in terms of MAE and RMSE of each mapping model fitted by the testing sample are also showed in Table 4.2. For the original Chinese DQOL, the GLM model created by stepwise procedure using DQOL domain scores plus age and gender (referred to as "DQOL_GLM model") had both the lower MAE (0.0977) and RMSE (0.1276) compared to the OLS model created by stepwise procedure using the same set of explanatory variables (MAE=0.0979, RMSE=0.1282). Therefore, the DQOL_GLM model was selected as the preferred mapping model for the Chinese DQOL.

For the short version of Chinese DQOL, among the 8 models met the first stage selection criteria, the CLAD model (MAE=0.0964, RSME=0.1359) using the domain scores of the short version of DQOL as dependent variables (referred to as "short_CLAD model") had the lowest MAE, and the OLS model (MAE=0.0981, RMSE=0.1269)created by stepwise procedure using domain scores of the short version of DQOL plus age and gender as dependent variables (referred to as "short_OLS model") had the lowest RMSE. Since the MAE and RMSE results were not consistent, we compared the individual level prediction precision of these two candidate models (as shown in Table 4.3). The short_OLS model had about 18.75% and 40.16% of the prediction errors smaller than 0.026 and 0.058, respectively, and 21.76% of the prediction errors larger than 0.139.The short_CLAD model had 13.26% and 31.49% of the prediction errors smaller than 0.026 and 0.058, respectively, and 20.41% of the prediction errors larger than 0.139. The Wilcoxon Rank Sum test (p-value <0.001) of the absolute difference between the two models showed that compared to the short_CLAD model, the short_OLS model using the domain scores plus age and gender as dependent variables had better individual level prediction precision (i.e., a statistically significantly higher proportion of small prediction errors). Thus the short_OLS model was selected as the preferred mapping model for the short version of DQOL.

In addition, compared to the DQOL_GLM model, the short_OLS model had lower RMSE while higher MAE (Table 4.2); while, the Wilcoxon Rank Sum test (p-value <0.001) showed that the short_OLS model had better individual level prediction precision than the DQOL_GLM model (Table 4.3). Thus, our recommendation is to use the short_OLS model when all item-level data are available.

4.3.3 Mapping algorithms

The following two equations show the final mapping models for the original DQOL and short version of DQOL, respectively:

U_{DQOL}= 1- exp (-6.6081+0.0245*Age-0.2548*Gender+0.4962*$DQOL_{lifesatisfaction}$score +0.7838*$DQOL_{diabetes\ impact}$ score) +0.0001

U_{short_DQOL}=1.2544-0.0024*Age+0.0206*Gender-0.1116*Domain-R score where, U_{DQOL} andU_{short_DQOL} are the predicted EQ-5D scores derived from the final mapping models for the original DQOL and short version of DQOL, respectively, Age is a continuous variable, Gender is a dummy variable coded as 0 for female and 1 for male, $DQOL_{lifesatisfaction}$ score and $DQOL_{diabetes\ impact}$ score are the domain scores of the life satisfaction and diabetes impact domains of the DQOL, respectively, and domain-R score is the domain score of the short version of the DQOL which contains 20 items (see Appendix 3.1).

For example, if a 51-year old female diabetic patient has the $DQOL_{lifesatisfaction}$ score, $DQOL_{diabetes\ impact}$ score, and the domain-R scores of 2, 3 and 2.3, respectively, the algorithm for the DQOL will estimate an EQ-5D-3L score of 0.8668, and the algorithm for the short version of DQOL will estimate an EQ-5D-3L score of 0.8959.

Theoretically, for the mapping algorithm for the DQOL, the predicted EQ-5D-3L scores can range from-4.7240for an 80-year old female who has $DQOL_{lifesatisfaction}$ and $DQOL_{diabetes\ impact\ t}$scores both of 5 to 0.9939 for a 20-year old male who has $DQOL_{lifesatisfaction}$ and $DQOL_{diabetes\ impact}$ scores both of 1. Since the Chinese EQ-5D-3L value set ranges from -0.149 to 1, in practice, any predicted value lower than -0.149 needs to be truncated at -0.149.

Similarly, for the mapping algorithm for the short version of DQOL, the predicted utility scores can theoretically range from 0.5044 for an 80-year old female who has domain-R score of 5 to 1.1154 for a 20-year old male who has domain-R score of 1. However, the utility score cannot exceed 1. Therefore, in practice, any predicted value larger than 1 needs to be truncated at 1.

Table 4.4 compares the EQ-5D-3L scores predicted by the two mapping algorithms. The mean and median of predicted utility scores were 0.899 and 0.914, respectively, for the DQOL_GLM model, and were 0.897 and 0.899, respectively, for the short_OLS model. Means and Medians of predicted utility scores of the two mapping algorithms were both smaller than those of the observed EQ-5D-3L scores.

4.4 Discussion

The present study aimed to develop mapping algorithms to map the diabetes-specific instrument, DQOL (and its short version), onto the generic preference-based instrument, EQ-5D-3L. Using community-based survey data of Chinese type 2 diabetic patients, we estimated and compared models with different specifications and estimators. The GLM model using age, gender, and DQOL life satisfaction and DQOL diabetes impact domain scores as dependent variables was selected as final mapping model to predict theEQ-5D-3L derived utility scores in the absence of preference-based measures in those existing studies which reported DQOL scores at the domain level. When DQOL item level scores are accessible, we recommend to use the OLS model using age, gender, and the domain score of the domain-R of the short version of DQOL as dependent variables to predict the EQ-5D-3L derived utility scores; since this model had better individual level prediction precision than the DQOL_GLM model.

To our knowledge, this is the first study that maps the DQOL measure onto the EQ-5D. Compared to results of existing studies that mapped disease-specific measures onto EQ-5D-3L scores in diabetes and other chronic disease areas, our final mapping models were acceptable in terms of MAE and RMSE. Chen et al. [24] mapped the Diabetes-19 onto the EQ-5D-3L using data from a multi-country diabetic patient sample, and the MAE and RMSE of the final mapping model (OLS model) that estimated using an internal testing sample were 0.131 and 0.177, respectively. In other chronic disease areas, the MAE estimated using the internal and external testing samples ranged from 0.0736 to 0.1662 [22, 34-36] and from 0.0933 to 0.1640 [23, 34], respectively, and the RMSE estimated using the internal and external testing samples ranged from

0.0947 to 0.2221 [22,34-36] and from 0.1107 to 0.2235 [23, 34], respectively.

In our study, age and gender were retained in the final mapping model through the backward stepwise selection process. Compared to who were young and male, elderly and female patients had lower predicted EQ-5D-3L scores. This finding is in line with the findings in existing mapping studies [34, 35, 37, 38] and the recommendation in Good Practices of mapping studies provided by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) [39]. Age and gender can provide additional information that cannot be captured by the DQOL and can improve the mapping models' fit [39]. In the practice of applying the mapping algorithm in the CUAs, including age and gender can make the mapping algorithm more sensitive, especially when the CUA targets on a certain age or gender population.

Another notable aspect is that not all domain scores were kept in both final mapping models for the DQOL and the short version of DQOL. This indicates the conceptual differences between the DQOL and EQ-5D-3L descriptive system. For example, the DQOL has items about satisfaction with the patients' accepted treatment, sleep quality, sexual activities and so on, which cannot be captured by the EQ-5D. Mapping both the DQOL (or its short version) and the

EQ-5D onto a common yardstick may help to solve this issue [40]; however, our data did not support this type of analysis because we did not use a common yardstick in the survey. In addition, previous studies show that 15-Dimension measure (15-D), which is a preference-based generic instrument contains 15 dimensions [41], is more sensitive than the EQ-5D-3L in measuring diabetic patients' QoL [24, 42]. Therefore, exploring the mapping algorithm between the DQOL (or the short version of DQOL) and 15-D in a future study is necessary.

The present mapping study has several limitations. First, limited by the log transform of the dependent variable and linear prediction of the GLM using log link function, the mapping algorithm (i.e., the DQOL_GLM model) for the DQOL underestimated the utilities for elderly patients who reported severe health states (i.e., have relatively high DQOL domain scores). Limited by the linear prediction of the OLS model, the mapping model (i.e., the short_OLS model) for the short version of DQOL overestimated the utilities for young patients with very mild health states (i.e., have relatively low short version of DQOL domain scores).

Second, our mapping models were established based on a sample of community-based type 2 diabetic patients from Chinese urban areas, which limited the generalizability of our results. In addition, from the summary of the baseline characteristics reported in Chapter 2, we can see that there was a low proportion of patients in our training sample with serious complications and using insulin. Thus, the mapping results can only be generalized to similar populations.

Finally, similar to Chapter 3, we used baseline data as the estimation sample to develop the mapping models and used the one-year follow-up data as the testing sample to assess model performance. These two samples were from the same population and shared similar demographics. We did not apply the commonly used split-sample approach (i.e. use the baseline data of a sub-sample of patients who were randomly chosen from all individuals for model estimation and use the final follow-up data of the remained individuals for model performance evaluation) in this study mainly because the largest number of explanatory variables across different model specifications was 16, and using the largest sample size for model estimation can achieve higher statistical power. In addition, this study mainly employed patients' responses to the Chinese DQOL and EQ-5D-3L measures to estimate and test mapping models. Table 4.1 shows that the estimation sample and the testing sample were statistically significantly different in terms of Chinese DQOL domain scores and EQ-5D-3L index scores, which decreased the risk of bias of using the own sample for model performance evaluation. Re-sampling methods, such as cross-validation and bootstrapping are also commonly used when an external population is not available; however, these methods could only maximize squeeze information from the same population. Therefore, to better validate the mapping algorithms, further external validation is needed.

References

[1] World Health Organization. (2016). Global Report on Diabetes. From: http://apps.who.int/iris/bitstream/10665/204871/1/9789241565257_eng.pdf (Accessed on January 25, 2017).

[2]NCD Risk Factor Collaboration (NCD-RisC). (2016).Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 387: 1513–1530.

[3]Chan M. (2017). China's Burgeoning Epidemicof Diabetes-Associated Mortality. JAMA.317(3): 264-266.

[4] International Diabetes Federation. Diabetes in China-2015. From: http://www.idf.org/membership/wp/china. (Accessed on January 25, 2017).

[5] World Health Organization. Integrated chronic disease prevention and control. From: http://www.who.int/chp/about/integrated_cd/en/. (Accessed on January 25, 2017).

[6] International Diabetes Federation. China Spends RMB 173.4 Billion (US$ 25 Billion) A Year on Diabetes Treatment. From: http://www.idf.org/china-spends-rmb-1734-billion-us25-billion-year-diabetes-treatment. (Accessed on January 25, 2017).

[7] Gerstein HC, Werstuck GH. (2013). Dysglycaemia, vasculopenia, and the chronic consequences of diabetes.Lancet Diabetes Endocrinol. 1(1):71-78.

[8] Diapedia. Patient reported outcomes among patients with diabetes and cancer: an overview and research agenda. From: http://www.diapedia.org/associated-disorders/6104721115/patient-reported-outcomes-among-patients-with-diabetes-and-cancer-an-overview-and-research-agenda. (Accessed on January 25, 2017).

[9]PROQOLID. Diabetes Quality Of Life Measure (DQOL). From:https://eprovide.mapi-trust.org/instruments/diabetes-quality-of-life-measure#languages. (Accessed on December 18, 2016)

[10]Ding Y, Kong D, Ni Z, Deng H. (2004). Culture adaption and revision of diabetes-specific quality of life scale (QDOL). Chin J Behav Med Sci. 13(1): 102-103.

[11] Ding Y, Ni Z, Zhang J, Chen G, Feng H. (2000). The Assessment on Reliability and Validity of Adjusted Diabetes Quality of Life(A-DQOL) Scale.Chin J PrevContrChron Non-commun Dis. 8(4): 160-162.

[12]Liang M. (2014) Quality of life and its impact factors of community-based patients with type 2 diabetes in Beijing. Master degree thesis.Beijing university of Chinese medicine.

[13] Chen A, Su A, Bai H. (2009). Factors impact on the quality of life of type 2 diabetic patients. Chin Community Doctors. 11(19): 38.

[14] Hou Y, Yang Q. (2016). Current research status of health-related quality of life in patients with diabetes mellitus.Chin J Clinicians. 10 (3): 433-436

[15]Canadian Agency for Drugs and Technologies in Health. (2015). Insulin Pumps for Adults with Type 1 Diabetes: A Review of Clinical Effectiveness, Cost-effectiveness and Guidelines. From: https://www.cadth.ca/insulin-pumps-adults-type-1-diabetes-review-clinical-effectiveness-cost-effectiveness-and-guidelines. (Accessed on January 25, 2017).

[16] National Institute for Health and Care Excellence. (2015). Type 2 diabetes in adults: management. From: https://www.nice.org.uk/guidance/ng28/resources/type-2-diabetes-in-adults-management-1837338615493. (Accessed on January 25, 2017).

[17] Pharmaceutical Benefits Scheme. (2016). Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (PBAC). From: https://pbac.pbs.gov.au/. (Accessed on January 25, 2017).

[18] International Society for Pharmacoeconomics and Outcomes Research.Pharmacoeconomic Guidelines Around The World. From: https://www.ispor.org/PEguidelines/index.asp. (Accessed on January 25, 2017).

[19] Whitehead SJ and Ali S. (2010). Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 96 (1): 5-21.

[20]McKenzie L, van der Pol M. (2009). Mapping the EORTC QLQ C-30 onto the EQ-5D-3L instrument: the potential to estimate QALYs without generic preference data.Value Health. 12(1):167-171.

[21]Kharroubi SA, Edlin R, Meads D, Browne C, Brown J, McCabe C. (2015). Use of Bayesian Markov chain Monte Carlo methods to estimate EQ-5D-3L utility scores from EORTC QLQ data in myeloma for use in cost-effectiveness analysis.Med Decis Making. 35(3):351-360.

[22]Xie F, Pullenayegum EM, Li SC, Hopkins R, Thumboo J, Lo NN. (2010). Use of a Disease-Specific Instrument in Economic Evaluations:Mapping WOMAC onto the EQ-5D-3L Utility Index. Value Health.;13(8):873-878.

[23] Kent S, Gray A, Schlackow I, Jenkinson C, McIntosh E. (2015). Mapping from the Parkinson's Disease Questionnaire PDQ-39 to the Generic EuroQol EQ-5D-3L: The Value of Mixture Models.Med Decis Making. 35(7):902-911.

[24]Chen G, Iezzi A, McKie J, Khan MA, Richardson J. Diabetes and quality of life: Comparing results from utility instruments and Diabetes-39.Diabetes Res ClinPract. 109(2):326-333.

[25] Hardouin JB, ConroyR, Sébille V. (2011). Imputation by the mean score should be avoided when validating a Patient Reported Outcomes questionnaire by a Rasch model in presence of informative missing data. BMC Med Res Methodol.11:105-117.

[26] Shrive FM, Stuart H, Quan H, Ghali WA. (2006). Dealing with missing data in a multi-question depression scale: a comparison of imputation methods.BMC Med Res Methodol.6:57-66.

[27] EuroQol Group. How to use EQ-5D. From:http://www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html. (Accessed on February 1, 2017).

[28] EuroQol Group. EQ-5D-3L Self-complete version on paper. From: http://www.euroqol.org/eq-5d-products/eq-5d-3l/self-complete-version-on-paper

.html. (Accessed on February 1, 2017).

[29]Zhao FL, Yue M, Yang H, Wang T, Wu JH, Li SC. (2010). Validation and comparison of EuroQol and short form 6D in chronic prostatitis patients.Value Health. 13(5): 649–656.

[30] Liu GG, Wu H, Li M, Gao C, Luo N. (2014). Chinese time trade-off values for EQ-5D-3L health states.Value Health.17(5):597-604.

[31] Chai T and Draxler RR. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geosci. Model Dev. 7(3): 1247-1250.

[32] Mulhern B, Meadows K. (2013). Investigatingthe minimally important difference of the Diabetes Health Profile (DHP-18) and the EQ-5D-3L and SF-6D in a UK diabetes mellitus population.Health. 5(6):1045-1054.

[33] Tsiplova K, Pullenayegum E,Cooke T, Xie F. (2016). EQ-5D-derived health utilities and minimally importantdifferences for chronic health conditions: 2011 CommonwealthFund Survey of Sicker Adults in Canada. Qual Life Res. 25:3009–3016.

[34] Boland MR, van Boven JF, Kocks JW, van der Molen T, Goossens LM, Chavannes NH, Rutten-van Mölken MP. (2015). Mapping the clinical chronic obstructive pulmonary disease questionnaire onto generic preference-based EQ-5D-3L values. Value Health.18(2):299-307.

[35] Kay S and Ferreira A. (2014). Mapping the 25-item National Eye Institute Visual Functioning Questionnaire (NEI VFQ-25) to EQ-5D-3L utility scores.Ophthalmic Epidemiol. 21(2):66-78.

[36] Kay S, Tolley K, Colayco D, Khalaf K, Anderson P, Globe D. (2013). Mapping EQ-5D-3L utility scores from the Incontinence Quality of Life Questionnaire among patients with neurogenic and idiopathic overactive bladder.Value Health.16(2):394-402.

[37] Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. (2014). Mapping EQ-5D-3L utility scores from the PedsQL™ generic core scales.Pharmacoeconomics.32(7):693-706.

[38] Hernández Alava M, Wailoo A, Wolfe F, Michaud K. (2014). A comparison of direct and indirect methods for the estimation of health utilities from clinical outcomes.Med Decis Making. 34(7):919-930.

[39]Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, Botteman M, Busschbach J. (2017). Mapping to Estimate Health-State Utility fromNon–Preference-Based Outcome Measures: An ISPORGood Practices for Outcomes Research Task Force Report.Value Health.20(1):18-27.

[40]Rowen D, Brazier J, Tsuchiya A, Alava MH. (2012). Valuing states from multiple measures on the same visual analogue sale: a feasibility study.Health Econ. 21(6):715-729.

[41] 15-D Instrument. From: http://www.15d-instrument.net/15d/. (Accessed on January 25, 2017).

[42]Kontodimopoulos N, Pappa E, Chadjiapostolou Z,Arvanitaki E, Papadopoulos AA,Niakas D. (2012). Comparing thesensitivity of EQ-5D, SF-6D and 15D utilities to the specificeffect of diabetic complications.Eur J Health Econ.13:111–120.

Tables

Table 4.1 Age, gender, and quality of life scores for the estimation and testing samples.

	Estimation sample (N=2,763)	Testing sample (N=2,293)	P-value [a]
Age mean (SD), years	62.01 (11.47)	61.78 (11.26)	0.356
Female (%)	1529 (55.34)	1268 (55.35)	0.995
EQ-5D score mean (SD)	0.889 (0.139)	0.915 (0.133)	<0.001
Chinese DQOL score mean (SD)	2.068 (0.386)	2.015 (0.337)	<0.001
Life satisfaction domain score mean (SD)	2.378 (0.421)	2.300 (0.403)	0.026
Diabetes impact domain score mean (SD)	2.115 (0.479)	2.063 (0.438)	<0.001
Social/vocational related worries domain score mean (SD)	1.382 (0.491)	1.352 (0.436)	<0.001
Diabetes related worries domain score mean (SD)	1.886 (0.745)	1.903 (0.730)	0.327
Short version of DQOL score mean (SD)	2.050 (0.464)	1.998 (0.404)	<0.001
Domain-T [b] score mean (SD)	2.116 (0.559)	2.093 (0.562)	0.752
Domain-R [c] score mean (SD)	2.067 (0.512)	1.979 (0.440)	<0.001

a. Chi-square test for frequency, and one-way ANOVA for means;
b. Domain-T contains 4 treatment-related items;
c. Domain-R contains 20 items.
SD Standard deviation; *EQ-5D* Euroqol 5 dimensions; *DQOL* Diabetes quality-of-life.

Table 4.2MAE and RMSE of mapping models fitted using the testing sample

	Domain Scores		Domain scores plus interactions		Domain score plus age and gender		Domain scores plus interactions plus age and gender	
	Chinese DQOL	Short version	Chinese DQOL	Short version	DQOL	Short version	DQOL	Short version
OLS model								
MAE	0.1010 (0.0831)	0.1019 (0.0798)*	0.1003 (0.0821)	0.1018 (0.0790)	0.0979 (0.0828)*	0.0981 (0.0804)*	0.0969 (0.0825)	0.0975 (0.0802)
RMSE	0.1308 (0.2057)	0.1294 (0.2015)*	0.1296 (0.2043)	0.1289 (0.2038)	0.1282 (0.2031)*	***0.1269 (0.1981)****	0.1272 (0.2021)	0.1262 (0.2003)
GLM model								
MAE	0.1008 (0.0801)	0.1017 (0.0783)*	0.1010 (0.1356)	0.1020 (0.0783)	***0.0977 (0.0822)****	0.0984 (0.0804)*	0.1037 (0.2194)	0.0984 (0.0804)*
RMSE	0.1287 (0.2037)	0.1284 (0.2018)*	0.1708 (0.7632)	0.1286 (0.2001)	***0.1276 (0.2049)****	0.1271 (0.2011)*	0.2427 (1.3979)	0.1271 (0.2011)*
CLAD model								
MAE	0.1053 (0.1023)	***0.0964 (0.0959)****	0.1066 (0.1000)	0.1156 (0.1036)	0.1032 (0.0968)	0.0997 (0.0927)*	0.1021 (0.0974)	0.1042 (0.0962)*
RMSE	0.1468 (0.2309)	0.1359 (0.2210)*	0.1461 (0.2259)	0.1552 (0.2314)	0.1415 (0.2232)	0.1381 (0.2202)*	0.1411 (0.2235)	0.1418 (0.2205)*

The values presented are mean estimates (standard deviation).

* Indicates the associations between EQ-5D score and domain scores were negative.

DQOL Diabetes quality-of-life; *OLS* Ordinary least squares; *MAE* Mean absolute error; *RMSE* Root mean square error; *GLM* Generalized linear models; *CLAD* Censored least absolute deviations.

Table 4.3 Individual level prediction precision estimated using the testing sample			
Prediction error	Chinese DQOL Domain score plus age and gender GLM model [a] n (%)	Short version Domain Scores CLAD model [b] n (%)	Short version Domain score plus age and gender OLS model [a, b] n (%)
0≤|∆|≤0.026	157 (6.85)	304 (13.26)	430 (18.75)
0.026≤|∆|≤0.058	588 (25.64)	418 (18.23)	491 (21.41)
0.058≤|∆|≤0.139	1,113 (48.54)	1,103 (48.10)	873 (30.07)
|∆|>0.139	435 (18.97)	468 (20.41)	499 (21.76)

a. P –value< 0.001 for the Wilcoxon Rank Sum test between the two models.
b. P –value< 0.001 for the Wilcoxon Rank Sum test between the two models.
|∆| Absolute difference between the predicted and observed values.
MAE Mean absolute error; *RMSE* Root mean square error; *CLAD* Censored least absolute deviations; *OLS* Ordinary least squares; *DQOL* Diabetes quality-of-life, *GLM* Generalized linear model.

Table 4.4 Observed and predicted EQ-5D scores for the testing sample			
EQ-5D score	Observed	DQOL Domain score plus age and gender GLM model	Short version Domain score plus age and gender OLS model
Mean	0.915	0.899	0.897
Median	1	0.914	0.899
First and third quartiles	0.869, 1	0.877, 0.942	0.861, 0.936
SD	0.133	0.064	0.057

EQ-5D EuroQol 5 dimensions; *DQOL* Diabetes quality-of-life; *GLM* Generalized linear model; *OLS* Ordinary least squares; *SD* Standard deviation.

Appendices

Appendix 4.1 Regression results of the OLS and GLM models using DQOL domain scores plus age and gender as explanatory variables

	OLS model [a]		GLM model [b]	
Variable	Coefficient (SE)	P-value	Coefficient (SE)	P-value
$DQOL_{lifesatisfaction}$	-0.0529 (0.0066)	<0.0001	0.4961 (0.0651)	<0.0001
$DQOL_{diabetes\ impact}$	-0.0870 (0.0058)	<0.0001	0.7837 (0.0618)	<0.0001
Age	-0.0021 (0.0002)	<0.0001	0.0245 (0.0023)	<0.0001
Male	0.0205 (0.0048)	<0.0001	-0.2548 (0.0508)	<0.0001
Constant	1.3196 (0.0202)	<0.0001	-6.6074 (0.2247)	<0.0001

a. The observed EQ-5D scores was used as dependent variable, and backward stepwise selection was applied;

b. 1-EQ-5D score+0.0001 was used as the dependent variable, and Gamma distribution, log link function, and backward stepwise selection were applied.

OLS Ordinary least squares; *GLM* Generalized linear models; *DQOL* Diabetes quality-of-life; *SE* Standard error.

Appendix 4.2 Regression results of models using short version of DQOL domain scores as explanatory variables

	OLS model [a]		GLM model [b]		CLAD model [c]	
Variable	Coefficient (SE)	P-value	Coefficient (SE)	P-value	Coefficient (SE)	95% CI
Domain-T					0.0272 (0.0105)	(0.0064,0.0479)
Domain-R	-0.1072 (0.0047)	<0.0001	0.9014 (0.0484)	<0.0001	0.1660 (0.0146)	(0.1370, 0.1949)
Constant	1.1078 (0.0099)	<0.0001	-4.1483 (0.1015)	<0.0001	-0.3420 (0.0430)	(-0.4272, -0.2568)

a. The observed EQ-5D scores was used as dependent variable, backward stepwise selection was applied;

b. 1-EQ-5D score+0.0001 was used as the dependent variable, Gamma distribution, log link function and backward stepwise selection were applied;

c. 1-EQ-5D score was used as the dependent variable, full model was estimated by bootstrapping 100 times.

DQOL Diabetes quality-of-life; *OLS* Ordinary least squares; *GLM* Generalized linear models; *CLAD* Censored least absolute deviations; *SE* Standard error; *CI* Confidence interval, *Domain-T* Treatment satisfaction domain score; *Domain-R* Domain score of the domain contains the other 20 items of the short version of DQOL.

Appendix4.3 Regression results of models using short version of DQOL domain scores plus age and gender as explanatory variables

	OLS model [a]		GLM model [b]		CLAD model [c]	
Variable	Coefficient (SE)	P-value	Coefficient (SE)	P-value	Coefficient (SE)	95% CI
Domain-T					0.0123 (0.0081)	(-0.0038, 0.0284)
Domain-R	-0.1116 (0.0047)	<0.0001	0.9776 (0.0499)	<0.0001	0.1609 (0.0111)	(0.1389, 0.1828)
Age	-0.0024 (0.0002)	<0.0001	0.0255 (0.0022)	<0.0001	0.0032 (0.0001)	(0.0023, 0.0042)
Male	0.0206 (0.0048)	<0.0001	-0.2596 (0.0496)	<0.0001	-0.0465 (0.0102)	(-0.0668, -0.0262)
Constant	1.2544 (0.0173)	<0.0001	-5.8171 (0.1897)	<0.0001	-0.4711 (0.0398)	(-0.5500, -0.3922)

a. The observed EQ-5D scores was used as dependent variable, backward stepwise selection was applied;

b. 1-EQ-5D score+0.0001 was used as the dependent variable, Gamma distribution, log link function and backward stepwise selection were applied;

c. 1-EQ-5D score was used as the dependent variable, full model was estimated by bootstrapping 100 times.

DQOL Diabetes quality-of-life; *OLS* Ordinary least squares; *GLM* Generalized linear models; *CLAD* Censored least absolute deviations; *SE* Standard error; *CI* Confidence interval; *Domain-T* Treatment satisfaction domain score; *Domain-R* Domain score of the domain contains the other 20 items of the short version of DQOL.

Appendix Table 4.4 Regression results of the GLM and CLAD models using short version of DQOL domain scores plus interactions plus age and gender as explanatory variables

	GLM model [a]		CLAD model [b]	
Variable	Coefficient (SE)	P-value	Coefficient (SE)	95% CI
Domain-T			-0.0933 (0.0739)	(-0.2399, 0.0534)
Domain-R	0.9776 (0.0499)	<0.0001	0.2188 (0.1057)	(0.0091, 0.4285)
Domain-T ^{2}			0.0049 (0.0113)	(-0.0175, 0.0273)
Domain-T* Domain-R			0.0354 (0.0249)	(-0.0139,0.0848)
Domain-R ^{2}			-0.0296 (0.0230)	(-0.0753, 0.0160)
Age	0.0255 (0.0022)	<0.0001	0.0032 (0.0006)	(0.0020, 0.0044)
Male	-0.2596 (0.0496)	<0.0001	-0.0482 (0.0118)	(-0.0716, -0.0248)
Constant	-5.8171 (0.1897)	<0.0001	-0.4142 (0.0698)	(-0.7507, -0.0776)

a. 1-EQ-5D score+0.0001 was used as the dependent variable, Gamma distribution, log link function and backward stepwise selection were applied;

b. 1-EQ-5D score was used as the dependent variable, full model was estimated by bootstrapping 100 times.

GLM Generalized linear models; *CLAD* Censored least absolute deviations; *DQOL* Diabetes quality-of-life; *SE* Standard error; *CI* Confidence interval; *Domain-T* Treatment satisfaction domain score; *Domain-R* Domain score of the domain contains the other 20 items of the short version of DQOL.

Chapter 5

Discussion and Conclusion

5.1 Overview

This thesis focused on identifying the long-term determinants and improving the measurement of the quality of life (QoL) of Chinese type 2 diabetic patients (T2DP). It consists of three individual projects that addressed topics related to the minimally important difference (MID) of the EuroQol 5 dimensions 3-level measure (EQ-5D-3L), clinically relevant and statistically significant determinants of change in EQ-5D-3L scores over a one-year study period among this patient population, shortening the Chinese version diabetes quality of life measure (DQOL), and mapping the Chinese DQOL (and its short version) onto the EQ-5D-3L. The present chapter summarizes the key findings from each project and discusses limitations due to the data employed in this thesis, as well as implications for further research.

5.2 Key findings

Chapter 2 estimated that the MID of the EQ-5D-3L index was 0.0262 for Chinese T2DP, using the anchor question “compared to one year ago, how you would rate your health in general now?" administered at the one-year end follow-up together with its responses. Using this estimated MID as a cut-point, we identified whether the change in patients' EQ-5D-3L scores from the baseline to the one-year end mark was clinically relevant (decrease or increase was greater than the MID) or not (absolute change was equal to or smaller than the MID). Regression results showed that age, being female, having chronic diseases, having a lower level of education, having a low income, never doing physical exercise, and having a high HbA1c level were clinically relevantly and statistically significantly associated with worsening QoL as measured by the EQ-5D-3L over one year.

Chapter 3 shortened the 46-item diabetes-specific QoL measure, Chinese DQOL, based on the classical test theory (CTT) and the item research theory (IRT), each combined with the exploratory factor analysis (EFA). The two short versions and the original Chinese DQOL performed comparably in the validation analyses in terms of the confirmatory factor analysis (CFA) and the criterion validities. Therefore, the short version developed based on the IRT, which kept 24 items, was selected as our final recommended short version for the Chinese DQOL because it can lighten response burden without reducing the psychometric properties, as compared to the 32-item short version developed based on the CTT.

Chapter 4 mapped the Chinese DQOL and its 24-item short version onto the EQ-5D-3L index. Mapping models using estimators including the ordinal least square (OLS), generalized linear model (GLM), and censored least absolute deviations(CLAD) were explored because the EQ-5D-3L index (dependent variable) was not normally distributed and only ranged from -0.149 to 1. Different model specifications using different combinations of domain scores, pair-wise interactions of domain scores, and age and gender were explored. According to the model performance assessed by the pre-set criteria, the GLM model using the life satisfaction domain score, diabetes impact domain score, age, and gender was selected as the mapping algorithm between the original Chinese DQOL and the EQ-5D-3L index. The OLS model, using score of the domain containing 20 items of the short version of the DQOL, age, and gender, was selected as the mapping algorithm between the 24-item short version and the EQ-5D-3L. Since the mapping model for the short version of the DQOL had better individual level prediction precision compared to the one for the original Chinese DQOL, we suggest using the mapping algorithm for the short version when item level scores of the original Chinese DQOL are available.

5.3 Limitations due to data

The interesting findings and methodological and other issues in this thesis have been discussed in detail in Chapters 2-4. This section discusses the limitations due to the data employed in this thesis.

The first weakness of the survey has to do with the generalizability of the findings. The survey recruited community-based diabetic patients who were relatively healthier than the entire Chinese T2DP population. Most of the patients did not have severe mobility impairments. In addition, even though diabetes complication information was not collected during the survey, we could still expect a relatively low proportion of patients with complications. This is because the mean diabetes duration of our patient sample was 7.9 years, and diabetes complications are more likely to occur in patients with a longer duration of diabetes (for example, diabetic retinopathy and nephropathy are more common in patients who have had diabetes for more than 10 years [1-3]). Therefore, findings of Chapters 2-4 cannot necessarily be generalized to the entire Chinese T2DP population.

The second weakness is the methodology regarding the validation/testing data selection. In the item reduction (Chapter 3) and mapping (Chapters 4) studies, we used the baseline data as the training sample and data from the one-year end follow-up as the validation/testing sample; therefore, the training sample and the validation sample were not independent of each other. We did this mainly for two reasons. One was to get a higher statistical power in model estimation processes by using the largest possible sample we had access to. In addition, the Chinese DQOL was only administered at the baseline and at the one-year end follow-up. It was reasonable to expect that, after a one-year gap, patients' health states as measured by the Chinese DQOL would be different from the baseline, and they should have no clear impression of the Chinese DQOL at the year-end. However, since the training and validation/testing data were from the same patient sample and shared very similar demographics, the validation analyses may not able to provide robust results. Currently, the split-sample, cross-validation, and Bootstrap are the commonly used validation approaches (especially in cross sectional studies) when external data is not available [4]. However, the split-sample method decreases the statistical power [5]; and the cross-validation and the Bootstrap can only maximize the utilization of the information of the same population. Neither the approach employed in this thesis nor the three above-mentioned methods can provide a truly external validation. The short version of the Chinese DQOL and the mapping algorithms need to be validated in external populations in the future.

5.4 Implications and future research

The determinants of change in QoL of Chinese T2DP over a one-year period identified in this thesis were in line with those found in Hongkong and other countries' studies [6, 7]. Great importance should be attached to these factors in diabetes management and patient education. In addition, as we mentioned in Chapter 2, the one-year study period was relatively short to observe the change in changeable factors (such as medication, compliance with blood glucose monitoring, and health relative behaviors) and how these factors affect changes in T2DP' QoL. Future diabetes-related research should consider longer time horizons.

We mapped the Chinese DQOL onto the EQ-5D-3L index, which built a bridge between existing studies which only measured the QoL using the Chinese DQOL and the cost-utility analysis (CUA). Future studies can focus on how to apply the algorithms in the CUA and comparing the results of the CUA using utilities estimated from mapping algorithms and that obtained from preference-based measures.

We applied different psychometric theories to shorten the Chinese DQOL and developed a 24-item short version. Though this short version needs further external validation, we believe that it can reduce the burden on both patients and interviewers in future practice. We also found that a lack of generally accepted fail criteria for the CTT and IRT tests was a challenge of implementing item reduction; however, this area was beyond the scope of this thesis. This methodological issue remains to be addressed in future psychometric research.

This thesis was devoted to improving the measurement of the QoL of Chinese T2DP. Currently, the majority of commonly used QoL measures, including both diabetes-specific and generic measures, in Chinese studies were translated from measures developed in foreign settings. Those foreign populations are usually different from the Chinese population in terms of culture, beliefs, and living habits, which significantly affect a population's health preferences [8]. Therefore, developing QoL measures in the Chinese setting will be more meaningful than translating, validating, and improving foreign QoL measures. It is a long and costly process to completely develop and validate a brand new measure; thus, research in this area requires more support from the Chinese government in the future.

References

[1]Butt S, Hall P, Nurko S. Diabetic Nephropathy. From: http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/nephrology/diabetic-nephropathy/ (Accessed March 26, 2017).

[2] Fowler MJ. (2008). Microvascular and Macrovascular Complications of Diabetes. Clinical Diabetes. 26(2): 77-82.

[3]Cortez DN, Reis IA, Souza DA, Macedo MM, Torres H. (2015). Complications and the time of diagnosis of diabetes mellitus in primary care. Acta Paul Enferm. 28(3):250-255.

[4] Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. (2001). Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J ClinEpidemiol. 54(8):774-781.

[5] Marsden PV, Wright JD. (2010). Handbook of Survey Research (Second Edition). Bingley: Emerald Group Publishing.

[6] Wan E.Y., Fung C.S., Choi E.P., Wong C.K., Chan A.K., Chan K.H., Lam C.L. (2016). Main predictors in health-related QoL in Chinese patients with type 2 diabetes mellitus. Qual Life Res. 25(11):2957-2965.

[7]Maatouk I, Wild B, Wesche D, Herzog W, Raum E, Muller H, Rothenbacher D, Stegmaier C, Schellberg D, Brenner H. (2012). Temporal Predictors of Health-Related QoL in Elderly People with Diabetes: Results of a German Cohort Study. PLoS ONE 7(1): e31088.

[8] Jin X, Liu GG, Luo N, Li H, Guan H, Xie F. (2016). Is bad living better than good death? Impact of demographic and cultural factors on health state preference. Qual Life Res. 25(4):979-986.