

University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Author (Year of Submission) "Full thesis title", University of Southampton, name of the University Faculty or School or Department, PhD Thesis, pagination.

Data: Author (Year) Title. URI [dataset]

University of Southampton

Faculty of Social Sciences

SCHOOL OF ECONOMIC, SOCIAL AND POLITICAL SCIENCES

SOCIAL STATISTICS AND DEMOGRAPHY

Exploring Type Two Diabetes Mellitus in the Kingdom of Saudi Arabia:

Studying the Socio-economic Environment

By

Abrar Hasan Fallatah

Thesis for the degree of Doctor of Philosophy

January 2020

University of Southampton

Abstract

Faculty of Social Sciences

School of Economic, Social and Political Sciences

Thesis for the degree of Doctor of Philosophy

Exploring Type Two Diabetes Mellitus in the Kingdom of Saudi Arabia:

Studying Socio-economic Environment

by

Abrar Hasan Fallatah

Type two diabetes mellitus (T2DM) is a global public health challenge, yet the socio-economic factors that contribute to this remain largely unexplored. This high burden of T2DM has profound economic and societal impact, particularly in the Gulf Cooperation Council (GCC) countries, where the oil-boom in the Arabic Gulf region has led to unprecedented growth and urbanisation. Saudi Arabia stands out as one of the countries with the highest prevalence of T2DM worldwide. Rapid socio-economic development in the Kingdom has resulted in lifestyle changes and concomitant changes in non-communicable disease epidemiology. Despite this, there are few national studies exploring the relationship between socio-economic factors and T2DM in Saudi Arabia. In this thesis I address this deficit by investigating the association between education level and the prevalence, diagnosis, and control of T2DM in Saudi Arabia. I explore and evaluate the impact of demography and geographical location together with socio-economic, lifestyle, healthcare access and behavioural factors such as self-care on T2DM outcomes.

I first explore the association between education level and T2DM prevalence in Saudi Arabia and show an inverse association. The prevalence of T2DM is elevated among groups with low education level, together with those who are older, obese, and lived in Southern region. Additionally, the findings support projections in which the prevalence of T2DM continues to increase up to 2030. I then investigate individuals with diagnosed T2DM and analyse the association between education level and other risk factors on the diagnosis of T2DM. These analyses confirm that age and low household income are highly associated with achieving diagnosis of T2DM. Based on these findings, I then develop two prediction models for undiagnosed diabetes and pre-diabetes by considering education level and socio-economic factors. Development of the prediction models established that age, administrative region, and body mass index are the three main predictors of undiagnosed diabetes, although gender and hypertension also emerged as predictors for the compound outcome of undiagnosed diabetes or pre-diabetes.

In the final analysis, I investigate whether education level is associated with achieving control of diabetes through interaction with health-care access and self-care factors. Administrative region, prescribed diet and adherence to medication emerged as key predictors of controlled diabetes. The association with administrative region reflected differences in patterns of T2DM management relating to medication and behavioural modification interventions. My findings indicate that T2DM management forms vary between regions and that the success of lifestyle management approaches is influenced by patients' socio-economic characteristics. The work described in this thesis establishes that education level is associated with the prevalence of diagnosed T2DM.

However, among other socio-economic factors, administrative region stands out as a significant predictor of diagnosed T2DM, undiagnosed T2DM and achieving effective control of T2DM. These findings have important implications for primary prevention of T2DM in Saudi Arabia through population approaches, especially in poorer or more deprived groups. They inform strategies to improve the early diagnosis of T2DM in such groups and suggest approaches to gain better glycaemic control in individuals with diagnosed disease. A combination of these two approaches will prevent complications and improve the quality of life for patients with T2DM.

Table of Contents

Table of Contents	i
Table of Tables	vii
Table of Figures	xi
Research Thesis: Declaration of Authorship	xiv
Dedication... ..	xv
Acknowledgements	xvii
Definitions and Abbreviations.....	xix
Chapter 1 Introduction.....	1
1.1 Background and rationale	1
1.2 Contribution to knowledge	3
1.3 Aims of the thesis	5
1.4 Research questions	5
1.5 Structure of this thesis	6
Chapter 2 Understanding the Risk Factors for the Prevalence, Late Diagnosis, and Inadequate Control of T2DM: A Literature Review.....	9
2.1 Introduction.....	9
2.2 Definitions of diabetes mellitus and diagnostic criteria	9
2.3 Global burden of T2DM and future trends	14
2.4 Overview of T2DM trends in Eastern Mediterranean region	18
2.5 Diabetes prevalence in the GCC region.....	24
2.6 Overview of diabetes studies in Saudi Arabia.....	27
2.7 Conceptual framework for understanding the association between socio-economic factors and T2DM outcomes	33
2.7.1 The Social Determinants of Health Model (SDH)	33
2.8 Risk factors for T2DM outcomes	36
2.8.1 Demographic factors	36
2.8.2 Genetic factors	38
2.8.3 Lifestyle and behavioural factors	40
2.8.4 Socio-economic factors	43

Table of Contents

2.8.4.1	Socio-economic factors and the risk of developing T2DM	43
2.8.4.2	Socio-economic factors and the risk of late diagnosis of T2DM.....	46
2.8.4.3	Socio-economic associations with inadequate control of T2DM	53
2.8.5	Environmental and community factors	55
2.9	Summarising the literature review: Conceptual and analytical frameworks	56
2.9.1	The prevalence of T2DM framework.....	58
2.9.2	The diagnosis of T2DM framework.....	59
2.9.3	The controlled of T2DM framework	61
2.1	Conclusion.....	63
Chapter 3 Socio-economic and Lifestyle Patterns of the Saudi Arabian Context and Health Care Services in Diabetes Management.....		65
3.1	Introduction	65
3.2	Saudi Arabia population.....	65
3.3	Socio-economic drivers and behavioural patterns.....	70
3.3.1	Administrative regions.....	70
3.3.2	Lifestyle patterns	72
3.3.3	Socio-economic characteristics	75
3.4	Saudi health care services.....	79
3.4.1	Health system structure.....	79
3.4.2	Development and challenges of healthcare services	80
3.5	Diabetes Management and Screening: policy, implementation and prevention programmes.....	83
3.5.1	Diabetes policies and plans.....	83
3.5.2	Public health concerns of T2DM	86
3.5.3	Screening and management guidelines for T2DM	87
3.6	Conclusion.....	89
Chapter 4 Data and methods		91
4.1	Introduction	91
4.2	Data sources.....	91

4.2.1	Saudi Health Interview Survey (SHIS).....	91
4.2.1.1	Survey design and flow.....	91
4.2.1.2	Weighting methodology.....	93
4.2.1.3	Ethical considerations	95
4.2.2	Saudi Arabia censuses and population estimates.....	95
4.2.3	United Nations population projections.....	97
4.3	Statistical methods.....	101
4.3.1	Data analysis plan and structure.....	101
4.3.2	Definition of thesis outcomes	103
4.3.3	Inverse probability weight (IPW).....	106
4.3.4	Binary logistic regression model	108
4.4	Conclusion	109
 Chapter 5 The Prevalence of Type Two Diabetes Mellitus in Saudi Arabia: Quantifying the Association of Socio-economic Factors and Predicting Future Trends .111		
5.1	Introduction.....	111
5.2	Background and rationale	114
5.3	Methods	121
5.3.1	Weighting method and sensitivity analysis.....	121
5.3.2	The association between education level and prevalence of diabetes	121
5.3.3	Projection of the prevalence of T2DM.....	128
5.3.3.1	Method 1: United Nations population projections.....	128
5.3.3.2	Method 2: Cohort component projection.....	129
5.4	Results	130
5.4.1	Weighted and unweighted estimates of T2DM prevalence	130
5.4.2	Association between education level and the prevalence of diabetes	132
5.4.2.1	Descriptive results	132
5.4.2.2	Analytical results	135
5.4.3	Projection of the prevalence of T2DM.....	140
5.4.3.1	United Nations population projections.....	140
5.4.3.2	Cohort Component projections.....	145

Table of Contents

5.5	Discussion and conclusion	148
5.5.1	Summary of the key findings	149
5.5.2	The association between education level and the prevalence of T2DM	149
5.5.3	Projection of T2DM prevalence	152
Chapter 6	Diagnosed and Undiagnosed Type Two Diabetes in Saudi Arabia: Exploring the Socio-economic Factors.....	157
6.1	Introduction	157
6.2	Methods.....	159
6.2.1	Study design and participants.....	159
6.2.2	The association between education level and diagnosis diabetes.....	160
6.3	Results.....	164
6.3.1	Descriptive results.....	164
6.3.2	Analytical results.....	169
6.4	Discussion and conclusion	174
Chapter 7	Development of Diabetes and Pre-Diabetes Prediction Risk Models: Identifying Undiagnosed Cases among the Saudi population	178
7.1	Introduction	178
7.2	Background and rationale.....	178
7.3	Methods.....	187
7.3.1	Development of prediction model	188
7.3.2	Validation of prediction model	190
7.3.3	Evaluation of prediction model performance	192
7.4	Results.....	195
7.4.1	Prediction model of undiagnosed diabetes.....	195
7.4.2	Prediction model of undiagnosed diabetes and pre-diabetes	206
7.5	Discussion and conclusion	215
Chapter 8	Socio-economic Factors Related to Controlled Diabetes in Saudi Arabia: Studying the Medication and Lifestyle Behavioural Patterns for Successful Diabetes Management.....	220

8.1	Introduction.....	220
8.2	Methods	226
8.2.1	The association between education level and inadequate control of T2DM	226
8.2.2	Medication and lifestyle patterns among diabetic patients	230
8.3	Results	234
8.3.1	The association between education level and controlled diabetes	234
8.3.2	Medication and behavioural forms for management diabetes.....	242
8.4	Discussion and conclusions	250
8.4.1	Association between education and controlled diabetes	251
8.4.2	Prescribed diet and physical activity level	253
Chapter 9	Discussion and conclusions.....	257
9.1	Contributions of this thesis	257
9.2	Summary of key findings	259
9.2.1	The association between education level and T2DM outcomes	259
9.2.2	Future projection of T2DM prevalence estimates.....	265
9.2.3	Prediction models for undiagnosed T2DM, diabetes and pre-diabetes.....	266
9.3	Strengths and limitations of the study	267
9.4	Policy implications and recommendations	269
9.5	Potential area for future research	272
9.6	Conclusion	273
	Appendix A Saudi Health Interview Survey Questions	275
	Appendix B Saudi National Reference for Diabetes Mellitus Guidelines in Primary Health Care	281
	Appendix C Sensitivity analysis for weighting methodology	284
	Appendix D Projection models for T2DM prevalence	285
	Appendix E Internal validation of the undiagnosed diabetes prediction model using the split sample validation method	289
	Appendix F Prediction model for undiagnosed pre-diabetes	292
	Appendix G Characteristics of observed and non-observed diabetes patients	295
	List of References	297

Table of Tables

Table 2.1 Summary of diagnostic criteria for diabetes and intermediate hyperglycaemia based on the WHO and ADA (WHO, 2006a, American Diabetes, 2014)	12
Table 2.2 Diagnosis of diabetes (Sacks, 2011)	13
Table 2.3 Summary of global estimates of diabetes prevalence studies	16
Table 2.4 Top 10 countries for prevalence of Diabetes Mellitus in the Middle East and North African Region in 2013 (Majeed et al., 2014)	21
Table 2.5 Summary of diabetes studies in the Middle East countries based on a national survey data by sex – age group	23
Table 2.6 Summary of diabetes studies in the GCC countries based on a national survey data by sex – age group	26
Table 2.7 Diabetes Studies in Saudi Arabia based on national surveys.....	28
Table 2.8 Summary of diabetes studies in Saudi Arabia.....	31
Table 2.9 Summary of socio-economic factors and T2DM published studies in the GCC and different income levels countries	49
Table 3.1 Average Household Monthly Income by Administrative Region and Nationality of Household Head	78
Table 4.1 The number of Saudi national population 1992 – 2018	99
Table 4.2 Total number of Saudi population (National and Non-national resident)	100
Table 4.3 Classification of diabetes outcome based on HbA1c test and survey question	105
Table 5.1 Summary of diabetes prevalence projected models	117
Table 5.2 Analytical frameworks for the prevalence of T2DM outcome	122
Table 5.3 Study explanatory variables list.....	127
Table 5.4 Prevalence of diabetes status in the Saudi Health Interview Survey data, KSA, 2013.....	131
Table 5.5 Characteristics of survey participants of prevalence of T2DM (unweighted percentages)	133

Table of Tables

Table 5.6 Generalised binary logistic regression odds ratios and 95% confidence interval of the prevalence of T2DM (Household weighted models).....	137
Table 5.7 Total number (cases) of Saudi diabetic males	142
Table 5.8 Total number of Saudi diabetic females.....	142
Table 5.9 Total number of Saudi diabetics.....	143
Table 5.10 Number of male diabetic patients in two timeframes 2023 and 2028	145
Table 5.11 Number of female diabetic patients in three timeframes 2018, 2023 and 2028 ...	146
Table 5.12 Number of diabetic patients in two timeframes 2023 and 2028	148
Table 6.1 The analytical framework for diagnosis diabetes outcome	162
Table 6.2 Study explanatory variables list	163
Table 6.3 Characteristics of survey participants of diagnosed and undiagnosed T2DM	166
Table 6.4 Generalised binary logistic regression odds ratios and 95% confidence interval of diagnosed T2DM (Household weighted models)	171
Table 7.1 Summary of previously published prevalence of T2DM prediction models (undiagnosed) and their risk factors	183
Table 7.2 Summary of previously published prevalence of diabetes and prediabetes prediction models (undiagnosed) and their risk factors.....	186
Table 7.3 Confusion Matrix/ Classification Algorithm for clinical prediction models.....	189
Table 7.4 Outcome and explanatory variables codes	194
Table 7.5 Characteristics of survey participants with undiagnosed T2DM	197
Table 7.6 Summary statistics for different prediction models based on model selection and performance measures	200
Table 7.7 Risk factors and final development prediction model of undiagnosed T2DM.....	202
Table 7.8 Confusion matrix for undiagnosed T2DM	203
Table 7.9 Summary statistics for different prediction models based on model selection and performance measures	204

Table 7.10 Observed and predicted T2DM status using the test score as a predictive variable in the groups defined by Hosmer-Lemeshow Goodness-of-fit test	205
Table 7.11 Characteristics of survey participants with undiagnosed diabetes and prediabetes	207
Table 7.12 Summary statistics for different prediction models based on model selection and performance measures.....	210
Table 7.13 Risk factors and final development prediction model of undiagnosed T2DM and pre-diabetes	211
Table 7.14 Confusion matrix of undiagnosed pre-T2DM	212
Table 7.15 Summary statistics for different prediction models based on model selection and performance measures.....	213
Table 7.16 Observed and predicted occurrence of prediabetes using the test score as a predictive variable in the groups defined by the Hosmer-Lemeshow Goodness-of-fit test	214
Table 8.1 The analytical framework of controlled diabetes	230
Table 8.2 Study explanatory variables list.....	233
Table 8.3 Characteristics of diabetes patients with controlled and uncontrolled diabetes	235
Table 8.4 Binary logistic regression odds ratios and 95% confidence interval for controlled diabetes (Household weighted models).....	239
Table 8.5 Diabetes management patterns and control of diabetes in Saudi patients	242
Table 8.6 Diabetes management patterns among Saudi patients by administrative region	243
Table 8.7 Generalised binary logistic regression odds ratios and 95% confidence interval for prescribed diet (Household weighted models)	245
Table 8.8 Relative risk of physical activity levels risk factors	248
Table 8.9 Predicted probabilities of the likelihood of physical activity levels by different groups	249
Table 9.1 Summary of three main T2DM outcomes results.....	260

Table of Figures

Figure 2.1 Trends of Diabetes Mellitus from the Global Report on Diabetes (WHO, 2016a)	19
Figure 2.2 The number of diabetic patients in different continents in 2017 and 2045	19
Figure 2.3 Map of Middle East and North Africa Region	20
Figure 2.4 Gulf Cooperation Council Countries Map.....	24
Figure 2.5 Dahlgren and Whitehead model of the social determinants of health (Bambra et al., 2009).....	34
Figure 2.6 Thesis conceptual framework.....	34
Figure 2.7 Analytical framework for the relationship between education level and development T2DM	59
Figure 2.8 Analytical framework for the relationship between education level and diagnosis of T2DM	61
Figure 2.9 Analytical framework for the relationship between education level and inadequate control of T2DM.....	62
Figure 3.1 Kingdom of Saudi Arabia map	66
Figure 3.2 Age-sex population pyramid in 2017	68
Figure 3.3 Total numbers of Saudis and non-Saudis resident population (1974 - 2017)	69
Figure 3.4 Average Monthly Household Income for the total Saudi population by administrative regions using Household Income and Expenditure Survey	78
Figure 3.5 Current structure of the health care sectors in Saudi Arabia (Almalki et al., 2011)...	80
Figure 4.1 Saudi Health Interview Survey flow	93
Figure 4.2 Flow diagram for SHIS stages and sampling weights.....	94
Figure 4.3 Survey flow and diabetes outcomes in this thesis.....	103
Figure 5.1 Projected number of people with T2DM for selected years among Saudi male adults 25 years old age and over in Saudi Arabia.....	141

Table of Figures

Figure 5.2 Trends and future projections of diabetic patient specific number among female adults in Saudi Arabia 2015 - 2030.....	143
Figure 5.3 Projected total number of people with T2DM for selected years among adults 25 years old age and over in Saudi Arabia.....	144
Figure 5.4 Prevalence of T2DM for selected years in three different age groups in Saudi Arabia.....	144
Figure 5.5 Future projection of (T2DM) patient number among Saudi males in three different timeframes; 2018, 2023, and 2028	146
Figure 5.6 Trends and future projections of the diabetic number among female adult in Saudi Arabia 2018 - 2028	147
Figure 5.7 Projected number of people with T2DM for selected years among adults 25 years old age and over in Saudi Arabia.....	148
Figure 6.1 Flow chart for participants' inclusion in the study	159
Figure 6.2 Proportions of diagnosed and undiagnosed diabetes by education level	168
Figure 6.3 Proportions of diagnosed and undiagnosed diabetes by administrative region	168
Figure 6.4 Proportions of diagnosed and undiagnosed diabetes by gender group	169
Figure 7.1 Flow chart of prediction model validation Using Bootstrap Method	192
Figure 7.2 Receiver-operating characteristic (ROC) curves for various prediction models in the presented study.....	200
Figure 7.3 Receiver operating characteristics (ROC) curve showing performance of original sample and bootstrap sample for undiagnosed diabetes	204
Figure 7.4 Calibration of the test score: observed and expected of undiagnosed T2DM.....	206
Figure 7.5 Receiver-operating characteristic curves for various prediction models in the presented study.....	210
Figure 7.6 Receiver operating characteristics (ROC) curve showing performance of original sample and bootstrap sample for undiagnosed diabetes and prediabetes	212
Figure 7.7 Calibration of the test score: observed and expected of undiagnosed prediabetes.....	214
Figure 8.1 Flow chart for participants' inclusion in the study	221

Figure 8.2 Analytical framework for the relationship between education level and inadequate control of T2DM.....	229
Figure 8.3 Proportion of patients taking medication by prescribed diet status.....	247
Figure 8.4 Proportion of patients taking medication by home sugar testing status.....	247
Figure 8.5 Predicted probabilities of physical activity levels by education level among Saudi diabetic patients	249
Figure 8.6 Predicted probabilities of physical activity levels by working status among Saudi diabetic patients	250

Research Thesis: Declaration of Authorship

Print name: Abrar Hasan Fallatah

Title of thesis: **Exploring Type Two Diabetes Mellitus in the Kingdom of Saudi Arabia: Studying Socio-economic Environment**

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signature:

Date:

Dedication

I wholeheartedly dedicated this thesis to my beloved parents,

my mother (Aziza) and my father (Hasan).

*The journey of this doctoral research surrounded with their inspiration, encouragement,
and motivation.*

With my sincere thanks for your constant support, patience, kindness, and love.

Acknowledgements

First, I would like thanks my God (Allah) for giving me the strengths to complete this research. I would like to express my deepest gratitude to my supervisors' Dr Andrew Channon, and Professor Paul Roderick for kindly provided me constant support and encouragement throughout this study. They enhanced my research skills and taught me about appropriate ways to develop my doctoral research project, and their valuable suggestions and comments on statistical methodology. In addition, I would like to extend my thanks to Professor Allan Hill and Professor Nuala McGrath who reviewed my upgrade examination report and provide me useful feedback and constructive comments on the report. My PhD thesis benefitted also from the epidemiological training analyses from Professor Dankmar Bohning based on the Statistical Science Research Institute of the University of Southampton. I extent also my thanks to Professor Nikos Tzavidis and Dr Angela-Luna Hernandez for their helpful explanations in my statistical analyses.

Grateful acknowledgements are extended to The National Diabetes Prevention and Control program and the Saudi Ministry of Health for providing the access to the Saudi Health Interview Survey (SHIS) database. I would like grateful thank Professor Tawfiq Khoja and Dr Mohammed Al-Saeedi for their help to get the access of national dataset. I would like also thank Dr Fatimah Al-Slail, the director of the National Diabetes Program for her support in access to the database. I extent my thanks and appreciation to the General Authority of Statistics in Riyadh for their encouragement and support to extract the Saudi Censuses data.

I also acknowledge the Saudi Cultural Bureau in London, Ministry of Higher Education, Kingdom of Saudi Arabia, and King Abdulaziz University for providing me with funding for my doctoral research project. I extent my sincere thanks and appreciation to my friends and my colleagues in the research office for their support. Last but not least, many thanks and warm gratitude to my beloved parents who always supported me with their continuous prayers and best wishes for my success. To my sisters, and my brother, thanks for your wonderful company and always visiting me at Southampton.

Definitions and Abbreviations

<i>HbA_{1c}</i>	Glycated Haemoglobin
ADA	American Diabetes Association
BMI	Body Mass Index
DALY	Disability Adjusted Life Year
DM	Diabetes Mellitus
EMR	Eastern Mediterranean Region
FPG	Fasting Plasma Glucose
GASTAT	General Authority of Statistics
GCC	Gulf Cooperation Council
GDP	Gross Domestic Product
GOSI	General Organisation for Social Insurance
IDF	International Diabetes Federation
IGT	Impaired Glucose Tolerance
IPAQ	International Physical Activity Questionnaire
IPW	Inverse Probability Weight
KSA	Kingdom of Saudi Arabia
MENA	Middle East and North African
MLSD	Ministry of Labour and Social Development
Mmol/l	Millimoles per litre
MOH	Ministry of Health
NCD	Non-Communicable Diseases
OGTT	Oral Glucose Tolerance Test
OPEC	Organisation of Petroleum Exporting Countries
ORs	Odds Ratios

Definitions and Abbreviations

PHCC	Primary Healthcare Centres
SDH	Social Determinants of Health
SES	Socioeconomic Status
SHIS	Saudi Health Interview Survey
SMOH	Saudi Ministry of Health
T2DM	Diabetes Mellitus Type II
UAE	United Arab Emirates
UK	United Kingdom
USA	United States of America
WHO	World Health Organisation

Chapter 1 Introduction

This thesis contributes to a better understanding of the socio-economic, demographic, spatial and lifestyle patterns underlying the high prevalence, undiagnosed and uncontrolled of T2DM in the Kingdom of Saudi Arabia.

1.1 Background and rationale

Diabetes mellitus (DM) is one of the most common chronic diseases and its prevalence is dramatically increasing worldwide. Globally, 43% of all diabetes related deaths occurred in those under 70 years old and more than 1.6 million people a year die from diabetes (WHO, 2016). In 2012, diabetes was the eighth leading cause of death for both sexes. Diabetes complications include blindness, stroke, heart attack, kidney failure, leg amputation and nerve damage (WHO, 1999). These complications influence quality of life and increase the likelihood/risk of early mortality. The highest proportion of diabetes-related deaths occurred after the age of 50 in middle-income countries and between 60 – 69 in high income countries (WHO, 2016a). The WHO estimated that the number of adults living with diabetes increased four-fold from 108 million in 1980 to 422 million in 2014. Some 40% of this increase is the product of ageing and population growth (WHO, 2014).

Type II diabetes mellitus (T2DM) and Type I diabetes mellitus (T1DM) are both highly prevalent in Saudi Arabia (Al-Rubeaan, 2015). **T1DM** is an insulin deficiency, which usually presents early with symptoms occurring before 20 years of age. Its management requires regular injections with insulin (Alberti and Zimmet, 1998). In contrast, **T2DM** (non-insulin dependent diabetes) is a metabolic disease resulting in chronically high blood sugar levels. The WHO reported that at 13.7%, the Eastern Mediterranean Region (EMR) has the highest diabetes prevalence worldwide (WHO, 2016a). However, these measures of diabetes prevalence underestimate the scale of the health problem presented by the disease as population surveys revealed that much T2DM remains undiagnosed. It has been estimated that 24% to 62% of people living with diabetes are undiagnosed and untreated (Jeon and Murray, 2008). In high-income countries the proportion of T2DM that remains undiagnosed is similarly high at between 30% and 50% (WHO, 2015). The Eastern Mediterranean region has undergone rapid economic growth and urbanisation leading to improvements in education and health care. These positive developments, however, have also led to a shift towards the consumption of western-style fast food and an increasingly sedentary lifestyle. These factors have fuelled the increase in prevalence of T2DM and contributed to an increase in obesity in the region.

The prevalence of obesity in the Middle East and North African (MENA) region is high (Mehio Sibai et al., 2010) and over 40% of the Gulf Cooperation Council (GCC) countries have high rates of obesity, with prevalence consistently higher in females (Mehio Sibai et al., 2010). Some cultural region-specific factors have further exacerbated the high levels of T2DM in the EMR. These include high multi-parity, a lack of physical activity by females and high temperatures that discourage outdoor activity in general. The highest prevalence of physical inactivity is found in high-income countries where it is twice that in low-income countries. Based on the WHO regional classification, however, the highest levels of physical inactivity are observed in the EMR for both adults and adolescents (WHO, 2016a). These considerations point to the importance of comprehensive studies on the role of these factors in the prevalence of T2DM in the GCC region. In particular, there needs to be a focus on the extent to which diet and physical activity can be influenced through nutrition education and exercise programs, as such interventions could be easily implemented within regional cultural norms (Badran and Laher, 2012).

T1DM is less prevalent than type two diabetes, however, like T2DM, its onset is influenced by multiple genetic and environmental factors (Robertson and Rich, 2018). Global estimates of diabetes incidence reveal that T1DM is increasing, particular among younger age groups (Tuomilehto, 2013). The 8th annual report from the IDF (2017) estimated the number of individuals aged under 20 years living with T1DM at over 1.1 million (Al-Rubeaan, 2015), with 96,000 children and adolescents under 15 years diagnosed annually. A Saudi national study covering the period 2001-2007 estimated the prevalence of T1DM at 109.5 per 100,000 among children and adolescents aged up to 19 years, with children aged between 13 and 16 years showing the highest prevalence (243 per 100,000) (Al-Herbish et al., 2008). Another national study, covering the period 2001-2007, estimated the overall prevalence of diabetes among children and adolescents aged up to 18 years at 10.84%, made up of 0.45% with T1DM and 10.39% with T2DM (Al-Rubeaan, 2015). A national diabetes prevalence study conducted by Al-Rubeaan estimated the prevalence among Saudi individuals aged 30 years and above at 25.4%, with prevalence significantly higher among males than females (29.1% vs. 21.9%). This study also provided useful age-sex prevalence data as follows (age group | Male, Female): (30 – 39 | 15.1%, 9.2%) , (40 – 49 | 27.8%, 21.0%) , (50 – 59 | 44.5%, 36.3%) , (60 – 69 | 49.8%, 43.6%) , (70+ | 45.8% , 40.3%) (Al-Rubeaan, 2015).

Overall, however, data on the incidence and prevalence of T1DM in SA is limited, with few published reviews on diabetes epidemiology. Likewise, there are very few official Saudi surveys on T1DM (Abdulaziz Al Dawish et al., 2016, Al-Herbish et al., 2008). I decided, therefore, that I would restrict the scope of my thesis to an analysis of the prevalence and diagnosis of T2DM, reaching a better understanding of risk factors for the disease, and combining these to gain insights into better ways to control this epidemic in the Kingdom.

In 2010, the Saudi Burden of Disease report showed that diabetes was ranked third in a list of the most important risk factors for disability adjusted life years (DALYs)¹ with an increase of 8% in the proportion of years of lost to disability (YLDs) due to diabetes (El Bcheraoui et al., 2014). As a result, the Saudi government, together with the Ministry of Health (MOH), have made extensive efforts to reduce the prevalence of T2DM and its risk factors. These include several prevention and intervention programmes alongside increased health education to raise awareness of the importance of adopting a healthy lifestyle. These initiatives, however, are hampered by our lack of understanding of how social, economic, and cultural changes interact to increase the risk of T2DM. To date, there have been limited studies of T2DM in relation to social determinants in Saudi Arabia and the GCC region. Although European and other Western countries have carried out much research into the social determinants of T2DM, most developing nations are at an early stage in this field, largely due to the complexity of measuring socio-economic status and inequality factors related to T2DM. It is even questionable whether these characteristics can be investigated in some developing countries.

1.2 Contribution to knowledge

To date, most research into the prevalence of T2DM in Saudi Arabia has only evaluated demographic and lifestyle factors (Al-Rasheedi, 2015). Therefore, the first analytical study, presented in Chapter 5, investigates the association between education level and T2DM prevalence in Saudi Arabia. In this analysis, I adopt an approach similar to two established frameworks - the “social determinants of health” (Marmot and Health, 2007) and the “social model of health” (Dahlgren and Whitehead, 1991) as the best examples of approaches to mapping risk factors that include socio-economic status. I use this model to forecast future trends of diabetes prevalence in Saudi Arabia. This chapter therefore addresses a significant research gap, enriching and improving our knowledge regarding the relationship between socio-economic factors and diabetes in Saudi Arabia. It will contribute to the development of policies and inform intervention programs targeted at individuals who are more likely to develop diabetes. In the second analytical study, presented in Chapters 6 and 7 of this thesis, I investigate the association between education level and the diagnosis of T2DM and develop a risk score model for undiagnosed T2DM.

¹ Centres for Diseases Control, “Basic about Diabetes” found at <https://www.cdc.gov/> accessed on 12 June 2016

Given the increasing impact of T2DM on healthcare spending and individual well-being discussed in the previous section, this second analysis is aligned with the Saudi Vision goal of improving the quality of life for all Saudi citizens. In the first part of this chapter, I analyse differences between diagnosed and undiagnosed diabetes in terms of their relationship to socio-economic and health care access factors. I build on these findings to develop a risk model for undiagnosed diabetes cases based on those socio-economic factors in the GCC region identified as important in the previous chapter.

Although a simple diabetes risk score model for the region has been already been developed (Al-Lawati and Tuomilehto, 2007), this model did not include a consideration of socio-economic characteristics and was confined to Oman. In contrast, the model I develop uses data from Saudi Arabia to develop a predictive model for T2DM and gives a likelihood score based on socio-economic factors as well as other known risk factors (e.g. age, weight etc.). These results will further inform the current diabetes prevention program in Saudi Arabia by better identifying high priority socio-economic groups thereby increasing the efficiency of screening and early intervention programs. This aim is closely aligned with the goals of The Diabetes Prevention and Control programme in Saudi Arabia. These highlight the need to develop guidelines for diabetes screening and decrease the incidence of new cases. Although existing risk score models in the region have been internally validated, recent research has not been able to validate such models using datasets from outside the GCC region. For the purposes of this chapter, therefore, I used internal validation approaches based on bootstrapping to validate my Saudi diabetes risk score model.

In the final analytical chapter of this thesis, Chapter 8, I investigate the relationship between education level and the control of diabetes. Through this analysis, I attempted to understand how socio-economic, demographic, and behavioural factors operate at the level of the individual with conformed diabetes to affect outcomes. These findings will contribute to Saudi diabetes management programmes in particular, as consideration of socio-economic and behavioural factors in diabetes intervention programmes could help health care providers. Not only would defining the relationship between controlled of diabetes and these factors influence an individual's diabetes status, but it would also help reduce the risk of diabetes complications and early mortality. The three analytical strands described above therefore make an important contribution to the design of effective diabetes prevention programmes and wider public health policies in Saudi Arabia.

1.3 Aims of the thesis

The overarching aim of this thesis is to study T2DM within Saudi Arabia, primarily focusing on socio-economic factors that influence disease susceptibility and outcomes. The first aim is to study three different outcomes related to T2DM. The first of these relates to T2DM susceptibility and the factors influencing why an individual has diabetes. The second outcome addresses factors influencing an individual's knowledge of their T2DM, focusing on why people are aware or not of their diabetes states. The final outcome relates to factors that predict successful management of T2DM in an individual. The second aim is to predict future trends in T2DM prevalence and identify the key factors that should be targeted to reduce the high burden of T2DM in the Kingdom. The third aim is to develop a diabetes risk score model, including a socio-economic dimension, to identify individuals who are at risk of having the disease. The final aim is to identify different patterns of diabetes management in the Kingdom based on medication and behavioural lifestyle factors.

1.4 Research questions

The proposed study will address the following research questions within the Saudi population in four analytical chapters:

1. What is the association between socio-economic factors and the prevalence, diagnosis, and control of T2DM? How are these associations affected by controlling for other factors including demographic, administrative region, socio-economic, lifestyle, healthcare access and self-care factors?
2. What is the association between education level and the prevalence of T2DM? More specifically, is this association modified by gender? How are such associations affected by controlling for demographic, administrative region, socio-economic and lifestyle factors?
3. What is the association between education level and diagnosed diabetes? Is this association affected by interaction between education level and gender? More specifically, does this relationship persist after controlling for demographic, regional, socio-economic, and healthcare access factors?

4. Which socio-economic factors should be included in a prediction model to yield a risk score to identify individuals who are at risk of having undiagnosed or pre-T2DM? Do socio-economic and health care factors contribute to models for prediction of undiagnosed diabetes and high-risk of diabetes in the whole population (excluding those with diagnosed diabetes)?
5. What factors relate to effective control of diabetes in individuals, with specific focus on educational level? How do demographic, administrative region, health care access, self-care, and other socio-economic factors related to controlled diabetes?

1.5 Structure of this thesis

This thesis contains nine chapters. The current chapter reviews the background to the research and rationale for the study, discusses how I expect it to contribute to our knowledge, details the specific aims of the study and the research questions addressed. **Chapter 2** surveys the burden presented by diabetes by reviewing T2DM prevalence at various levels including global, regional – i.e. the EMR, MENA and GCC regions - and specifically Saudi Arabian estimates. In it, I highlight different definitions of diabetes, their corresponding diagnostic criteria and review risk factors associated with the three T2DM outcomes (overall prevalence, diagnosed disease and controlled diabetes). The chapter summarises Saudi studies that have investigated the prevalence of diabetes and its risk factors. I conclude the chapter by presenting a conceptual framework for understanding the relationship between the three main diabetes outcomes and their associated risk factors. In **Chapter 3**, I discuss the Saudi Arabian context to provide background information. It highlights socio-economic and behavioural patterns such as educational attainment, lifestyle factors, and environmental factors. Finally, I review health care services in Saudi Arabia and conclude with a discussion of Saudi policy regarding the prevention, screening, and management T2DM.

Chapter 4 provides details of the Saudi Health Interview Survey (SHIS), the main dataset used throughout this thesis. In it, I outline the survey flows, design, and questions, and discuss the implications of the study design and its limitations. Other aggregate datasets from the United Nations (UN) and the General Authority of Statistics (GASTAT) are also presented in this chapter. I conclude with a description of the statistical methods and data analysis used to address each research question; however, some detailed discussion of statistical methods is deferred to the most relevant analytical chapter.

The first analysis is presented in **Chapter 5**. Here, I assess the implications of weighting estimates and generate the HbA1c weight. I then assess the association between education level and T2DM in the KSA by applying binary logistic regression, controlling for demographic factors, administrative region, and other socio-economic and lifestyle factors. Finally, I use UN projections and cohort component projections to project T2DM prevalence in Saudi Arabia. Building on the findings from the previous chapter, **Chapter 6** presents the second analysis, in which I aim to investigate the association between educational level and diagnosed T2DM. This analysis applied nested binary logistic regression to investigate the relationship, controlling for socio-economic, demographic, and healthcare access factors.

In **Chapter 7**, I describe the development of a risk prediction model for undiagnosed T2DM and pre-diabetes. Binary logistic regression was used for model development and I applied a bootstrapping method for model validation. The models were evaluated on the basis of sensitivity, specificity, Area under the Curve (ROC) and the Hosmer-Lemeshow Chi-square test. **Chapter 8** is the final analytical chapter, and investigates factors influencing how well diagnosed individuals control their blood sugar level (i.e. achieve glycaemic control), with a particular focus on the relationship between control of diabetes and education level. I applied logistic regression models to investigate the association between controlled diabetes and education level, adjusting for demographic, administrative region, socio-economic, healthcare access and self-care behaviour factors. I go on to investigate the relationship between medication management strategies and control of diabetes and then examine how these strategies relate to the administrative region. I conclude with an analysis of the interaction between the effectiveness of different behavioural interventions - such as diet and exercise programmes - and socio-economic factors.

In **Chapter 9** I summarise the main results of this thesis based on the results presented in the four analytical Chapters (5, 6, 7 and 8) and discuss the strengths and limitations of my approach. I assess the possible impact of my findings and propose a number of recommendations for policy interventions aimed to reduce the risk of T2DM and its complications in Saudi Arabia. I conclude with suggestions for follow-on studies and prospects for future research into T2DM in Saudi Arabia.

Chapter 2 Understanding the Risk Factors for the Prevalence, Late Diagnosis, and Inadequate Control of T2DM: A Literature Review

2.1 Introduction

In this chapter, I review the evidence supporting the association between socio-economic factors and T2DM. In **Section 2.2**, I summarise the definitions of and diagnostic criteria for diabetes, followed by a review of the current burden and future trends for T2DM both globally and in the EMR/GCC regions in **sections 2.3 and 2.4**. **Section 2.5** is a summary of diabetes studies carried out in Saudi Arabia. In **Section 2.6**, I briefly overview the definitions and theories that underpin the Social Determinants of Health and present a framework for better understanding the interaction between T2DM and socio-economic factors. I first present the strategy for identifying appropriate reviews and articles that address my research questions. I follow this section with a consideration of how demographic, genetic, lifestyle and behavioural, socio-economic, and environmental factors also affect T2DM outcomes. I then consider the risk factors for T2DM and discuss the association between T2DM and socio-economic factors in more detail. I summarise findings on how socio-economic factors affect T2DM outcomes in countries with different income levels. The chapter ends in **Section 2.8**, with a description of the thesis framework.

2.2 Definitions of diabetes mellitus and diagnostic criteria

Diabetes mellitus is a common chronic disease resulting from an inability either of the pancreas to produce insulin or effector organs to respond to insulin. Left untreated, the condition causes symptoms in multiple organ systems (Alberti and Zimmet, 1998). In 1999, the WHO defined diabetes mellitus as “a metabolic disorder of multiple aetiology characterised by chronic hyperglycaemia with disturbances of carbohydrate, fat and protein metabolism resulting from defects in insulin secretion, insulin action, or both” (WHO, 1999, p.1). Diabetes often presents with symptoms such as thirst, weight loss and blurring of vision. The long-term effects of diabetes can be severe, including amputation of limbs, blindness and development of nervous system specific complications such as neuropathy and retinopathy (WHO, 1999). Complications are largely due to acceleration of vascular disease processes that increase the risk of coronary heart disease (CHD), heart attack, stroke and chronic kidney disease (WHO, 1999).

Diabetes mellitus is divided into three main types: type I diabetes (T1DM), type II diabetes (T2DM) and gestational diabetes, however there are other rarer forms of diabetes such as malnutrition-related diabetes. As mentioned in Chapter 1, T2DM is much more common than T1DM, typically presents later and its incidence increases with advancing age (Alberti and Zimmet, 1998). Often driven by obesity or metabolic derangement leading to insulin resistance, it is initially asymptomatic.

T2DM can be diagnosed opportunistically from routine blood testing, by formal screening or at a later stage when presented with complications. If poorly controlled, T2DM can lead to episodes of hyperglycaemia (excessively high BG: blood glucose) or hypoglycaemia (excessively low BG). Hypoglycaemia can lead to coma; prolonged hyperglycaemia leads to vascular damage. **Gestational diabetes (GD)** refers to high blood glucose or hyperglycaemia first diagnosed during pregnancy together with its impact on the pregnant women and the infant (American Diabetes, 2014). In addition to these three clinically important types of diabetes, there is increasing awareness of pre-diabetic or intermediate states of diabetes characterised by **Impaired Glucose Tolerance (IGT)** and **Impaired Fasting Glycaemia (IFG)**. Such individuals have a high risk of going on to develop T2DM (Alberti and Zimmet, 1998, WHO, 2006a). Here, I focus on T2DM and its risk factors.

The most widely used criteria for diagnosis of T2DM are: fasting plasma glucose (FPG) ≥ 126 mg/dL or 2-h plasma glucose ≥ 200 mg/dL during an oral glucose tolerance test (Sacks, 2011). **Table 2.1** summarises the evolution of diagnostic criteria defined by the WHO and American Diabetes Association (WHO, 2006a). The WHO first published diagnostic criteria for diabetes and its classification in 1965. These were revised several times based on fasting blood glucose level and/or oral glucose tolerance test results. The purpose of the **Oral Glucose Tolerance Test (OGTT)** is to “evaluate the efficiency of the body to metabolise glucose and for many years has been used as the gold standard for diagnosis of diabetes.” The WHO in 2006 recommended that the OGTT should be retained as the primary diagnostic test for IGT and diabetes, as several studies found that the FPG diagnosis criteria is less sensitive and fails to detect 30% diabetes cases (WHO, 2006a).

More recently, the WHO has considered the level of glycated haemoglobin HbA_{1c} both as a diagnostic test for diabetes and as a screening tool for high risk individuals (WHO, 2011). **Glycated haemoglobin** reflects the average “Plasma Glucose” concentration over the previous 8-12 weeks. A review of the evidence by the World Health Organisation (WHO, 2011) led them to conclude that the HbA_{1c} test approaches the sensitivity and specificity of OGTT and FPG methods. However, the HbA_{1c} test has advantages in that it can be performed at any time without the requirement for fasting overnight before the test (**Table 2.2**). My literature search strategy used a combination of words, phrases, and search terms relevant to risk factors and diabetes outcomes.

Keywords relating to risk factors were: “Socio-economic status, educational level, income status, employment status, gender, marital status, sex/gender, age, ethnicity, social and economic factors, risk factors, residence place, region, hypertension, smoking, physical activity, family history of diabetes, health care access factors, self-care factors, obesity, public health, global, EMR, MENA, Gulf Cooperation Council (GCC), Saudi Arabia, high-middle-low income countries, and developed-developing countries”. These were coupled with the following diabetes outcome related terms: Hyperglycaemia, high blood sugar, blood glucose, self-reported diabetes, develop diabetes, measured diabetes, prevalence of diabetes, diagnosed diabetes, controlled diabetes, type II diabetes, diabetes type two, undiagnosed diabetes, pre-diabetes. My search strategy was primarily carried out on the following databases: Medline, POPLINE, JSTOR, Google Scholar and Science Direct. My goal was to identify references to association between the prevalence, diagnosis, and control of three T2DM outcomes and socio-economic factors in published studies. While the emphasis was on studies in Saudi Arabia, I also wanted to find comparable studies across a range of other countries representing a range of income and development levels.

Table 2.1 Summary of diagnostic criteria for diabetes and intermediate hyperglycaemia based on the WHO and ADA
(WHO, 2006a, American Diabetes, 2014)

		WHO 1965	WHO 1980	WHO 1985	WHO 1999	ADA 2003	WHO 2011
Normal	Fasting glucose	Not specified	Not defined	Not defined	< 6.1 mmol/l	---	<i>HbA_{1c}</i> can be used as diagnostic test for diabetes with a recommended value of 6.5% as the cut point for diagnosing diabetes.
	2 h glucose*	< 6.1 mmol/l			Not specified but < 7.8 mmol/l implied	---	
Diabetes	Fasting glucose	Not specified	≥ 8.0 mmol/l And/or	≥ 7.8 mmol/l or	≥ 7.0 mmol/l or	≥ 7.0 mmol/l or	
	2 h glucose	≥ 7.2 mmol/l	≥ 11.0 mmol/l	≥ 11.0 mmol/l	≥ 11.1 mmol/l	≥ 11.0 mmol/l	
IGT	Fasting glucose	Referred to as borderline state	< 8.0 mmol/l and	< 7.8 mmol/l and	< 7.0 mmol/l and	Not required	
	2 h glucose	6.1 – 7.1 mmol/l	≥ 8.0 mmol/l And < 11.0 mmol/l	≥ 7.8 mmol/l and < 11.1 mmol/l	≥ 7.8 mmol/l and < 11.1 mmol/l	≥ 7.8 mmol/l and < 11.1 mmol/l	
IFG	Fasting glucose	Not defined	Not defined	Not defined	≥ 6.1 mmol/l and < 7.0 mmol/l	5.6 to 6.9 mmol/l	
	2 h glucose				and <7.8 mmol/l (if measured)	Measurement not recommended (but if measured should be <11.1 mmol/l)	
Fasting blood glucose; *OGTT: 2 hours after ingestion of 75g oral glucose load							

Table 2.2 Diagnosis of diabetes (Sacks, 2011)

Glucose measurement	Advantages	Disadvantages
Fasting Plasma Glucose (FPG)	• Glucose assay easily automated • Widely available • Inexpensive • Single sample	• Patient must fast ≥ 8 hours • Large biological variability • Sample not stable • Numerous factors alter glucose concentration as stress or illness. • Concentration varies with source of the sample (venous, capillary, or arterial blood) • Concentration in whole blood is different from that in plasma • FPG less tightly linked to diabetes complication than (A1c) • Reflects a single point in time glucose homeostasis
Oral Glucose Tolerance Test (OGTT)	• Sensitive indicator of risk of developing diabetes • Early marker of impaired glucose homeostasis	• Lacks reproducibility • Extensive patient preparation • Time-consuming and inconvenient for patients • Expensive • Influenced by numerous medications • Subject to the same limitations as FPG, sample not stable, need to be performed in the morning
Glycated Haemoglobin (HbA_{1c})	• Subject need not be fasting • Sample may be obtained any time of the day • Very little biological variability • Sample stable • Not altered by acute factors as exercise, and stress. • Reflect long-term blood glucose concentration. • Assay standardised across instruments. • Accuracy of the test is monitored. • Single sample namely whole blood. • Concentration predicts the development of microvascular complications of diabetes. • Used to guide treatment	• May be altered by factors other than glucose as ethnicity • May not be available in some laboratories/ areas of the world • Cost

2.3 Global burden of T2DM and future trends

Rapid socio-economic development has had a key role in enhancing individual's lives by providing better food resources. However, this improved access to nutrition has influenced eating habits resulting in a transition to an unhealthy diet. Consequently, T2DM is becoming one of the major public health issues of the 21st century, with the increase in prevalence outstripping previous estimates (Whiting et al., 2011). One measure of the impact of this epidemic is that globally around 5 million deaths could be attributed to diabetes among those aged 20 – 99 years in 2015 and 2017 (Cho et al., 2018, Ogurtsova et al., 2017), representing about 12.8% of all-cause mortality within this age group.

Furthermore, approximately 1.8 million or 36.5% of these deaths were in those aged under 60. Africa had highest proportion of deaths attributed to diabetes in this age group. The absolute highest number of deaths T2DM attributable deaths was in the Western Pacific Region (1.7 million) and was lowest in South and Central America (0.27 million). Almost three quarters of deaths related to diabetes occurred in those aged under 69 years (Ogurtsova et al., 2017). In high-income countries, the proportion of diabetes deaths in those aged under 60 was 29.6%, in marked contrast to the 72.7% among low-income countries.

Previous studies are in almost complete agreement that the increase in T2DM prevalence will continue in coming years. The WHO estimated the prevalence of T2DM for the 20 – 79 age group could increase from 5.1% to 6.3% (Sicree and Shaw, 2007), and that globally between 200 – 300 million people may be diagnosed with T2DM over the next ten years (King et al., 1998). These match estimates from the International Diabetes Federation (IDF) that the prevalence of T2DM in adults will increase from 194 million in 2003 to 333 million by 2025 (Sicree et al., 2003). King et al. (1998) investigated the prevalence of T2DM globally among those aged 20 and over at three time points: 1995, 2000, and 2025. They estimated that the prevalence would increase from 4.0% in 1995 (135 million), to 5.4% in 2025 (300 million).

Recent studies have extended these projections to 2030. A report by Wild et al. (2004b) projected that T2DM prevalence would increase from 2.8% in 2000 (171 million) to 4.4% in 2030 (366 million). However, Shaw et al. (2010) examined the period 2010 – 2030 and estimated a rise from 285 million (6.4%) in 2010 to 439 million (7.7%) in 2030. A more recent projection from the IDF is that the diabetic population will increase by about 51% between 2011 and 2030, from 366 to 552 million (Whiting et al., 2011). These estimates in the prevalence of diabetes are in broad agreement; where there are differences, these can be explained by the diversity of datasets used, the methods used for projections and the age groups examined. For example, Wild and Ogurtsova built their model based on extrapolation from similar countries using United Nations population estimates due to a lack of data on prevalence by gender and age. Estimates of prevalence in the report by Shaw were on individuals aged 20-79 and used varying diagnostic criteria, such as WHO or ADA definitions, relying on self-reported or registry data for unavailable blood glucose measurements. The studies by King, Whiting and Ogurtsova used logistic regression modelling to generate age-specific estimates from the UN population projections.

A later global report estimated that the number of people with diabetes would increase from 382 million to 592 million between 2013 and 2035 (Guariguata et al., 2014). The Guariguata model for adults 20 – 79 years was based on previous estimates adjusted for geography, economic development, and ethnicity. The IDF global report by Ogurtsova et al. (2017) expected the number of people affected by T2DM to rise from 415 million to 642 million between 2015 and 2040. The most recent global report by Cho et al. (2018) estimated that 451 million people aged 18 – 99 years were affected by T2DM in 2017, projecting an increase to 693 million in 2045. Strikingly, this report also suggests that some 50% of people with T2DM are undiagnosed, a further 374 million people have IGT and 21.3 million have gestational diabetes. A study by Beagley et al. (2014) showed that estimating undiagnosed diabetes, both regionally and globally, is important for developing policies to address the high prevalence of diabetes. Critically, they observed that the mortality rate is similar for diagnosed and undiagnosed DM. Espelt et al. (2011) drew attention to the very significant proportion of DM that is undiagnosed. It has been estimated that globally 45.8% of the adult population has undiagnosed diabetes (Guariguata et al., 2014), further highlighting the importance of estimating undiagnosed diabetes in driving health policy.

Moreover, Ogurtsova estimated the prevalence and number of people with diabetes at 6.7% (318 million) in 2015, rising to 7.8% (481 million) in 2040. The age-specific prevalence estimates in the Cho report were generated using logistic regression for countries with available data and extrapolation for countries where suitable data was unavailable. Estimates of the global burden of diabetes used published studies from 1990 to 2016 and took account of age group, gender, geographical region, and World Bank reported income levels. **Table 2.3** below shows the global estimates of current and projected diabetes prevalence.

Table 2.3 Summary of global estimates of diabetes prevalence studies

Study	Diabetes diagnostic criteria	Age groups in years	Global estimates	
			Years	Estimate
King et al. (1998)	WHO 1985	20 +	1995 2025	135 million 300 million
Wild et al. (2004a)	WHO 1985	All age groups	2000 2030	171 million 366 million
Shaw et al. (2010)	WHO 1985, ADA 1997 and self-reported	20 - 79	2010 2030	285 million 439 million
Whiting et al. (2011)	WHO 1985, ADA 1997, <i>HbA_{1c}</i>	20 – 79	2011 2030	336 million 552 million
Guariguata et al. (2014)	Self-reported and measured: WHO	20 - 79	2013 2035	382 million 592 million
Ogurtsova et al. (2017)	Reported and measured: fasting blood glucose, oral glucose tolerance tests, glycated haemoglobin (HbA _{1c}), self-report, medical record, random blood glucose.	20 - 79	2015 2040	415 million 642 million
Cho 2018	Reported and measured: fasting blood glucose, oral glucose tolerance tests, glycated haemoglobin (HbA _{1c}), self-report, medical record, random blood glucose.	18 - 99	2017 2045	451 million 693 million

Some of these reports compared projections for the number of people living with diabetes in developed and developing countries. These studies show that the increase in T2DM prevalence will disproportionately affect developing countries. One estimate is that while in developed countries the number of people with diabetes will increase from 51 to 72 million (+ 42% increase), in developing countries the increase will be from 28 to 228 million (+170% increase) (King et al., 1998). The countries with the highest number of people with diabetes in 2025 will be the US, China, and India. One clear distinction is that most people with diabetes in developing countries will be in the 45-64 age group, while in developed countries most people with diabetes will be aged 65 or over. Similarly, Shaw et al. (2010) estimated that the number of people with diabetes will rise by 69% in developing countries as opposed to 20% in developed ones. Overall, therefore, these results illustrate the continued high prevalence of diabetes, particularly in developing countries, and show how age and a country's income level are important factors.

The majority of people with diabetes in low- and middle income countries are aged under 50 (59%) (Guariguata et al., 2014). This is supported by a similar observation (Whiting et al., 2011), that in low-and middle income countries most people with T2DM are aged between 40 – 60 years. The picture is very different in high-income countries, where the highest number of people with diabetes (74%) are over 50 years old (the Guariguata study) or aged 60 years or above (in the Whiting study). Cho's report presents similar findings, with individuals with diabetes aged between 65 and 69 of age some 3 times more likely to have T2DM in high-income than in low-income countries. Diabetes prevalence estimates peaked at 22% among the 75 – 79 aged group in high-income countries; 19% in the 60 – 74 age group in middle-income countries; and 8% in low-income countries among 55 – 64 years old (Cho et al., 2018). The age composition presents absolute values; however, proportions could vary widely due to the different age structure between high- middle- and low- income countries.

With regards to gender difference, the prevalence of diabetes is expected to increase for both sexes through to 2045 (Cho et al., 2018). A previous report by King projected a greater increase in prevalence by 2025 among females, particularly in developed nations and in urban areas, while a later global report by Wild showed a similar trend extended to 2030. These studies indicate that urbanisation level and gender are likely to be important determinants of future diabetes prevalence and the number of people living with diabetes. The evidence discussed above paints a picture of the global impact of the growing diabetes burden. These estimates are important to support the substantial community interventions needed to deflect the huge financial implications of these projections through the adoption of healthier lifestyles and a reduction in obesity.

One of most important findings is that the prevalence of diabetes will increase among individuals aged 65 and above due to demographic change (Wild et al., 2004b). Rapid urbanisation contributes to a sedentary lifestyle that is associated globally with high diabetes prevalence. These high rates of urbanisation together with demographic patterns resulting from increases in life expectancy will lead to further increases in T2DM prevalence (Guariguata et al., 2014). To conclude, it should be an urgent priority for governments to develop policies for the prevention and control of diabetes. Three key requirements of such policies are the need to reduce exposure of populations to diabetes risk factors; early diagnosis of disease; and early access to intervention and medication for those with T2DM. Studies projecting national T2DM prevalence will be reviewed in Chapter 5.

Additionally, in GCC countries most affected by T2DM, the prevalence is 10% above the global average (Al-Rasheedi, 2015) and the GCC countries stand out as having the highest prevalence of T2DM among MENA countries (Majeed et al., 2014). Further, the top ten countries with the highest diabetes prevalence worldwide, three of them are from the GCC region namely: Saudi Arabia, Kuwait and Qatar (Guariguata et al., 2014). Projections for Saudi Arabia show an increase in adjusted prevalence by 24.5% in 2030, which will place it in sixth position globally (Guariguata et al., 2014).

A combination of risk factors contributes to the high prevalence of diabetes in these countries including lack of physical activity along with poor access to sports facilities. The following sections will focus on these trends and risk factors in across the EMR and GCC, with particular focus on Saudi Arabia, and will examine the effect of immigration. The majority of the estimates discussed here only considered country nationals and not any immigrants to the country. Where GCC countries have experienced waves of immigration, the estimates of diabetes prevalence might appear to be inflated if the immigrant population has different socio-economic status or cultural norms. This will be considered in more detail in the next section.

2.4 Overview of T2DM trends in Eastern Mediterranean region

T2DM has become a considerable public health problem in the Eastern Mediterranean region (EMR), with an overall prevalence of diagnosed T2DM of 7.7% (Al Johani et al., 2015). Collectively, the Middle East and North African region (MENA) has one of the highest diabetes rates in world (Majeed et al., 2014). The last WHO Global Report on Diabetes (2016) stated that the prevalence of diabetes in the EMR has increased dramatically from 5.9% in 1980 to 13.4% in 2014 (**Figure 2.1**) (WHO, 2016a). Furthermore, the IDF estimated² that the number of diabetic patients will increase from 425 in 2017 to 629 million in 2045 and that the highest increase will be seen in the MENA and across wider African regions with obvious implications for mortality rates and health expenditure (**Figure 2.2**). Additionally, the EMR recorded the highest mortality rate due to high blood glucose in adults (people aged 20 and over) at 139.6 deaths per 100,000 population (WHO, 2016a). The prevalence of T2DM in the adult population, currently 9.2%, is expected to rise to 11.6% by 2035 (Majeed et al., 2014).

² Arab Health Market Series (2018) Diabetes, the Pulse: ninth Edition. Available at: <https://www2.colliers.com/en-AE/Research/Dubai/Diabetes-Arab-Health-Market-Series-The-Pulse-9th-Edition> (Accessed 20 December 2018)

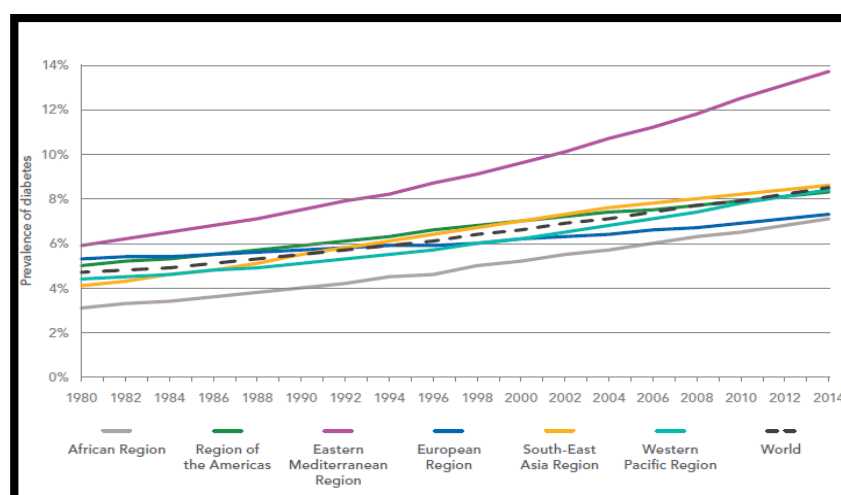


Figure 2.1 Trends of Diabetes Mellitus from the Global Report on Diabetes (WHO, 2016a)

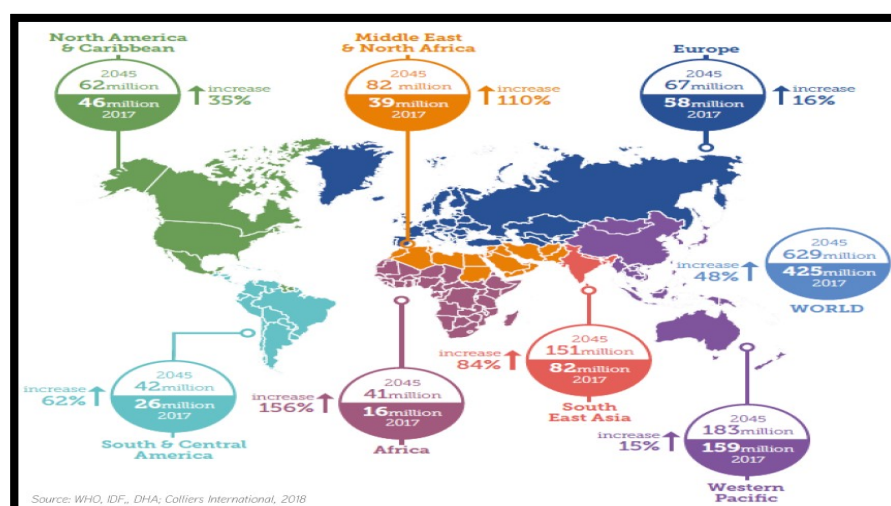


Figure 2.2 The number of diabetic patients in different continents in 2017 and 2045 ³

The previous section highlighted that the Middle East region (Figure 2.1) has the highest diabetes prevalence globally and that across the region, with a rapid increase of diabetes prevalence in both sexes (Khatib, 2004). According to IDF estimates, at 12.5% the MENA region had the highest diabetes prevalence worldwide in 2011. This was predicted to increase to 14.3% by 2030 (Whiting et al., 2011).

³ <https://www2.colliers.com/en-AE/Research/Dubai/Diabetesity-Arab-Health-Market-Series-The-Pulse-9th-Edition>

Chapter 2

Most countries in the Middle East have faced massive and rapid socioeconomic changes over the last few decades. These changes in social and economic status have played a key role in the epidemiological transition that has increased the diabetes burden in the region (Musaiger, 2004, Taylor, 2009). The WHO estimated that the disease burden represented by non-communicable diseases (NCDs) in the EMR would rise from 47% to 60% by 2020 (Khatib, 2004), driven largely by life-style changes causing increases in cardiovascular disease, obesity and T2D. In the Arab world, the number of individuals with diabetes is expected to almost double by 2035 (Musaiger et al., 2011) (IDF, 2013)⁴ (Musaiger et al., 2011; IDF, 2013). Another recent report on Middle Eastern countries observed that 25.8 million individuals in the region had diabetes, with National prevalence up to 8.8% higher than the global average and further increases expected through to 2045 (Al Busaidi et al., 2019)(Al Busaidi et al, 2019).



Figure 2.3 Map of Middle East and North Africa Region⁵

The International Diabetes Federation (IDF) estimates reveal that of the top ten countries worldwide in terms of diabetes prevalence, three are from the GCC region: namely Saudi Arabia, Kuwait, and Qatar (IDF, 2013). Of the top ten countries for diabetes prevalence in the MENA region (listed in Table 2.4), the top five are GCC states led by Saudi Arabia alongside the UAE, Kuwait, Bahrain, and Qatar (Majeed et al., 2014).

⁴ **International Diabetes Federation.** IDF Diabetes Atlas, 6th edition. Brussels, Belgium: International Diabetes Federation, 2013. Available from: <https://idf.org/e-library/epidemiology-research/diabetes-atlas/19-atlas-6th-edition.html>

⁵ https://www.researchgate.net/figure/Figure1-Map-of-the-Middle-East-and-North-Africa-region-as-defined-in-in-this-supplement_fig1_259294603

Table 2.4 Top 10 countries for prevalence of Diabetes Mellitus in the Middle East and North African Region in 2013 (Majeed et al., 2014)

Country	Prevalence (%)
Saudi Arabia	23.87
Kuwait	23.09
Qatar	22.87
Bahrain	21.84
United Arab Emirates	18.98
Egypt	16.80
Lebanon	14.99
Oman	14.24
Jordan	11.40

Along with the increase in T2DM prevalence in the MENA region, there is also a shift in the age of onset, with a marked increase in prevalence among younger age groups from age 20 and even extending to children in some countries (Khatib, 2006). Obesity appears to be the key factor driving the increased prevalence among children (Moussa et al., 2008, Al-Rubeaan et al., 2015). This is a worrying trend that highlights deficiencies in programmes to monitor and control the disease. The prevalence of T2DM has risen dramatically in the last two decades, driven by increase in the prevalence of risk factors for NCDs (Abuyassin and Laher, 2016). For example, in 2004, the regional WHO office based in the Eastern Mediterranean region estimated that 43% of the adult population (≥ 20 year-old) were overweight/obese, with a corresponding prevalence of diabetes at 14.5% (Khatib, 2004). Although it is seen that genetic factors play a significant role in the high prevalence of diabetes (T2DM) in the region, it is the dual increase in obesity and physical inactivity resulting from rapid urbanisation that are the key determinants of the rapid increase (Badran and Laher, 2012).

Chapter 2

In addition to genetics, physical activity, and obesity, these include: multiparity, advanced age, urbanisation, lack of health education and low educational level. These factors will be further justified in the following sections and chapter three. In order for a comprehensive literature review on population-based diabetes prevalence studies in the region to be undertaken, a literature search was carried out using different databases and terms such as: Google Scholar, Science Direct. The terms include “diabetes mellitus, Eastern Mediterranean region, Middle East countries, type two diabetes, prevalence” followed by a review of titles and abstracts for each article identified.

Results of these studies are presented in the following tables. Inclusion studies were national population-based surveys; males and females; prevalence of diabetes and diagnostic criteria reported; multistage stratified random sampling used. Studies were excluded if they were regional, applied to only one gender, were self-reported diabetes prevalence studies, or were duplicates reported in different journals. The selected studies used reliable and representative sampling techniques as well as standard diagnostic criteria for T2DM. Most excluded pregnant women and some did not distinguish between type 1 and 2 diabetes. In this thesis, the diabetes prevalence in adults is based on T2DM, as this represents around 90% of all diagnosed diabetes cases and is the most common type among adults.

Table 2.5 summarise the prevalence of diabetes in Middle Eastern countries from the studies I identified. It is important to recognise that they are not directly comparable due to differences in study design, diagnostic criteria, and population. Overall, the available data confirm that the prevalence of T2DM is high, with highest rates in GCC countries, particularly in Saudi Arabia. The lowest prevalence was reported in Egypt (3.7%), however this was an old study. There were few nationally representative studies for other Middle Eastern countries, however among the few identified, overall prevalence was high in Jordan (17.4%) and Iraq (19.7%). Lower rates were reported in Yemen, Tunisia, and Libya, however these studies covered older-aged groups.

A recent review of diabetes prevalence in the Middle East by Al Busaidi et al. (2019) reported that the age-adjusted prevalence of diabetes was highest in Saudi Arabia (17.7%), followed by the UAE (17.3%), Egypt (17.3%), Bahrain (16.5%), Qatar (16.5%), Kuwait (15.8%), Sudan (15.7%), Lebanon (12.7%), Oman (12.6%), Jordan (11.8%), Libya (10.4%), Iran (9.6%), and Iraq (8.8%). Syria (8.2%) and Yemen (5.4%) have the lowest prevalence in the region. The review confirms that the Gulf Cooperation Council (GCC) countries have the highest diabetes prevalence among ME countries and are among the top ten countries globally. In the next section, I focus on the prevalence of diabetes and its associated risk factors in the GCC countries.

Table 2.5 Summary of diabetes studies in the Middle East countries based on a national survey
data by sex – age group

Study	Country	Age group	Diagnostic criteria	Prevalence of diabetes (%)
Abdella et al. (1998)	Kuwait	20 years and more	OGTT and FPG	14.8
Al Khalaf et al. (2010a)	Kuwait	20 years and more	FBG	21.4
Alarouj et al. (2013)	Kuwait	20 - 65	FPG	17.9
Al Zurba and Al Garf (1996)	Bahrain	20 years and more	FPG	25.5
Malik et al. (2005)	UAE	20 years and more	OGTT and FPG	21.4
Saadi et al. (2007)	UAE	18 years and more	FPG	10.5
Al-Lawati et al. (2002)	Oman	20 years and more	FPG	11.6
Al-Moosa et al. (2006)	Oman	20 years and more	FPG	11.6
Al-Lawati et al. (2015)	Oman	20 years and more	FPG and OGTT	10.4 – 21.1
Bener et al. (2009)	Qatar	20 years and more	OGTT and FPG	16.7
Mansour et al. (2014)	Iraq	19 – 94	FPG	19.7
Al-Habori et al. (2004)	Yemen	25 - 65	FPG	16.7
Gunaid and Assabri (2008)	Yemen	35 years and more	FBG and OGTT	10.4
Abdul Rahim et al. (2001)	Palestine	30 - 65	OGTT	12.0
Ajlouni et al. (2008)	Jordan	25 years and more	FPG	17.4
Albache et al. (2010)	Syria	25 years and more	FPG	15.6
Lfotouh et al. (2008)	Egypt	18 years and more	OGTT	3.7
Hegazi et al. (2015)	Egypt	20 - 79	--	15.6
Noor et al. (2015)	Sudan	18 - 95	FBG	1.3
Romdhane et al. (2014)	Tunisia	35 - 70	FPG	15.1
Kadiki and Roaeid (2001)	Libya	30 - 64	OGTT	14.1

2.5 Diabetes prevalence in the GCC region

The Cooperation Council for the Arab States of the Gulf (GCC) was established in 1981 between six countries located on the Arabian Gulf coast (Figure 2.4); these are Saudi Arabia, Oman, Bahrain, Qatar, Kuwait and the United Arab Emirates (UAE). They have distinct and homogeneous characteristics in terms of religion and culture, leading to shared social, environmental and lifestyle factors (Aljefree and Ahmed, 2015). Building on this unique relationship, the GCC aimed to develop joint investment by members in areas such as health and education. The GCC is dominated by Saudi Arabia in terms of size and population (Saudi population 32 million compared to around 9.3 million in UAE, 1.5 million in Bahrain, 4.5 million in Oman, 2.7 million in Qatar and 4.2 million in Kuwait).⁶



Figure 2.4 Gulf Cooperation Council Countries Map⁷

Collectively, the GCC countries are the major supplier of oil in the world and accordingly have the largest economies in the ME region. This wealth has resulted in rapid socioeconomic development in recent decades, leading to dramatic changes lifestyle and dietary habits (Mabry et al., 2010), providing a likely explanation for the high prevalence of diabetes in the region. As discussed above, with the exception of Oman, all GCC countries are within the top ten countries in terms of diabetes prevalence worldwide (Whiting et al., 2011). In the next section (section 2.6), I present the published national population-based studies of diabetes prevalence in Saudi Arabia. In the following section (2.7) I go on to consider a conceptual framework for understanding the association between the prevalence of diabetes and socioeconomic factors. The association between T2DM prevalence and demographic, socioeconomic and lifestyle factors, together with a consideration of the wider healthcare system, is discussed in more detail in the next chapter.

⁶ <https://gccstat.org/en/#>

⁷ <https://www.crowe.com/-/media/Crowe/LLP/folio-pdf/Deal-making-in-the-GCC-Webinar-Handout-AS-18600-004Dpdf.pdf>

The previous summary table highlights the prevalence of T2DM in and beyond the GCC countries based on national survey data (Table 2.6). Overall, the table reveals the high prevalence of diabetes in these countries and that prevalence is higher in females than males in all countries except Oman. The highest prevalence was reported in Bahrain; however, this study covered older age groups (40 to 69 years-old) who would be expected to have a higher prevalence of diabetes. The prevalence of T2DM in the GCC national studies ranged between 11% - 25.5% with Saudi Arabia having the highest prevalence followed by the UAE, Kuwait, and Qatar. The lowest prevalence was observed in Oman.

Table 2.6 Summary of diabetes studies in the GCC countries based on a national survey data by sex – age group

Study	Country	Age group	Diagnostic criteria	Prevalence of diabetes (%)		
				Females	Males	Both sexes
Al-Lawati et al. (2002)	Oman	20 years and more	FPG	11.3	11.8	11.6
Al-Moosa et al. (2006)		20 years and more	FPG	11.3	11.8	11.6
Al-Lawati et al. (2015)		20 years and more	FPG and OGTT	-	-	10.4 – 21.1
Abdella et al. (1998)	Kuwait	20 years and more	OGTT and FPG	14.8	14.7	14.8
Al Khalaf et al. (2010b)		20 years and more	FPG	-	-	21.4
Alarouj et al. (2013)		20 - 65	FPG	23.6	19.2	32.4
Al-Mahroos and McKeigue (1998)	Bahrain	40 – 69	OGTT	35.3	25.4	29.4
Al Zurba and Al Garf (1996)		20 years and more	FPG	25.0	26.4	25.5
Bener et al. (2009)	Qatar	20 years and more	OGTT and FPG	18.1	15.2	16.7
Malik et al. (2005)	UAE	20 years and more	OGTT and FPG	22.3	20.4	21.4
Saadi et al. (2007)		18 years and more	FPG	-	-	10.5
Al-Nozha et al. (2004)	Saudi Arabia	30 - 70	FPG	21.5	26.2	23.7
Almajwal et al. (2009)		30 years and more	FPG	18.6	15.9	17.2
Al-Rubeaan et al. (2014a)		30 years and more	FPG	-	-	25.4

These studies reveal that the increase in average income across the GCC region and the accompanying rapid socioeconomic transformation has had negative consequences on the prevalence of NCDs, particularly T2DM. On its own, therefore, achieving a high economic level is not sufficient to improve the control and management of diabetes. The next section provides an overview of diabetes prevalence studies and risk factors in Saudi Arabia with a view to understanding how such resources might be better deployed to control the disease.

2.6 Overview of diabetes studies in Saudi Arabia

In the previous section, I established that the prevalence of diabetes is predicted to increase in most nations. In a study of the regional prevalence of T2DM, Guariguata et al. (2014) found that Saudi Arabia is among the countries with the highest prevalence, both within the MENA region and worldwide. In another recent Saudi study, Al-Rubeaan et al. (2014b) estimated that the prevalence of T2DM in those aged 30 years and above was 25.4%, of whom 40.3% were undiagnosed. El Bcheraoui et al. (2014) also investigated the T2DM prevalence in Saudi Arabia and found that 13.4% of Saudis have diabetes of whom 57.8% were undiagnosed and 20.2% were undergoing treatment that failed to control their condition.

The following **Table 2.8** summaries the age-sex-specific prevalence of T2DM among Saudi adults. Three national studies highlighted the increasing prevalence of T2DM in both males and females, with prevalence rates for males higher across most age groups. All these studies were nationally representative and observed that the prevalence increased from 14.8% - 25.4% among people aged 30 years or older. However, it should be noted that there is some variation in the diagnostic criteria used in these studies. The original criteria were based on high OGTT and fasting plasma glucose cut-off values and these have now changed to cut-off values based on low fasting glucose values. This could result in increased detection of diabetes meaning that recent studies will tend to give higher prevalence rates.

In this section, I highlight T2DM studies in Saudi Arabia that have investigated the demographic and health risk factors for T2DM. It will undertake a critical analysis of the arguments used to support the commonly accepted diabetes risk factors in SA namely: demographic, socio-economic, geographical location, and genetic factors. Moreover, it will argue that the socio-economic factors could have a beneficial impact on diabetes in SA, although this is highly dependent on other related risk factors. In order to understand this relationship better, I here explore previous Saudi studies that have investigated diabetes risk factors.

Table 2.7 Diabetes Studies in Saudi Arabia based on national surveys

Author and publication year	Dataset year	Sample size	Diagnostic criteria	Age range	Total prevalence	T2DM	Prediabetes
El-Hazmi et al. (1996a)	National Household Survey 1991	23,493	Diabetes: Fasting plasma glucose ≥ 6.7 mmol/L (≥ 120 mg/dL) and/ or 2-h post-glucose load ≥ 10.0 mmol/L (≥ 180 mg/dL); Pre-diabetes/ IGT: normal fasting blood glucose level (<6.7 mmol/L) and 2-h post-glucose load between 6.7 and 10.0 mmol/L	≥ 30	T2DM = 14.75% IGT = 1.75%	Male = 17.32% Female = 12.18%	Male = 1.30% Female = 2.20%
Al-Nuaim (1997)	National epidemiological household survey (1990 - 1993)	37,000	Diabetes: OGTT: plasma glucose at 2-h > 11.2 mmol/L Pre-diabetes: OGTT: plasma glucose > 7.8 and < 11.1 at 2-h	≥ 15		Urban = 13%; Rural = 7.5% Male = 12%; Female = 14%	Urban = 10.5%; Rural = 8% Male = 7%; Female = 7.7%
Warsy and El Hazmi (1999)	Nationwide study during (1992 – 1995)	14,660	WHO criteria using the fasting and 2-hour post glucose load glucose levels	> 14	Male = 9.7% Female = 7.0%	Male = 5.39% Female = 3.65%	Male = 0.50% Female = 0.72%
Al-Nozha et al. (2004)	5-year National epidemiological health survey (1995 – 2000)	17,232	Diabetes: FPG ≥ 7.0 mmol/L (126 mg/dL) or random value at or > 11.1 mmol/L (200 mg/dL).; Prediabetes: FPG = 6.1 – 6.9 mmol/L (110 and 125 mg/dL)	30 - 70	T2DM = 23.7% IGT = 14.1%	Male = 26.2% Female = 21.5%	Male = 14.4% Female = 13.9%
Al-Rubeaan et al. (2015)	Nationwide, and household data (2007 – 2009)	18,034	Diabetes: FPG ≥ 126 mg/dL (7.0 mmol/L) Prediabetes: FPG = 100 mg/dL and 125 mg/dL (5.6 mmol/L – 6.9 mmol/L).	≥ 30	DM = 25.4% IGT = 25.5%	Male = 29.1%, Female = 21.9%	Male = 26.4% Female = 24.7%

In the last three decades, Saudi Arabia has faced economic growth and urbanisation leading to demographic transitions and an increasingly sedentary lifestyle which have had epidemiological consequences (Al-Nuaim, 1997). Despite this, few studies have investigated the high prevalence of T2DM in Saudi Arabia from a socio-economic perspective. Warsy and El Hazmi (1999) highlighted epidemiological studies that attempted to determine the key risk factors driving the high prevalence of DM in Saudi Arabia. These mainly focused on traditional risk factors as age, sex, BMI, and family history.

Table 2.8 below summarises previous diabetes studies in Saudi Arabia. Thirteen studies aimed to investigate the prevalence of diabetes and its risk factors. Eight studies were not nationally representative, focusing for example on rural areas (Bacchus et al., 1982), on Saudi males only (Fatani et al., 1987) and in urban areas (Anokute, 1990). One study relied on self-reporting of the diabetes diagnosis (Alqurashi 2011), one was only in the Central region (Al-Baghli et al., 2010, Al-Daghri et al., 2011) and one was centred on the Al-Kharj Military Hospital (Karim et al., 2000).

Overall, the prevalence of diabetes increased between 1987 to 2014 and however males were more likely to be diabetic, females were more likely to be diagnosed with the disease. In addition, it should be noted that previous researchers seldom classified different types of DM and often relied on self-reported surveys, with few studies using data from national surveys. Several studies claimed that older aged individuals are at higher likelihood of being diabetic compared to younger individuals (Al-Rubeaan et al., 2015, Fatani et al., 1987, El-Hazmi et al., 1995a, El-Hazmi et al., 1996a, Warsy and El Hazmi, 1999, Al-Nuaim, 1997, Al-Baghli et al., 2010). However, the situation with respect to gender is more complicated.

In terms of gender differences, several studies argued that Saudi males have a higher likelihood of being diabetic compared to Saudi females (Al-Rubeaan et al., 2015, Fatani et al., 1987, El-Hazmi et al., 1995a, El-Hazmi et al., 1996a, Warsy and El Hazmi, 1999). However, some studies found that the prevalence of diabetes was higher in females (Karim et al., 2000, Al-Nuaim, 1997, Al-Baghli et al., 2010). These differences between studies likely reflect diversity in the study populations, with some studies focusing on cities or regions. Furthermore, most studies reported that high BMI is positively associated with the prevalence of diabetes (Fatani et al., 1987, Warsy and El Hazmi, 1999, Al-Nuaim, 1997, Alqurashi et al., 2011, Al-Rubeaan et al., 2014b) and that a family history of DM in first-degree relatives is associated with high risk of diabetes (Al-Nuaim, 1997, Al-Rubeaan et al., 2015).

As previously discussed in Chapter 1, however, the combination of demographic, socio-economic and cultural contexts are important factors that influence the increasing prevalence of T2DM in SA. Analysing socio-economic and demographic characteristics would contribute to a better understanding of how best to apply economic resources to address the health needs of the Saudi population in relation to T2DM (Albejaidi, 2010). Previous research investigating the risk factors for T2DM have tended to underplay the importance of socio-economic factors such as educational level, employment status and household income. Yet as a result of rapid economic growth and urbanisation, Saudi Arabia now ranks as the country with the third highest prevalence of diabetes worldwide (Bennett, 2007).

This study found that monthly income and urbanisation were significantly associated with diabetes. Bennett noticed that diabetes prevalence in urban areas was higher in low and middle-income groups compared to high-income groups. This finding was also observed in Saudi national studies, highlighting that individuals who live in urban areas have a higher likelihood of being diabetic compared to those in rural areas (Al-Nozha et al., 2004, Al-Nuaim, 1997). Hence, the place of individual residence could be an important determinant of T2DM. In urban areas, individuals with low-income status have a higher likelihood of being diabetic or pre-diabetic than high-income individuals (Al-Rubeaan et al., 2015).

Limited studies considered socio-economic factors such as education level, employment status and income status. With regards to socio-economic factors in relation to diabetes in SA, one study found that low-educated individuals had a higher likelihood of being diabetic compared to high-educated individuals in the Eastern region (Al-Baghli et al., 2010). Additionally, income status is significantly associated with diabetes (Al-Rubeaan et al., 2015, Fatani et al., 1987, Al-Baghli et al., 2010). In the Western region, high-income males and those who lived in rural areas had a higher risk of being diabetic compared to low-income males (Fatani et al., 1987). Conversely, two later studies observed that low-income individuals have higher risk of being diabetic compared to high-income individuals (Al-Baghli et al., 2010, Al-Rubeaan et al., 2015). With respect to employment and marital status, unemployed individuals are more likely to be diabetic than those in work (Al-Baghli et al., 2010). Individuals who are currently or were previously married are more likely to be diabetic than those who never married in the Eastern region (Al-Baghli et al., 2010).

Despite these observations, few national studies have investigated the relationship between individual and regional risk factors and how these are related different T2DM outcomes. In the following section I will critically evaluate and review the literature surrounding the risk factors for T2DM. This section will first address one of the main conceptual models for understanding the association between socio-economic factors and health outcomes. Secondly, this section will divide based on this framework. In this thesis, I will attempt to fill this gap by investigating how the three main diabetes outcomes (prevalence, diagnosed, and control) relate to socio-economic factors based on national survey data.

Table 2.8 Summary of diabetes studies in Saudi Arabia

Study	Year of data collection	Sample size	Diagnostic criteria	Summary of key findings / estimates	Method	Exposure measures
Fatani et al. (1987)	Western province 1987	5222	Random Blood Glucose/ Oral Glucose Tolerance Test	- They found a high prevalence of diabetes in rural areas in the western region. - The overall prevalence was 4.3%, (5.9% for females and 2.9% for males). - The prevalence was highly associated with advanced age and high-income groups. Most diabetic patients were obese by 41.2%.	Chi-square test	Age, sex, BMI, and income
El-Hazmi et al. (1995a)	Riyadh city Central Region 1991-1993	3981	WHO criteria	- They investigated the prevalence of IGT and DM type II and I. - The prevalence of T2DM was 4.76% for men and 4.10% for women from two to 70 years old. - The IGT was high for females. While males have a higher prevalence of diabetes than females. - The prevalence of diabetes increased with age increases.	Multiple logistic regression	Sex and age
El-Hazmi et al. (1996a)	National Household Survey 1991 34 areas	23,493 (2– 70) years	Fasting plasma glucose $\geq$ 6.7 mmol/L ($\geq$ 120 mg/dL) and/ or 2-h post-glucose load $\geq$ 10.0 mmol/L ($\geq$ 180 mg/dL)	- They estimated the prevalence of IGT, diabetes type II and I for females and males. - The prevalence of type II diabetes was 17.3% in males and 12.2% in females between (31 – 70 years). - The prevalence of diabetes increased with age increases.	Linear model	Sex and age
El-Hazmi et al. (1996b)	Al-Baha population Southern Region	2060	Fasting blood glucose level $\geq$ 6.7 mmol/l ($\geq$ 120 mg/dl) and 2-h post glucose load $\geq$ 10.0 mmol/l ($\geq$ 180 mg/dl)	- In this study, they added BMI as an important indicator of DM. The prevalence of DM was associated with age, sex, and BMI. - The overall prevalence was 7.1% (type II = 6.9%, type I = 0.2%, IGT= 0.8%). - There is a significant association between diabetes and high BMI. Most diabetic patients were obese.	Chi-square test	Age, sex and BMI
Al-Nuaim (1997)	National epidemiologic al household survey Rural and Urban areas (1990 - 1993)	13,177 (15 years and above)	Random Blood Glucose/ OGTT: plasma glucose at 2-h $>$ 11.2 mmol/L	- The prevalence of diabetes in urban areas was higher than rural areas. He added residence place (rural and urban) – family history of DM – BMI. - Residence, family history, age and BMI were significantly associated with diabetes. - Females in rural and urban areas diagnosed with diabetes more than males. - The DM prevalence: urban area (12.0% males, 14.0% females) – Rural area (7.0% males, 8.0% females). - The prevalence of IGT: urban area (10% males, 11% females) – Rural area (8% males and females).	Multiple regression analysis	Age, sex, residence place, family history and BMI.
Al-Baghli et al. (2010)	Eastern Province	197,681 (age $\geq$ 30)	Fasting Plasma Glucose FPG $\geq$ 126 mg/dL (7.0 mmol/L)	- The prevalence of DM is 17.4% and IFG 2.7%. - The results indicated that individuals who are older age, being married, divorced, and widowed, illiterate, low-income, women and unemployed were positively associated with diabetes.	The univariate difference tested by using Chi-square test Multiple logistic regression analyses for explore the association and risk factors of DM and IFG	Age, sex, marital status, occupation, education level, and income level

Table 2.8 Summary of diabetes studies in Saudi Arabia (cont.)

Study	Dataset Year	Sample size	Diagnostic criteria	Summary	Methods	Exposure measures
El Hazmi (1998)	National survey (1992-1996)	25,337	Fasting blood glucose level ≥ 6.7 mmol/l (≥ 120 mg/dl) and 2-h post glucose load ≥ 10.0 mmol/l (≥ 180 mg/dl)	- The prevalence of diabetes increases with advanced age. - The prevalence of DM was 4.53% for females and 5.63% for males. - The IGT in females was higher than in males. - Although it is not surprising that the prevalence of DM increased more with advanced age groups, the middle-age group have the highest prevalence compared with other age groups.	Proportions and Chi-square test	Age, sex.
Warsy and El Hazmi (1999)	Nationwide study during (1992 -1996)	14,660 (age > 14)	WHO criteria using the fasting and 2-hour post glucose load glucose levels	The prevalence of diabetes was 9.7% males and 7.0% females.	Proportions and Chi-square test	Age, sex and BMI
Karim et al. (2000)	Al-Kharj city	3,747 (age ≥ 35)	Random Blood Glucose/ Oral Glucose Tolerance Test	The prevalence of DM was 5.3% in females and 2.6% in males. DM increases with age.	Age proportions of DM.	Age and sex.
Al-Nozha et al. (2004)	5-year National epidemiological health survey (1995 – 2000)	17,232 (30 – 70 years)	FPG is ≥ 7.0 mmol/L (126 mg/dL) or a random value at or above 11.1 mmol/L (200 mg/dL)	- The prevalence of diabetes was 23.7% (diagnosed). - High prevalence of DM in urban areas compared with rural areas (25.5% - 19.5%). (Further longitudinal studies are required) 26.2% males – 21.5% females.	Age-sex proportions of DM.	Sex and residence place, age, region
Al-Daghri et al. (2011)	Central region Riyadh Cohort	9,149	fasting plasma glucose (FPG) level ≥ 7.0 mmol/L (126 mg/dL)	The overall prevalence of diabetes was 23.1% (age-adjusted = 31.6%, 34.7% males, 28.6% females).	Age-sex proportions of DM.	Age and sex
Alqurashi et al. (2011)	2009 Cross-sectional Jeddah city	6024	Self-reported	- The prevalence of diabetes is 30.0% (34.1 in males – 27.6% in females). - Females have a higher prevalence of diabetes than males for ages less than 50 years old. - Obesity has a higher prevalence within females than males.	Linear regression	Age, sex and BMI.
Al-Rubeaan et al. (2015)	Cross-sectional National household survey (2007 - 2009)	18,034 (age ≥ 30)	FPG is ≥ 126 mg/dL (7.0 mmol/L)	- The overall prevalence of DM is 25.4%, while 40.3% are unaware of having diabetes. - Risk factors of DM are history of gestational diabetes, family history of DM in first-degree relatives, residence place and income. - BMI associated with diabetes and prediabetes. - Urban areas have higher prevalence of DM than rural areas, and individuals with monthly income <4,000 and 4,000 – 8,000 SR. - IGT prevalence associated with individuals with 4,000 – 8,000 SR.	ANOVA test Logistic regression analysis	Age, income, BMI, GDM, gender, residence place, family history

2.7 Conceptual framework for understanding the association between socio-economic factors and T2DM outcomes

This section provides a critical evaluation of the risk factors for the three main T2DM outcomes (prevalence, diagnosed and controlled) drawing on relevant literature with a focus on socio-economic factors. This section first presents one of the main frameworks - the Social Determinants of Health (SDH) model - by (Dahlgren and Whitehead, 1991). This framework has been used to gain a better understanding the complex relationship between socio-economic factors and various health outcomes. However, in this chapter, I will address the main demographic, genetic, lifestyle, socio-economic, and environmental risk factors for T2DM using the SDH framework. I also review studies from other countries examining the impact of socio-economic factors on T2DM outcomes. I conclude by summarising the characteristics of T2DM based on past research and examining theories addressing the association between socio-economic factors and T2DM outcomes to provide a conceptual framework for this thesis.

2.7.1 The Social Determinants of Health Model (SDH)

Dahlgren and Whitehead established a model to understand the relationship between socio-economic factors and health outcomes as they relate to T2DM. A systematic review by Bambra et al. (2009) illustrated the importance of social determinants of health (SDH) using the Dahlgren and Whitehead “rainbow” model. This framework was developed to address the concept of health equity together with its policy implications (Dahlgren and Whitehead, 1991) and can be used to help identify a wider range of SDH and better understand their interactions. The model identifies four major layers of social determinants. The two external layers present environmental factors (external factors) as living and working circumstances. The two inner layers identify factors relating to the individual (internal factors) such as education, unemployment, housing, and access to healthcare services.

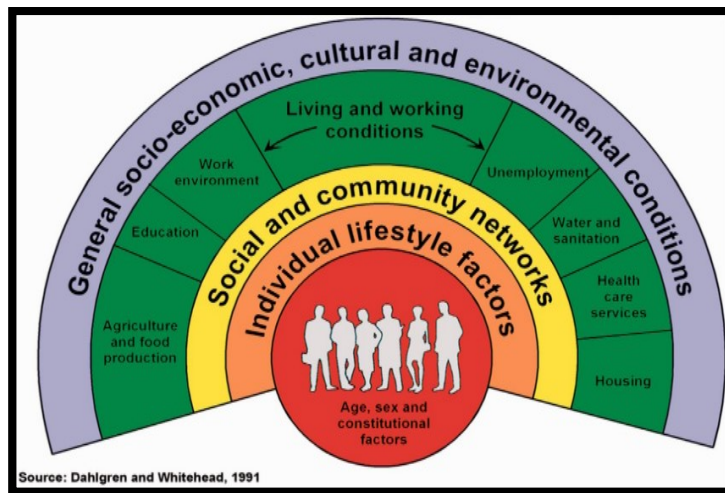


Figure 2.5 Dahlgren and Whitehead model of the social determinants of health (Bambra et al., 2009)

I used this model as a framework for integrating the factors identified as important for T2DM from the literature review. For example, lifestyle, health care access and self-care factors are likely to be associated with an individual's socio-economic characteristics and related to T2DM (see **Figure 2.6**). Based on my review of previous studies, individuals with lower SES are more likely to have poor health outcomes. This lack of equality in health status in most populations reflects the unequal distribution of socio-economic resources between individuals. The SDH model is therefore useful for understanding the true relationship between unequal access to health care for socially disadvantaged groups (Marmot and Health, 2007).

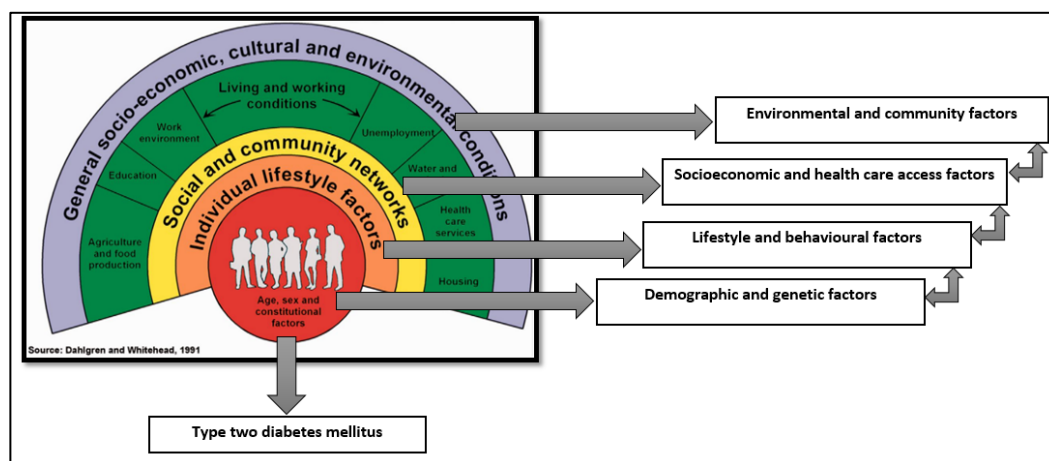


Figure 2.6 Thesis conceptual framework

In the article entitled: *“Closing the gap in a generation: health equity through action on the social determinants of health.”* Marmot has defined social determinants of health as: *“the conditions in which people are born, grow, live, work and age, including the health system. The circumstances are shaped by the distribution of money, power, and resources at global, national, and local level. The social determinants of health are mostly responsible for health inequalities – the unfair and avoidable differences in health status seen within and between countries.”* (Marmot 2008).

For example, women tend to live longer than men due to biological differences. However, in some societies, women have a similar or lower life expectancy than men, which could indicate that there are social conditions that negatively affect females (Marmot and Health, 2007). Researchers have tended to focus on factors directly implicated in the high prevalence of diabetes such as obesity and lack of education. This approach has neglected key social determinants such as poverty, despite the fact that public health policies that provide targeted economic support could provide better healthcare access and address some of the other risk factors in the longer term.

By understanding and addressing these determinants through policy intervention, inequality between the poor and rich individuals within society can be minimised. A better understanding of the social determinants of health could therefore reduce health inequalities, improve the health status of the whole of society and reduce the burden of disease for future generations (Marmot, 2005). In conclusion, it is clear from **Figure 2.5** and the above considerations that:

- Social determinants of health are key to understanding the variation in health outcomes between societies
- Social determinants have a wide influence on population health and are key to tackling health inequalities within society
- The improvement of healthcare services is a high priority in reducing health inequality

Policy makers should therefore consider the relationship between social determinants and health by managing the mismatch between health care provision and human needs (Marmot, 2005). This improved provision should take account of socio-economic diversity at the individual, regional and national level. Based on this framework, the following sections provide a discussion of the risk factors for the prevalence, diagnosis, and controlled diabetes outcomes. The following studies used different measures of T2DM status and outcomes such as:

- Measured “survey-diagnosed diabetes using diagnostic criteria” using various diabetes diagnostic criteria e.g. OGTT, FPG, and HbA1c as discussed at the beginning of this chapter.

- Self-reported by asking: “Have you ever been diagnosed with diabetes by a doctor or health professional?” This is not considered as measured diabetes status.
- Measured and self-reported for allowing classification of individuals as previously diagnosed or undiagnosed, controlled or uncontrolled.
- Hospital or medical records that included only diagnosed diabetes patients’ information.
- Some studies considered medication or anti-diabetic treatments.

2.8 Risk factors for T2DM outcomes

In this section I will address the main socio-economic, demographic, lifestyle, environmental and genetic risk factors for T2DM. The section aims to provide a critical evaluation of the risk factors for the three main T2DM outcomes drawing on relevant literature with a focus on socio-economic factors. After that, I evaluate the impact of healthcare access and self-care factors on the diagnosis and management of diabetes. I also review studies from other countries examining the impact of socio-economic factors on T2DM outcomes. I conclude by summarising the characteristics of T2DM based on past research and examining theories addressing the association between socio-economic factors and T2DM health outcomes to provide a conceptual framework for this thesis.

2.8.1 Demographic factors

In this section, I consider the association between demographic factors (age and gender) and diabetes prevalence/control of diabetes. With respect to age, a study by El Hazmi (1998) highlighted the high prevalence of diabetes in the middle-aged group (45 years old) in Saudi Arabia and how prevalence of DM increases with age. In a UK study, Connolly et al. (2000) investigated the high prevalence of diabetes in the middle-aged group and found a significant association between SES and T2DM. All studies discussed above indicated diabetes prevalence increased with age. The reviews by Sicree and Shaw (2007) and Whiting et al. (2011) discussed above describe how T2DM prevalence increases in the middle-aged group, but also underline the importance of considering gender in view of the biological differences between males and females.

In fact, most previous studies reveal a relationship between gender and T2DM prevalence, however the size and timing of these effect vary according to the study setting. The most consistent finding is that males are more likely to have T2DM than females. Examples of this are two Saudi studies – those of El Hazmi (1998) and Warsy and El Hazmi (1999) - and a study by Connolly et al. (2000) in the UK, all of which report that T2DM prevalence is lower in women than men. These authors further investigated the relationship by considering social class. This revealed that there was a significant association between T2DM and SES, for both men and women and that this was most marked in the 40-49 age band. However overall women showed a lower prevalence of T2DM than men irrespective of age or SES, they appeared more sensitive to changes in the deprivation score (a proxy for SES). These studies were based on different age bands, which complicates direct comparisons between them.

Similarly, Robbins et al. (2005), Reisig et al. (2007) and Agardh et al. (2011) published studies showing that SES has a strong inverse association with the prevalence of T2DM in women across different populations. A review of diabetes in Europe by Espelt et al. (2011) found subgroups of females, characterised by lower levels of physical activity, high BMI, and other psychosocial factors, who have a higher incidence of T2DM than males. This reflects the greater variation of diabetes prevalence due to SES status in women discussed above. Connolly et al. (2000) had already suggested that the relationship to SES reflected the level of exposure to modifiable risk factors. In accord with this view, Robbins et al. (2005) found that correcting for these factors reduced the association between SES and T2DM. For females, the bulk of the residual association was with occupational status. However, for males, only the poverty income ratio emerged as a significant risk factor after correction.

Studies on the influence of demographic factors such as age on the control of diabetes present a complex picture. Bi et al. (2010) reported that in the Jiangsu region of China, older adults with T2DM were less likely to have good glycaemic control. Other studies have also reported that middle-aged and older adults showed poorer glycaemic control than younger adults (D'Souza et al., 2015, Al-Baghli et al., 2010, Alramadan et al., 2018, Ajlouni et al., 2008, Cowie et al., 2010, Al-Kaabi et al., 2008). However in another angle on this debate, a US study found that younger adults were more likely to be facing financial, working and family responsibilities as reflected in high scores on a Diabetic Distress Scale, perhaps contributing to a lack of awareness as to their health status (Wardian and Sun, 2014).

Similar results were obtained in a Canadian study showing that age can influence patient self-management indirectly through Diabetes-related distress (Aghili et al., 2016). A study conducted in Oman also found that older patients (>40) had better glycaemic control than younger patients (20-39) (Al-Lawati et al., 2012). Finally, in an alternative perspective contrasting with the evidence presented above, a number of studies found no evidence for an association between age and control of diabetes in either direction (Al-Hayek et al., 2012, Al-Kaabi et al., 2009, Al Balushi et al., 2014, Ajabnoor and Laajam, 1987, Al-Nuaim et al., 1998, Abdelmoneim and Al Homrany, 2002, Malik et al., 1999).

Several studies have investigated the association between gender and control of diabetes with mixed findings. One study reported that male patients were more likely to exhibit good glycaemic control (AOR = 2.84) (Abdelmoneim and Al Homrany, 2002, Al-Baghli et al., 2010). However, a similar Omani study found the opposite effect, with female patients more likely to show glycaemic control (AOR = 1.5) (Al-Lawati et al., 2012). Other studies have not found any significant associations between gender and glycaemic control (Al-Hayek et al., 2012, Al-Kaabi et al., 2009, Al-Kaabi et al., 2008, Al Balushi et al., 2014, D'Souza et al., 2015, Ajabnoor and Laajam, 1987, Al-Nuaim et al., 1998, Malik et al., 1999, Alramadan et al., 2018). The evidence highlights that the association between demographic factors and controlled diabetes could be related to differences between study populations. To conclude, the majority of the previous evidence highlighted that there are likely to be gender differences in relation to control of diabetes, however these are complicated by other demographic and SES factors.

2.8.2 Genetic factors

In addition to age and gender, ethnicity is another factor to take into consideration when investigating the association between SES and prevalence of DM (Agardh et al., 2011, Calixto and Anaya, 2014, Piccolo et al., 2014). A study by Robbins et al. (2005) found that ethnicity is a significant factor in developing DM among low SES individuals. More specifically, Piccolo et al. (2014) convincingly showed that Africans and Hispanics with low SES have a higher prevalence of T2DM than Caucasian individuals. However, studies by Bhopal et al. (2002) and Walker et al. (2014) suggested that further evaluation of the interaction between socio-economic factors and ethnicity as they relate to prevalence of T2DM was required, particularly with regard to gender-related differences and age stratification. Kumari et al. (2004) found significant association between T2DM and ethnicity; for example, South Asian countries have a high prevalence of T2DM compared with European countries. Furthermore, T2DM prevalence is increasing in European countries with a high immigrant population, predominantly those with poor SES (Espelt et al., 2011).

It is known that a family history of T2DM is an important risk factor and that this also applies in the Saudi population (Warsy and El Hazmi, 1999). Genetic factors likely contribute significantly to the high prevalence of T2DM within certain ethnic groups in e.g. South East Asian countries (Hussain et al., 2007, Agardh et al., 2011). Further studies of genetic risk factors within specific ethnic groups might therefore help clarify the relationship between T2DM and socio-economic factors (Agardh et al., 2011). Three studies have investigated the relationship between family history and glycaemic control with conflicting results. Malik et al. (1999) observed that good glycaemic control was associated with a family history of T2DM. However, another study reported that poor glycaemic control was associated with having a family history of diabetes (Ajlouni et al., 2008). A third study by Abdelmoneim and Al Homrany (2002) found no significant relationship.

Consanguinity in marriage is another important factor contributing to the high prevalence of T2DM (El-Hazmi et al., 1995b). Most research notes that there is a positive association between T2DM and consanguineous marriage in this region (El Mouzan et al., 2008, Midhet et al., 2010). For example, one study reported that 80% of relative marriages had a family history of T2DM compared to 20% in non-relative marriages (Badran and Laher, 2012). Thus, genetic risk factors contribute to the high prevalence of T2DM in the region (Abuyassin and Laher, 2016). These genetic influences have probably remained constant, however, and the lifestyle changes resulting from rapid urbanisation and economic growth have probably been the major cause of the increase in T2DM prevalence over the last few decades.

Few studies have investigated the association between ethnicity and control of diabetes. One Dutch study reported that Asian Americans have a higher likelihood of having diabetes than non-Hispanic Whites, and that ethnic minorities had poor outcomes that could not be attributed solely to differences in self-management (Lanting et al., 2008). A similar finding was reported in a US study showing that non-Hispanic Whites have better diabetes control compared than Mexican Americans and non-Hispanic Blacks (Rothman et al., 2008, Welch et al., 2006). Finally, although one study in the UAE by Malik et al. (1999) did find a significant relationship between ethnicity and glycaemic control, interpretation of the observation was complicated by ethnic differences in preference for hospital or health-centre management. The evidence presented has shown that ethnicity and family history were meaningful risk factors for prevalence and controlled diabetes.

2.8.3 Lifestyle and behavioural factors

Previous research has highlighted that socio-economic factors influence lifestyle and behaviour and that better understanding the association between lifestyle factors (such as diet, physical activity, obesity, diet and smoking) and T2DM is important in the prevention and better management of disease. For example, changes in diet and physical activity have led to an increase in the prevalence of obesity, a key risk factor for T2DM (Popkin et al., 2012). Consequently, health education has a key role in increasing awareness of unhealthy behaviour such as smoking, physical inactivity, and poor nutrition, central to health policies targeting T2DM (Wray et al., 2006). Physical activity, diet and smoking are significant lifestyle factors associated with T2DM. Furthermore, these factors are unequally distributed within different socio-economic groups (Solar and Irwin, 2007). Urbanisation is a related factor that leads to population changes in lifestyle and subsequent increases in obesity as discussed in Chapter 2. Hence, obesity is a key factor directly linked to the metabolic changes underlying the development of T2DM.

Physical inactivity is a major risk factor that is known to lead to an increase in the prevalence of both obesity and T2DM (Sicree and Shaw, 2007, Guariguata et al., 2014, Connolly et al., 2000, Hussain et al., 2007, Lidfeldt et al., 2007, Tamayo et al., 2010, Kumari et al., 2004, Espelt et al., 2011). It has been reported that there are high rates of physical inactivity in the MENA region, ranging from 21.6% - 86.8%, with females less active than males (Mehio Sibai et al., 2010). There is therefore a relationship between physical inactivity and obesity, leading to an increase in the prevalence of T2DM. A common factor underlying the high prevalence of obesity among females in the GCC countries is limited access to sports facilities, contributing to sedentary lifestyles (Al-Rasheedi, 2015). A study in Oman evaluated the reasons for the high prevalence of **obesity** and physical inactivity in the GCC countries (Al-Lawati and Jousilahti, 2004). The authors highlighted that the increase in obesity in Oman was related to the oil boom in the early 1980s. The discovery of oil was associated with rapid socio-economic developments and an adoption of western lifestyles such as: physical inactivity, dietary changes, and smoking habits. The increase in urbanisation due to internal migration of Omanis from rural and agricultural regions also contributed to an increasingly sedentary lifestyle.

Subsequently, obesity has decreased somewhat among Omani females due to an increase in educational attainment and reduction in fertility rates (Al-Lawati and Jousilahti, 2004). This trend began in the 1970s, when culturally driven gender differences in Oman began to decrease. In the early 1970s the illiteracy rate among Omani females was high, falling to 21.2% by the end of the decade, however still higher than the value for Omani men (15.8%). However, socio-cultural restrictions applying to public facilities remain in Oman, with mixed gender exercise facilities less accessible to women and women-only facilities scarce and expensive (Al-Lawati and Jousilahti, 2004). In addition, cultural norms make it less acceptable for women to walk or exercise alone outdoors. A Saudi study by Al-Nozha et al. (2007) reported that 96.1% of adults were physically inactive as defined using the International Physical Activity Questionnaire (IPAQ). The hot climate in most regions means that individuals are less likely to participate in physical activity, leading to a high dependence on, and cultural acceptance of, car use - even for short journeys (Al Othaimen et al., 2007, Al Johani et al., 2015).

Moreover, **obesity** is known to be one of the largest contributors to both morbidity and mortality as measured by disability affected life years (DALYs) (Ezzati et al., 2002). Several previous studies have highlighted the high prevalence of overweight and obese individuals within the EMR. A similar prevalence of obesity within the MENA region was confirmed by (Mehio Sibai et al., 2010). These authors reported obesity rates of 40% among women in oil rich states, much higher than in Europe (10-20%) and the United States (32.2%) (Control and Prevention, 2006, Ogden et al., 2006, Seidell, 2000). In the Arabian Gulf region, since the oil discovery in 1960, the GCC countries have experienced an increase in population size, per capita income and total wealth (Ng et al., 2011). This has led in turn to an increase in social and economic status associated with major lifestyle changes. As a result, the growth in unhealthy diet and physical inactivity are the two main risk factors driving the regional increase in obesity.

Over 40% of the GCC countries have high rates of obesity, with prevalence higher in women (Mehio Sibai et al., 2010). The prevalence of obesity is particularly high in Saudi Arabia, Kuwait and Qatar in the 30 - 60 age group (70 - 85% in males and 75 - 88% within females) (Ng et al., 2011). Mehio Sibai et al. (2010) highlighted that the prevalence of diabetes parallels obesity rates, and exceeds 23.0% among adults in Bahrain, the UAE, and Saudi Arabia. In Saudi Arabia over 50% of females are classified as obese, one of the highest rates worldwide. This is perhaps the most significant of the risk factors associated with the high prevalence of DM in the Kingdom (El Hazmi, 1998).

As mentioned above, **diet** is widely recognised as another important risk factor of T2DM, as a high intake of carbohydrate and fats is strongly associated with obesity and consequent higher risk for T2DM (Hussain et al., 2007, Calixto and Anaya, 2014, Espelt et al., 2011). Poor diet and obesity during early adulthood often results in diabetes in later life (Connolly et al., 2000, Sommer et al., 2015). Such unhealthy diets became more widespread in the MENA countries as they obtained more calories from high fat and protein diets due to an increase in the availability of meat products (Mehio Sibai et al., 2010). The economic boom resulting from oil production in Kuwait and Saudi Arabia contributed to this by driving the adoption of a western lifestyle. This is a major factor that has influenced dietary habits and physical activity and driven a decline in production of cereal products, fruit, and vegetables. This illustrates how behavioural modification has played a key role in the nutritional transition away from a traditional healthy diet.

Increases in **smoking** and changes in smoking habits are further examples of behavioural factors associated with an increase in the prevalence of T2DM (Connolly et al., 2000, Sommer et al., 2015, Espelt et al., 2011). There is a high prevalence of smoking across the MENA region (Mehio Sibai et al., 2010), which peaks at 55.8% in the middle-aged males. In addition to a high prevalence of smoking, GCC countries such as Bahrain, Kuwait, Qatar and the UAE suffer from similar high levels of physical inactivity, particularly among females (Mehio Sibai et al., 2010). In Saudi Arabia, however, the prevalence of smoking is low compared to other countries in the region (Dar-Odeh and Abu-Hammad, 2011), perhaps because there is greater awareness of the risks. However, there is a high prevalence of physical inactivity and obesity in the Kingdom, both of which are associated with high risk of T2DM.

For controlled diabetes, the converse of the role of **obesity** as a risk factor for controlled T2DM is that achieving weight loss (reduced body mass index) is strongly associated with control of diabetes. Uncontrolled diabetes patients showing poor glycaemic control were more likely to have a high BMI value ≥ 25 in comparison to controlled diabetes patients (Bae et al., 2016, Khattab et al., 2010, Ajlouni et al., 2008, Costanian et al., 2014). A study in Pakistan reported that obese patients tended to have irregular meal patterns leading to body-insulin instability and poor control of diabetes (Ercan and Kiziltan, 2013). Alramadan et al. (2018) reported a high prevalence of poor glycaemic control among patients with high waist-hip ratio.

However, other reports observed that body weight was not associated with HbA1c status and noted that not all high-BMI individuals had T2DM (Anari et al., 2016, Alramadan et al., 2018). Nonetheless, the bulk of studies support the inclusion of BMI in evaluation of an individual's health status. In conclusion, most studies have indicated that socio-economic inequality can influence health in general and prevalence of T2DM in particular. However, the relationship cannot be explained without considering the significance of behavioural risk factors such as smoking. Key lifestyle and behavioural risk factors are likely to be unequally distributed among different socio-economic groups in ways that might be unpredictable and that might vary between countries in response to different cultural norms.

2.8.4 Socio-economic factors

This subsection will shed light on the socio-economic risk factors for the development, diagnosis, and inadequate control of T2DM to address the thesis research questions. Some studies investigated the socio-economic as factors and while others considered them as indices. I review studies that have investigated the relationship between socio-economic factors and T2DM, particularly those that are the focus of this thesis.

2.8.4.1 Socio-economic factors and the risk of developing T2DM

Several research studies on T2DM prevalence carried out over the last thirty years have shown a significant relationship between socio-economic factors and T2DM. International studies in the USA and Germany have also observed an inverse association between SES and the prevalence of T2DM (Lidfeldt et al., 2007, Reisig et al., 2007). A review of global diabetes prevalence identified population size and current income status are important factors (Whiting et al., 2011). This study showed an inverse relationship between diabetes and income status, underpinning projections showing that the greatest increase of diabetes is likely to be in low-income countries (Whiting et al., 2011). Another systematic review of studies examining the association between socio-economic factors and T2DM incidence in Europe concluded that there was an inverse relationship between SES and diabetes (Espelt et al., 2011). However, the influence of this association was more marked in women than in men, with respect to both prevalence and mortality. In addition, socio-economic factors are more associated with diabetes among middle-aged groups (40 – 69 years).

Previous published studies highlight that the effect of socio-economic factors on diabetes prevalence are conditioned by the income level of the country where the study was conducted. Some studies showed an inverse relationship between socio-economic factors (education level, occupation status and income status) and T2DM. For example, cross-sectional studies attempted to identify the main risk factors behind the high prevalence of T2DM in the MENA region (Ajlouni et al., 2008, Costanian et al., 2014). One point to emerge from these studies was high-educated individuals were less likely to have diabetes compared to low-educated individuals (Hwang and Shon, 2014, Su et al., 2016). This might be because better education and financial resources resulted in improved access to health care. Additionally, divorced or widowed individuals were more likely to be diabetic (Costanian et al., 2014). The odds of having diabetes increased with obesity and a family history of diabetes. Conversely, two studies in India and Bangladesh found that high SES individuals had a higher prevalence of diabetes compared with low SES individuals (Corsi and Subramanian, 2012, Rahman et al., 2015). This implies that the association between socio-economic factors and diabetes could differ based on country income level.

Supporting the findings of Ajlouni and Costanian, Rabi et al. (2006) investigated the association between income status and T2DM in Canada. The authors found an inverse association between diabetes and income status. This association was more obvious in females with low-income status even after correcting for BMI and physical activity (Winkleby and Cubbin, 2003). Existing literature examining the relationship between socio-economic factors and diabetes also focused on gender differences. The results broadly indicate that females show a stronger inverse association between socio-economic factors and T2DM, particularly in the middle-aged groups in Spain, Sweden, Germany and Italy (Agardh et al., 2007, Icks et al., 2007, Gnani et al., 2008, Larranaga et al., 2005). A Finnish study evaluated that the association between income status and diabetes was most noticeable in the poorest women and richest men aged for those aged 70 years or over (Hiltunen, 2005).

A Korean study to evaluate the relationship between the prevalence of diabetes and SES based on household income and educational level (Kim et al., 2015), found that the relationship between educational level and diabetes was significant in both unadjusted and adjusted models. Additionally, this association was more obvious among females aged over 30, but again indicated that females with low SES are more likely to be diabetic than comparable males. Similarly, Collier et al. (2015) investigated gender-related differences in the relationship between socio-economic factors and T2DM in Scotland, and showed an inverse association between the prevalence of diabetes and level of deprivation (as a measure of SES) for women (Andersen et al., 2008).

In terms of individual occupation status, individuals with high occupational status have a lower risk of diabetes compared to low occupation status individuals in Japan and US (Nagaya et al., 2006, Kumari et al., 2004). A US population-based epidemiological study found a similar relationship among women (Robbins et al., 2005). Similar to the findings from the diabetes incidence studies, prevalence studies also reveal an inverse association between occupation status and diabetes (Agardh et al., 2004, Lawlor et al., 2007, Geyer et al., 2004).

With regard to the GCC region, limited studies have investigated the association between socio-economic factors and the prevalence of T2DM. Three such studies used education level, marital status, and occupation. Two of these observed that education level, family history and parity are associated with prevalence of diabetes (Al-Moosa et al., 2006, Al-Mahroos and McKeigue, 1998). These findings are in line with two Saudi studies showing that low-educated individuals have a higher prevalence of diabetes in Eastern region and Jeddah city, Western region (Murad et al., 2014, Al-Baghli et al., 2010). These results indicate that the high-risk groups for diabetes were first-degree relatives of diabetics, those who are or were previously married, the unemployed, and homemakers on a low monthly income. Additionally, one study argued that urbanisation level could be an important factor related to diabetes prevalence (Al-Moosa et al., 2006). Urban residences have higher prevalence of diabetes than rural residences in Oman and Saudi Arabia (Al-Nuaim, 1997, Al-Nozha et al., 2004, Al-Rubeaan et al., 2015). Studies conducted in Scotland and Spain found that individuals who lived in deprived areas had higher prevalence of diabetes (Evans et al., 2000, Larranaga et al., 2005). This evidence gives an indication that the urbanisation level of the place of residence can be significantly related to T2DM prevalence.

In an Omani study, the odds of having diabetes were increased in individuals who were married, of advanced age, or who had high waist circumference. The authors in this study proposed that low-educated individuals were less likely to be aware of diabetes and its risk factors (Al-Moosa et al., 2006). Key factors behind variation in the prevalence of diabetes were differences in physical activity and dietary patterns and internal migration from rural to urban regions. These factors were driven by rapid socio-economic growth in Oman that had contributed to an increase of car ownership, smoking and consumption of foods high in fat and calories (Al-Moosa et al., 2006). This study indicates that socio-economic factors and urbanisation level are likely to be important determinants of diabetes risk across the GCC region.

The authors demonstrated that the strength of the association between SES and diabetes was highly sensitive to adjustment for obesity, smoking habits, and healthcare access. This highlights the importance of behavioural and health care access factors as contributors to the association between socio-economic factors and diabetes. Most studies discussed above have shown that high SES individuals were less likely to be obese and more likely to have a healthy lifestyle. These factors will be discussed in more detail in the following sections. Additionally, the above findings strongly suggest that socio-economic factors are important secondary risk factors contributing to the increasing prevalence of diabetes and related outcomes. To conclude, previous studies indicated that socio-economic factors are important contributors to the risk of developing diabetes. It appears that the development of T2DM is associated with socio-economic status (SES) irrespective of whether it is measured by education level, occupation status or income. In the next section, I critically evaluate the association between socio-economic risk factors and diagnosis of T2DM, a key challenge in managing the condition to improve outcomes.

2.8.4.2 Socio-economic factors and the risk of late diagnosis of T2DM

The prevalence of diagnosed T2DM is high among individuals with low SES, older people, and in some ethnic groups. However, few studies have investigated the association between socio-economic factors and undiagnosed T2DM, defined as individuals with no previous history of diabetes who are found to have high blood sugar level in a population survey. Undiagnosed DM is a major public health issue, as it contributes to the development of morbidity and the associated personal and societal costs. The early screening and detection of diabetes is therefore a high priority (Wilder et al., 2005). Globally the prevalence of undiagnosed diabetes has been estimated by analysing the results of blood glucose tests carried out in the course of national nutrition or health surveys (see Wilder 2005, Beagley 2014). Undiagnosed T2DM accounts for nearly 50% of the 34 million people currently affected with diabetes in the EMR region (Majeed et al., 2014). Further investigations are therefore urgently needed to identify the major causes behind this high levels of undiagnosed diabetes. Mehio Sibai et al. (2010) carried out a review of studies on CVD risk factors in the MENA Region and reported that there is high proportion of undiagnosed diabetes. In particular, they identified that at 27.9%, Saudi Arabia had the highest rate of undiagnosed diabetes in the region. Thus, individuals were not aware of their diabetes, despite the accessibility of healthcare services.

Supporting this view, Al-Lawati et al. (2002) found that two-thirds of diabetics were undiagnosed and unaware of their condition in the Omani population. However, high-income MENA countries have a lower proportion of undiagnosed diabetes than low and middle-income countries in MENA (Majeed et al., 2014). In addition, across the whole region, many people are still unaware of diabetes and its complications and routine screening is largely absent (Majeed et al., 2014). It is therefore important that future studies should investigate the factors contributing to the high prevalence of undiagnosed DM in Saudi Arabia. Previous research by Robbins et al. (2005) and Wilder et al. (2005) indicated that high SES individuals are more likely to be diagnosed with DM than low SES individuals. However, Wilder pointed out that it is important to consider household income and health insurance as factors when analysing the relationship between diagnosed diabetes and socio-economic factors. For instance, Clark and Utz (2014) clearly illustrated that high-educated individuals in the USA are more likely to participate in healthcare services and adhere to a healthy lifestyle. As a result, low SES individuals were both more likely to have DM and less likely to be diagnosed.

A study investigating the prevalence and risk factors for T2DM in Qatar found that the most significant factors associated with diagnosed T2DM were low-educational level, high BMI, a family history of diabetes, high cholesterol and occupation (Bener et al., 2009). Significantly, the study also noted that in over half of the individuals with T2DM, the disease was undiagnosed. It should be noted, however, that the strong inverse association between education level and age in the GCC region means that diagnosed individuals are more likely to be low-educated or illiterate compared to undiagnosed individuals (Bener et al., 2009). Similarly, Ajlouni et al. (2008) investigated the prevalence of T2DM and its associated factors in Jordan. The study found that only 25% of the subjects with functional diabetes were aware of their condition. Males were less likely to be diagnosed with diabetes early and most diabetic males were younger than 50 years old.

A global review of diabetes epidemiology Sicree and Shaw (2007) concluded that individuals living in urban areas are more likely to be diagnosed with diabetes than those are living in rural areas. A previous Saudi study found that females are more likely to be diagnosed with T2DM in both rural and urban areas (Al-Nuaim, 1997). One possible explanation of this finding might be provided by a study showing that people living in disadvantaged areas in Canada are less likely to have routine blood or urine glucose tests (Tang et al., 2003).

A recent study showed that generally the prevalence of undiagnosed diabetes is lower in high-income countries due to the better quality of healthcare, however this protection might not extend to individuals who are living in rural areas, who can also suffer from a high prevalence of undiagnosed diabetes (Beagley et al., 2014). This results from regional variations in healthcare quality, showing that a high prevalence of undiagnosed diabetes can occur even in a high-income country. This perspective was discussed in Chapter 2 and highlights the importance of urbanisation level as a factor influencing the diagnosis of diabetes.

In conclusion, undiagnosed diabetes is now a significant public health issue. Reducing the burden of T2DM is a national priority and should be high on the policy agenda in Saudi Arabia. One important component of a strategy to tackle the problem will be population screening of individuals to ensure early detection of diabetes (Clark and Utz, 2014, Wilder et al., 2005). The efficient targeting of resources in such a screening programme will require the use of risk scores in clinical practice, particularly in high-risk diabetes countries as Saudi Arabia. In Chapter 7, I review diabetes risk score models for undiagnosed diabetes and risk factors in more detail.

Table 2.9 Summary of socio-economic factors and T2DM published studies in the GCC and different income levels countries

Country income level	Setting and data collection year	Design and sample size	Socio-economic factors	T2DM Outcome measure	Exposure measures used to adjust model	Summary of key findings or association estimates	Methods	Sex and age groups
GCC countries								
Abdella et al. (1998)	Kuwait (1995)	Cross-sectional N=3,003	---	Measured	Age, BMI, Systolic blood pressure, diastolic blood pressure, cholesterol, triglycerides, parental history of diabetes	The prevalence of DM=14.8%, a strong association between DM and age, obesity, family history, hypertension, cholesterol, triglycerides, and parental history	Multiple regression analysis and multivariate linear regression analysis	age ≥ 20 Males and females
Al-Lawati et al. (2002)	Oman (2000)	Cross-sectional N=5,838	---	Measured	Age, sex and residence place.	The prevalence of DM= 16%, high prevalence in urban, and males. Men =16%; women=15.4%	Two-sample t-test	age ≥ 20 Males and females
Al-Mahroos and McKeigue (1998)	Bahrain (1991)	Cross-sectional N=2,128	Educational level	Measured	Family history of diabetes, cholesterol, waist circumference, ethnicity, BMI, physical activity.	High prevalence of diabetes in low-educated, family history, high parity in women.	Multivariate logistic regression analysis	(40 - 69) Males and females
Al-Moosa et al. (2006)	Oman (2000)	Cross-sectional N=7,179	Marital status Educational level	Measured	Age, residence, waist circumference, systolic blood pressure.	The prevalence was 17.7% in Muscat compared to rural 10.5%. Urban area OR=1.7 (1.4 – 2.1); rural area OR=0.6 (0.4 – 0.9). Education related to the prevalence of diabetes.	Multiple logistic regression	age ≥ 20 Males and females
Bener et al. (2009)	Qatar (2007 - 2008)	Cross-sectional N=1,434	Educational level Occupation status	Measured	Age, family history of diabetes, BMI, physical activity, smoking habits, meals per day.	The prevalence was 16.7%; 53.2% in Qatari women and 31.2% in Qatari men. The DM was related to obesity, HTN, smoking and family history.	Multiple logistic regression	(20 - 59) nationals Males and females
Saadi et al. (2007)	UAE (2005 - 2006)	Random sample in Al-Ain city N=2,396	---	Measured	Age, BMI, waist circumference, hypertension.	The prevalence of DM = 10.2%. Only age and BMI.	Logistic regression analysis	Females and males Both ≥ 18
High-income								
Ezeamama et al. (2006)	USA 4-year follow-up	Cohort N=1,289	Educational level	Self-reported	Age, sex, smoking status and un-employment.	Inverse association with SES	Multivariate logistic regression	25 – 58 Not combined gender
Kumari et al. (2004)	United Kingdom (1985-1999)	Cohort N=10,308	Occupation status	Self-reported	Age, follow-up length, ethnicity, diabetes family history, height, BMI, systolic blood pressure, smoking status	Occupation status showed an inverse association with DM risk.	Cox regression analysis	Males (35 - 55)

Table 2.9 Summary of socio-economic factors and T2DM published studies in the GCC and different income levels countries

Country income level	Setting and data collection year	Design and sample size	Socio-economic factors	T2DM Outcome measure	Exposure measures used to adjust model	Summary of key findings or association estimates	Methods	Sex and age groups
High-income								
Nagaya et al. (2006)	Japan (1988 - 1991)	Cohort N=5,130	Occupation status	Measured	Age, sex, and survey year Age, BMI, drinking, smoking status, physical activity, and educational level	High occupation status has lower risk of diabetes incidence of being diabetic compared to low occupation status.	Cox's proportional Hazard models	Males (30 - 49)
Robbins et al. (2005)	USA (1971 – 1992)	Cohort N=11,069	Educational level, income, and occupation	Medical records		Low diabetes incidence with high income and education. No inverse association between diabetes incidence and occupation status.	Cox's Proportional hazard models	Males (29 -84)
								Females (29 -84)
Su et al. (2016)	China (2013-2014)	Cross-sectional N=5,532	Educational level	Measured	Sex, Age, ethnicity, medical services accessibility, annual household income	Higher educated level associated with higher awareness, treatment, adhere medications and monitoring of blood glucose not only individuals' socio-economic factors, but contextual SE factors.	Multilevel logistic regression	Males and females age (≥ 35)
Tang et al. (2003)	Canada (1996 – 1997)	Cross-sectional N=39,021	Education attainment and income adequacy	Self-reported	Age, sex, residence, BMI, physical activity	The prevalence of DM increases with decreasing income and education in both females and males. The prevalence of DM: 6.6% in men and 5.5% in women.	Multiple logistic regression	Males and females age (≥ 40) Not combined
Hwang and Shon (2014)	Korea (2010 – 2012)	Cross-sectional N=14,330	Education level income	Self-reported, measured and diagnosed	Age, sex, marital status, region, house ownership, BMI, physical activity, smoking status, alcohol consumption.	High prevalence of diabetes mellitus in low educated and income status individuals.	Logistic regression analysis	Combined and not combined models for males and females; age ≥ 30
Evans et al. (2000)	Scotland (1993)	Cross-sectional (N=366,849)	Deprivation index	Register records pre-diagnosed	Age	Higher prevalence of diabetes among deprived areas.	Logistic regression	35+ Men and women (jointly)
Espelt et al. (2008)	10 European countries (2000)	Cross-sectional (N=137,636)	Education level	Self-reported	Age	The SEP inequalities observed higher in women than men in some countries.	Poisson regression model for DM mortality Log-binomial model for DM morbidity	30 – 64 years Men and women (separately)
Geyer et al. (2004)	Germany (1995)	Cross-sectional (N = 124,151)	Occupation status	Hospital records, medication, and anti-diabetic	Age and sex	There is a relationship between occupation and DM.	Logistic regression	20+ Men and women (separately)

Table 2.9 Summary of socio-economic factors and T2DM published studies in the GCC and different income levels countries

Country income level	Setting and data collection year	Design and sample size	Socio-economic factors	T2DM Outcome measure	Exposure measures used to adjust model	Summary of key findings or association estimates	Methods	Sex and age groups
<i>High-income</i>								
Agardh et al. (2004)	Sweden (1992-1994) (1996 – 1998)	Cross-sectional (N = 8,108)	Occupation status	Measured (OGTT)	Age, BMI, smoking, activity, family history, coherence sense, decision-making at work	There is a relationship between occupation and DM (higher in women).	Multiple logistic regression	35 – 56 years Men and women (separately)
Escolar (2009)	Spain (2003)	Cross-sectional (N=6,708)	Self-perceived economic difficulties	Self-reported	Age, BMI, physical activity	Only women with economic difficulties was significantly associated with diabetes, even after adjusted by BMI and activity.	Regression analysis	16 years and over Women and men.
Hiltunen (2005)	Finland (1991)	Cross-sectional (N=379)	Education, occupation, income, and economic status	Diabetes treatment and OGTT	age, marital status	Income only associated with DM (highest-income men and lowest-income women).	Logistic regression	70+ Women and men
Lawlor et al. (2007)	United Kingdom (1999 – 2001)	Cross-sectional (N=4,277)	Adult: education, occupation. Childhood: father's occupation, household amenities	Diagnosed: self-reported Undiagnosed: Glucose plasma test	Age, smoking, BMI, activity, waist-hip ratio, blood pressure.	Education and occupation associated with diagnosed DM, and no differences for undiagnosed DM.	Logistic regression	60 to 79 years women
Larranaga et al. (2005)	Spain (2000)	Cross-sectional (N = 65,651)	Area deprivation index	Medical records	Age	SEP inequalities associated with T2DM and observed among women than men.	Logistic regression	24 + Men and women (separately).
Gnavi et al. (2008)	Italy (2003)	Cross-sectional (N=897,743)	Education level and income	Individuals with diabetes, medical and hospital registers	Age	There is relationship between diabetes and SEP and higher within women and individuals aged <65 years old.	Log-linear models	Men and women (separately) age 20+
Andersen et al. (2008)	United Kingdom (23 towns) (1999 – 2001)	Cross-sectional (N=4,286)	Deprivation index Individual life-course SEP	Self-reported Medical records	Smoking, age, activity, obesity, waist-hip ratio, cholesterol level, triglycerides, and hypertension.	There is a relationship between T2DM and deprivation level even after adjusted with individual life-course SEP.	Regression analysis	Women aged 60 - 79
Agardh et al. (2007)	Stockholm (1996 – 1998)	Cross-sectional (N = 7,949)	Education level and occupation, father's occupation	OGTT Measured	Physical activity, smoking status, BMI, family history of diabetes, coherence sense, decision-making at work	Father's occupation not associated with diabetes, occupation status associated with DM for both sexes, education level appeared in women only.	Logistic regression	Men and women (separately) 35 - 56
Icks et al. (2007)	Germany (1992 – 1998)	Cross-sectional (N = 7,466 and 5,583)	Education level	Self-reported	Age and BMI	The relationship between education and diabetes appeared only among women.	Multiple logistic regression	Men and women (separately) 25 – 69

Table 2.9 Summary of socio-economic factors and T2DM published studies in the GCC and different income levels countries

Country income level	Setting and data collection year	Design and sample size	Socio-economic factors	T2DM Outcome measure	Exposure measures used to adjust model	Summary of key findings or association estimates	Methods	Sex and age groups
High-income								
Wändell and Gåfvels (2004)	Sweden (2001)	Cross-sectional (N = 72,347)	SEP	Electronic records	Sex	High prevalence of diabetes among deprivation areas.	Logistic regression	Men and women (separately) 35 - 64
Connolly et al. (2000)	United Kingdom (1994)	Cross-sectional (N = 285,157)	Deprivation index	WHO criteria	Interactions with age and sex	The prevalence DM inversely associated with deprivation index		Whole population Men and women (jointly)
Middle-income								
Ajlouni et al. (2008)	Jordan (2004)	Cross-sectional N=1,121	Educational level	Measured	Age, sex, BMI, family history of diabetes	The prevalence of DM = 17.1% and 54% uncontrolled diabetes. Higher level of education was associated with a decrease in the odds of having diabetes.	Multiple logistic regression	Age (≥ 25) Males and females
Corsi and Subramanian (2012)	India (2005 - 2006)	Cross-sectional N=168,135	Educational level Household wealth Social caste	Self-reported	Age, sex, marital status, religion, BMI, residence	Highest SES group appear to be at greatest risk for diabetes. There is substantial geographic heterogeneity in the prevalence of DM.	Multilevel logistic regression models	Females (18-49) Males (18 - 54)
Costanian et al. (2014)	Lebanon (2008 - 2009)	Cross-sectional N=2,195	Marital status Educational level	Self-reported	Age, sex, insurance coverage, smoking status, alcohol consumption, physical activity, BMI, family history of diabetes	The prevalence of DM was 8.5%. Males, high-educated, older-age groups married found controlled diabetes compared to others.	Logistic regression analysis	Age (≥ 25) Males and females
Low-income								
Rahman et al. (2015)	Bangladesh	Cross-sectional	Employment status Marital status	Measured	Age, sex, residence place, family size, number of adults in the household who were ≥ 35, smoking status, BMI	The prevalence of DM = 10.8%; 7.2% in low SES, and 17.9% in high ESE. The prevalence of DM was more prevalent among high SES groups.	Multivariate logistic regression	Females and males Age (≥ 35)

2.8.4.3 Socio-economic associations with inadequate control of T2DM

To date, most research into T2DM treatment has focused on better ways to manage high levels of blood glucose. However, it could be argued that this should be complemented by an increased emphasis on understanding the socio-economic factors behind the high prevalence of undiagnosed and uncontrolled T2DM. This section will review studies that investigated the association between socio-economic factors and the prevalence of uncontrolled diabetes. Subjects with controlled diabetes are typically defined as individuals who are aware of their diagnosis but who have HbA1c < 7.0%. With regard to socio-economic factors, in high-income countries, individuals in low SES groups were more likely to have uncontrolled T2DM (Fisher et al., 2010). Other studies also observed that individuals in high SES groups are more likely to have controlled diabetes (Bi et al., 2010, Houle et al., 2016, Alramadan et al., 2018, Al-Baghli et al., 2010). D'Souza et al. (2015), however, reported that in Oman, high educated individuals (i.e. who had completed diploma/technical) showed poorer glycaemic control compared to low educated diabetics (i.e. only completed high school or lower). This Omani study also observed that interventions involving empowerment and self-management increased the likelihood of good glycaemic control.

Empowerment level is another significant factor associated with education level (Tol et al., 2013). As previously discussed, education level is associated with individuals better understanding their condition and achieving improved glycaemic control. Other studies observed that approaches based on empowerment contributed to improve self-management among T2DM patients (Tol et al., 2013, Schmuhl et al., 2014, Pulvirenti et al., 2014). This approach has been evaluated in a randomised/controlled clinical trial and shown to improve self-management and glycaemic control (Chen et al., 2015). In conclusion, therefore, education level is an important factor affecting the control of T2DM and low education is a risk factor for uncontrolled diabetes. However several other studies reported that low-educated individuals were more likely to have uncontrolled diabetes (Khattab et al., 2010, Ajlouni et al., 2008, Costanian et al., 2014), two UAE studies did not find any association between education level and glycaemic control (Al-Kaabi et al., 2008, Al-Kaabi et al., 2009). Two Saudi studies have investigated the socio-economic risk factors of inadequate diabetes control. The first study in the Eastern province observed better glycaemic control among individuals who were single and those in professional employment compared to those who were widowed, poor and self-employed (Al-Baghli et al., 2010). This study also observed better glycaemic control in urban areas compared to rural areas.

A more recent Saudi study Alramadan et al. (2018) investigated the risk factors of inadequate control among T2DM patients in three cities (Riyadh, Jeddah, and Al-Hofuf). This revealed poor glycaemic control among patients living in remote areas, housewives and the unemployed. Glycaemic control was not associated with nationality, income, and region (Alramadan et al., 2018), however these three cities are classified as advantaged areas in SA. To conclude, individuals with low SES are more likely to have uncontrolled diabetes compared to high-SES individuals. One reason for this is presumably that low-income individuals are less likely to have access to advanced health care services and routine screening for diabetes within health care centres.

Such variations in healthcare services provision, together with differences in access to healthy food and exercise facilities, highlight the importance of socio-economic factors in understanding uncontrolled diabetes (Stringhini et al., 2012). These authors argued for managing T2DM in a wider context, as they concluded that modifiable risk factors such as obesity and health behaviours account for about half the difference in T2DM prevalence between groups with high and low SES. A systematic review by Grintsova et al. (2014) noted that individual SES could affect adherence to treatment as well as healthcare access. The authors observed that the quality of diabetes care services was associated with SES, as services were usually worse in deprived areas. They concluded that it was important to consider differences at the level of the individual as well as regional provision. Assessing SES at the individual level highlights problems such as poor social support (Grintsova et al., 2014). At the same time, it is important to identify differences in resource availability between regions, for example the availability of specialised physicians.

Additionally, most medical studies have focused on investigating the individual socio-economic factors that lead to better self-management of DM (Clark and Utz, 2014). These authors argue that this narrow focus neglects wider factors such as environmental factors and cultural conditions, access to health care, education, employment, community, and social support. However, a 2011 review considered that the quality of health care services and blood sugar control were the most important factors influencing the association between diabetes mortality and socio-economic factors (Espelt et al., 2011). In the next section, I review studies examining the association between diabetes outcomes and non-modifiable risk factors such as age, gender, ethnicity, and family history.

In summary, studies have used various measures of SES such as: education status and occupation status separately or combined as adulthood socio-economic status, deprivation index by geographical location, father's occupation to define childhood socio-economic status, marital status, simple employment status, income status and household income status. The previous table is a summary of T2DM and socio-economic factors studies in GCC countries and international comparator countries with different income-levels (**Table 2.9**). Of all the socio-economic indicators, education level was the most commonly observed and significant in terms of association with diabetes. However, education level is essentially fixed at adulthood, it could impact on health status in later life. Furthermore, education levels within a large birth cohort might differ depending on the age of the subjects reflecting economic development. In this thesis, I have used both measured and self-reported outcomes as an inclusive measure of diabetes status to examine the association with SES. I chose education level as the primary socio-economic factor based on a review of the literature and its availability in the dataset.

2.8.5 Environmental and community factors

In this section, I discuss the impact of environmental factors such as climate and geography on the prevalence of diabetes. In a review of the impact of SES on autoimmune diseases (Calixto and Anaya, 2014), the authors explain that environmental factors are important determinants of health outcomes in different populations. They note that geographical differences, local cultures, and mobility are significant determinants of health that should be taken into consideration. Some environmental factors interact with sociocultural barriers. For example, a lack of public exercise facilities, hot climate and the restriction of employment opportunities for women could contribute to high prevalence of diabetes in the wider MENA region and the GCC region in particular (Mehio Sibai et al., 2010). Thus, environmental factors should be regarded as significant risk factors for T2DM and Saudi Arabia. The absence of an appropriate environment for female physical activity contributes to the high prevalence of obesity.

This effect is compounded by nutritional changes caused by a lack of healthy food shops coupled with ready access to fast food. We need a more detailed understanding of the association between health and SES if we are to address the problem of T2DM in Saudi Arabia. This perspective brings a focus on key issues such as quality of health services, transportation, and access to healthy food. For example, the influence of inequality on health outcomes might reflect a critical lack of human resources in the health care system. Brown et al. (2004) and Sommer et al. (2015) note that SES is a surrogate measure reflecting the quality and availability of healthcare services in different regions, coupled with access to transportation and exercise facilities.

The effect of regional variation or individual place of residence on diabetes outcomes both reflected and provided evidence that the diversity of socio-economic status between regions reflects disparities in access to and/or quality of resources such as healthcare and educational institutions. This inequality drives differences in health outcomes between individuals. The previous section evaluated the association between socio-economic factors and T2DM, independent of demographic, lifestyle, health care access, and self-care factors. As previously discussed, several studies support education level as the most important socio-economic risk factor for T2DM. Other significant contributory risk factors are occupation status, employment status and household income. Here, I present an overview of the social determinants of health and health inequality theories to highlight the importance of the association between socio-economic risk factors and health outcomes. I then synthesise this analysis into a framework to guide the methodology for analytical chapters.

2.9 Summarising the literature review: Conceptual and analytical frameworks

Overall, my review of the evidence shows that little is known about the association between socio-economic factors and T2DM in the Saudi Arabian context. Education level, employment/working status and monthly household income are the available socio-economic factors in our dataset. Of these, I chose to focus on education level as the primary socio-economic factor in my analysis as educational attainment levels have increased significantly over recent decades in the country (see section 3.3.3 in the following chapter). I hypothesised that the increase of educational attainment could be associated with the prevalence, diagnosis and controlled T2DM.

However, I also used employment/working status and monthly household income as secondary socio-economic predictors. In addition, gender is a strong modifier of the risk associated with most socio-economic factors and therefore a key variable to include in any analysis. The following table summarises all the possible risk factors that I identified from the literature.

Demographic factors	Age, sex, marital status.
Socio-economic factors	Education level, occupation status, employment status, income status, household family sized, annual household income.
Environmental and Community factors	Administrative region, residence place (urbanisation), Weather conditions, geographical location, insufficient mobility, availability of public and walking places, access to exercise facilities and healthy food shops.
Genetic factors	Family history of diabetes, and ethnicity.
Medical factors	Waist-ratio circumference, obesity (Body mass index), hypertension, cholesterol level, and triglycerides.
Lifestyle factors	Physical activity, diet, meal per day, and smoking habit.
Socio-cultural factors	Social norms, traditions and support factors, religious beliefs, empowerment level that effect on lifestyle and behaviour.
Healthcare access factors	Insurance coverage, accessibility to medical services, routine check-up, and visit for manage diabetes.
Self-care behavioural factors	Test blood sugar at home, loss weight (BMI), prescribed diet, and physical activity.
Self-care factors	Use of medications/ anti-diabetic, and insulin.

It is an inescapable fact that the high prevalence of T2DM in Saudi Arabia is related to socio-economic change. I decided to investigate this association in more detail by applying three frameworks based on my review of the literature relating to the risk factors for three different outcome measures. These frameworks address the four main sets of factors known to play an important role in the development of T2DM and that are available as variables in the SHIS dataset. These are demographic, lifestyle, health care access and self-care factors. In this section, I present the conceptual frameworks that I developed after reviewing the literature and previous analytical approaches. Based on the discussion presented above, the prevalence of diabetes is primarily dependent on demographic, community, and lifestyle factors. As an outcome, I consider T2DM at three different levels: prevalence (whether an individual has uncontrolled blood glucose); diagnosis (whether or not an individual has received a diagnosis of diabetes); and control (whether an individual treated for diabetes has regained control of blood glucose levels).

These outcomes are influenced by multiple factors that **associate** with each other. The first framework (Figure 2.7) shows the proposed relationship between education level and overall prevalence of T2DM (whether diagnosed or undiagnosed). The primary factors - those that contribute directly to the prevalence of T2DM - are generally grouped into four sets of factors (genetics, lifestyle, obesity, and demographic factors). In most cases, two or more of these interact to influence the likelihood and prevalence of T2DM. These primary factors are influenced in turn by secondary interdependent factors such as education level, geographic region, and socio-economic status.

2.9.1 The prevalence of T2DM framework

From the left-hand-side of (**Figure 2.7**), the second box indicate how an individual's education level can be associated with the prevalence of T2DM (Kim et al., 2015). For example, low-educated individuals are more likely to be diabetic compared to high-educated individuals. However, this relationship could be also influenced and associated with other factors such as demographic, socioeconomic, community and lifestyle factors. For example, the framework shows that education level has a significant influence on individual lifestyle behaviour.

These lifestyle factors also influence the prevalence of T2DM. For example, Saudi Arabia has faced rapid socio-economic development leading to an increasingly sedentary lifestyle compounded by dietary changes (Al-Rasheedi, 2015). These changes in turn have increased the risk of diabetes. For demographic factors such as gender, age, and marital status can be considered determining individual diabetes status. For example, older age groups are more likely to be diabetic compared to younger age groups (Al-Rubeaan, 2015). Hence, the association between education level and the prevalence of T2DM via demographic and lifestyle factors can lead to differences in diabetes risk status. For instance, individuals that are better-educated and/or younger are less likely to be diabetic than those who are older or educated to a lower level.

Additionally, socio-economic factors such as employment status and household income have a significant influence on the association between education level and the prevalence of T2DM. For example, having a high socioeconomic status reduces the likelihood of an individual being diabetic (Murad et al., 2014, Al-Baghli et al., 2010). Similar associations occur between community factors (such as place of residence, urbanisation level and region) and lifestyle that further affect the likelihood of T2DM in individuals. For example, regional differences in socio-economic level can affect the availability and quality of health care services (Sommer et al., 2015).

Genetic factors also influence the risk of T2DM, with a family history of disease an important predictor of risk (Warsy and El Hazmi, 1999). Overall, however, the prevalence T2DM is multifactorial, although individually it might be difficult to detect a direct association between individual factors and individual risk. I will address this framework in more detail in Chapter 5.

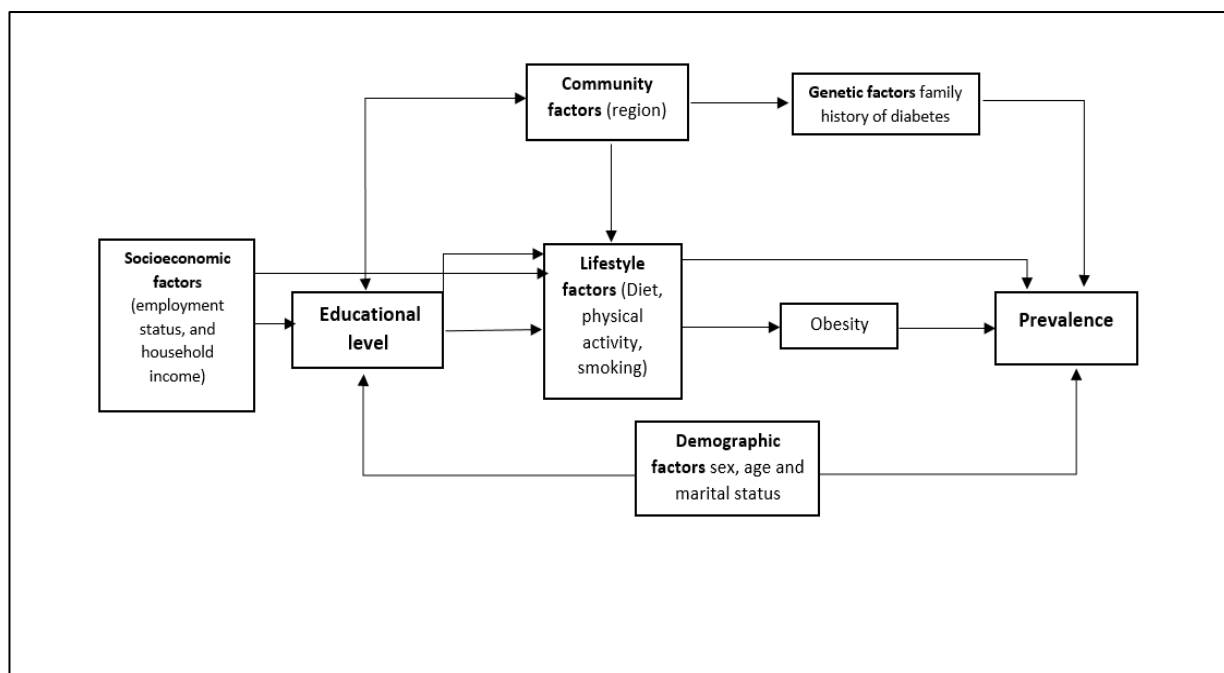


Figure 2.7 Analytical framework for the relationship between education level and development T2DM

2.9.2 The diagnosis of T2DM framework

The second framework (Figure 2.8) represents the association between education level and the diagnosis of T2DM. The hypothesis is that this is primarily related to demographic and community factors, with a contribution from socioeconomic and health care access factors. As is the case for the previous outcome (the prevalence of diabetes), a combination of these factors determine the likelihood of individual with diabetes receiving a diagnosis. For example, based on my review of the literature, well-educated individuals are less likely to have undiagnosed diabetes (Robbins et al., 2005, Wilder et al., 2005). The framework encompasses sets of factors used in previous studies identified in my review of the literature. These were added to this framework to explore the relationship between education level and diagnosis of T2DM.

The first set of factors I considered were demographic factors such as age, sex, and marital status, all of which interact in ways that could affect the association between education level and diagnosis of T2DM. For example, individuals in older age groups are more likely to be diagnosed (Bener et al., 2009), and previous literature suggests that individuals educated to a high level are more likely to be in the younger age groups. The second set of factors I considered were community factors such as an individual's place of residence, region, and geographical location. These factors affect an individual's lifestyle, as well as socioeconomic and health care access factors. For example, individuals who lived in the Southern region have a higher likelihood of undiagnosed diabetes compared to those who lived in Central region, reflecting regional diversity in socioeconomic status in Saudi Arabia.

In addition, socioeconomic factors can have a direct impact on the diagnosis of T2DM. For instance, individuals with high SES are more likely to receive advanced health care and are therefore more likely to have receive a diagnosis of diabetes. Health care access factors such as health care provider, health care facility type and the distance between home and health facility have an additional influence on the likelihood of T2DM diagnosis. These health care access factors are likely to be influenced in turn by socioeconomic and/or community factors, further influenced by demographic variables. For example, an individual's socioeconomic characteristics clearly influence their ability to access health care. This can operate directly, in terms of access to private healthcare, or indirectly via educational level. These factors will be further influenced by place of residence and regional differences in healthcare provision (AlOmar et al., 2018, El Mouzan et al., 2012). In the end, it is a combination of all these factors that determine the association between education level and diagnosis of T2DM. I address this framework in more detail in Chapters 6 and 7.

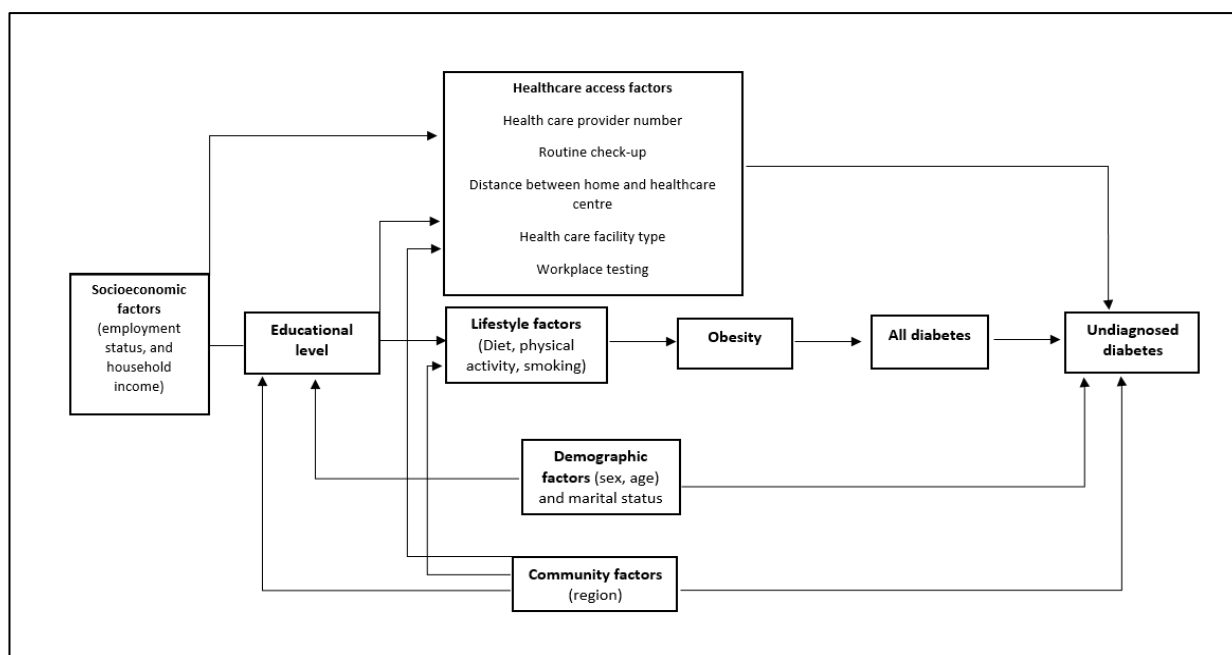


Figure 2.8 Analytical framework for the relationship between education level and diagnosis of T2DM

2.9.3 The controlled of T2DM framework

In the final framework (Figure 2.9), I consider the association between education level and control of T2DM. This framework extends the previous two frameworks to consider factors related to diabetes management, in particular the association between education level and control of T2DM. Though highly educated individuals with diabetes are more likely to achieve control of their blood sugar, this relationship could be influenced by a number of other factors. These were identified from previous literature and added to the framework. Similarly, demographic, region/community, and socioeconomic factors could also have an influence on the association between education level and control of T2DM. For example, individuals educated to a high level are more likely to exhibit self-care behaviour and have better access to health care, both of which are important determinants of blood glucose control (Fritschi et al., 2017). Additionally, patients are more likely to be in control of their blood sugar if they are young, live in an urban area, or have high socioeconomic status. I therefore included health care access and self-care factors to the framework. Self-care encompasses a range of behavioural factors, the most important of which are the ability to test blood sugar at home (Khattab et al., 2010), and the motivation to follow a weight-loss programme (BMI) involving a prescribed diet and physical activity (Alramadan et al., 2018).

Patients who manage their blood sugar, have a lower BMI, and exhibit better lifestyle behaviours are more likely to gain control of their blood sugar. These factors are also likely to be affected by other factors specific to the individual as discussed for the first two frameworks. These include internal factors related to individual motivation and self-care behaviours, as well as external factors such as quality of healthcare provision and prescribing practices (Espelt et al., 2011). Thus, the self-care behaviour of a patient is likely to be influenced by their education level, with individuals educated to a lower level less likely or less able to follow disease management plans. As a result, they are less likely to achieve control of their blood sugar.

External health care access factors are likely to be affected by the socioeconomic status of patients as well as regional factors related to policy and economics (Stringhini et al., 2012, Grintsova et al., 2014, Clark and Utz, 2014). They include self-care factors such as attendance at routine check-ups, visits for managing diabetes, and use of medication with particular emphasis on access to insulin. This final framework will be addressed in more detail in Chapter eight. Overall, I have applied these three frameworks to the analysis of the available datasets to gain insights into the association between the three T2DM outcome measures and education level. My results using each framework are discussed separately within each of the analytical Chapters (5, 6 and 8).

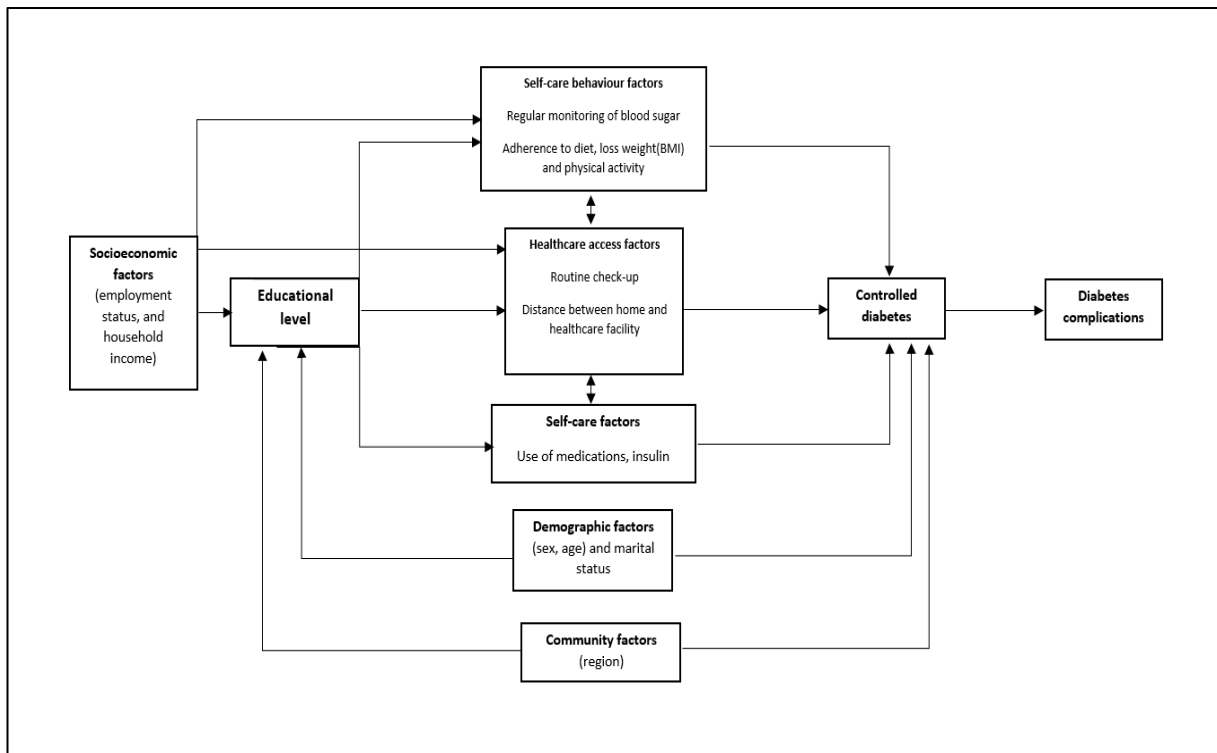


Figure 2.9 Analytical framework for the relationship between education level and inadequate control of T2DM

2.1 Conclusion

The epidemic of T2DM is causing major concern in countries across the world irrespective of income or development levels. The risk of T2DM is noticeably higher in the GCC region, particularly in Saudi Arabia. In this chapter, I reviewed the comprehensive literature relating to trends in diabetes prevalence both globally and within the MENA and GCC regions. This analysis revealed a deficit in research relating to the trends and drivers of the rapid increase in diabetes prevalence in Saudi Arabia and the wider MENA/GCC regions, as most studies of the relationship between socio-economic factors and diabetes have been conducted in European and other Western societies. However, there are few studies addressing how education level and socio-economic characteristics relate to diabetes. Dahlgren and Whitehead and Marmot underline the importance of reaching a better understanding of the complex relationship between government policy and health outcomes (Dahlgren and Whitehead, 1991, Marmot and Health, 2007). This perspective has only undergone limited investigation in the MENA and GCC regions, despite the fact that they are undergoing a period of rapid socio-economic, cultural and political change.

From my review of the literature, I conclude that the variation in diabetes outcomes between individuals is the result of complex interactions between socio-economic, demographic, lifestyle, and behavioural risk factors. A better understanding of these interactions and how they drive increasing T2DM prevalence is urgently required. Only then can we design optimal intervention strategies to reduce the incidence of T2DM, prevent its complications and minimise the societal impact. Although there are many peer-reviewed studies of the risk factors associated with T2DM in Saudi Arabia, there remains a gap in research into the association between T2DM and socio-economic factors. In this study, I address this gap by focusing on the association between education level and the three main diabetes outcomes (prevalence, diagnosis, and control) in four main analytical chapters.

Chapter 3 Socio-economic and Lifestyle Patterns of the Saudi Arabian Context and Health Care Services in Diabetes Management

3.1 Introduction

In this chapter I discuss background information relating to the Kingdom of Saudi Arabia divided into four aspects. The first provides a demographic description of the Saudi Arabian population using indicators such as population growth, urbanisation, and life expectancy. I then consider socio-economic factors such as educational attainment, employment and lifestyle factors that are relevant to the high prevalence of diabetes in the GCC region. The third aspect provides an overview of health care services in Saudi Arabia, which is followed by a more detailed consideration of the structure of the health system with particular reference to screening procedures, standards of care and official diabetes policy.

3.2 Saudi Arabia population

The Kingdom of Saudi Arabia is the largest country in the Arabian Peninsula. It is a high-income country with a Gross Domestic Product (GDP) per capita of US\$24,893 in 2013. Population growth in the Kingdom was 340% between 1974 and 2010, however there was considerable variation across geographical areas and between national and expatriate populations (Ashwan et al., 2012). With the second largest oil reserves supplying some 25% of world's oil, the Kingdom is one of the main international oil markets and a key member of the Organisation of Petroleum Exporting Countries (OPEC) (Ashwan et al., 2012). The climate in the Arabian Peninsula is hot and dry in the summer and mild but wet in the winter, similar to that of the Mediterranean region (Almazroui et al., 2012).

The Kingdom shares borders in the east with the UAE, Qatar, and Bahrain; in the north with Iraq, Jordan and Kuwait; in the south with Oman and Yemen; and the Red Sea in the west (see **Figure 3.1**). There are 13 administrative regions in Saudi Arabia resulting from a period of rapid urbanisation and economic growth: Al-Riyadh and Al-Qaseem in the central region; Makkah, Madinah and Tabuk in the western region; the Eastern Region; Northern Border, Hail and Al-Jouf in the north; and Jezan, Aseer, Najran and Al-Baha in the south.



Figure 3.1 Kingdom of Saudi Arabia map⁸

Rapid urbanisation has resulted in differences in development between regions. The high internal migration of families from rural to urban and urban to urban areas in Saudi Arabia is a key feature driving the rapid increase in the population of the four main cities: Riyadh, Medina, Jeddah and Dhahran. Overall, the proportion of individuals living in urban areas has increased from 46.0% in 1974 to 81.0% in 2004, however this percentage varies between administrative regions (Khraif, 2007). Another factor is that the Makkah region has the highest overseas and non-national population due to visiting pilgrims (Al Bassam, 2012).

⁸ https://commons.wikimedia.org/wiki/File:Saudi_Arabia_location_map.svg

The overall age distribution (irrespective of nationality) also varies significantly across the 13 Saudi regions as follows⁹. Al-Baha region had the lowest proportion of under-15-year olds (22.9%) and Al-Jouf region had the highest (29.5%). For the 15 up to 64-year-old age group, the lowest percentage was observed in Najran and Al-Jouf regions (67.6%) while the Eastern region was highest with 73.7%. The Eastern region had the lowest percentage of individuals aged over 65 (2.3%) with the highest percentage seen in Al Baha region (5.5%). Age structures are broadly the same across the regions, although some small differences.

The population age distribution for Saudi nationals also varied between regions but showed a similar pattern. As in the overall population, for individuals aged less than 15 years the lowest proportion was seen in Al Baha region (26.5%) and highest in Al-Jouf region (36.7%). In the 15 to 64 age group, the lowest percentage was again seen in Najran region (62.2%), however Makkah region was now the highest (67.7%). For those aged 65 and above, the highest percentage of Saudi individuals was again in Al-Baha region by 6.7%, but it was now lowest in Tabuk region (3.1%).

For the total Saudi population, the age dependency ratio (the number of dependants aged under 15 or over the age of 65 as a proportion of the total population) is lowest in the Eastern region (36%) and highest in Najran region (48%). Considering Saudi nationals only, the age dependency ratio¹⁰ was highest in the Al-Jouf region (67%) and lowest in Makkah region (48%). The gender distribution among Saudi nationals is broadly balanced and similar in all regions. There is, however, a significant difference in the in the gender balance among Saudi non-nationals that is also reflected in the overall gender distribution.

⁹ General Authority of Statistics (2016) https://www.stats.gov.sa/sites/default/files/en-demographic-research-2016_2.pdf [Accessed March 2020].

¹⁰ The age-dependency ratio
https://www.un.org/esa/sustdev/natlinfo/indicators/methodology_sheets/demographics/dependency_ratio.pdf

The Central Department of Statistics¹¹ (2003) estimated that the total Saudi population grew from 16.9 million in 1999 to 22.67 million in 2003 and 27.1 million in 2010. By mid-2014, the population size of Saudi Arabia was the largest among GCC¹² countries at 30 million. Projections are that this growth will continue, reaching 39.8 million by 2025 and 54.7 million people in 2050 (Almalki et al., 2011). Part of the rapid growth is explained by the high number of non-national male workers (**Figure 3.2**), with non-nationals accounting for 32.4% of the total population in 2014 according to the National Institute of Statistics (**Figure 3.3**). This also explains the preponderance of males in the total adult population (Shah, 2012).

In addition, there has been high residential mobility in the last three decades driven in part by an increase in household size and high rents (Khraif, 1994). Riyadh, Western and Eastern regions have the highest population size compared with other regions in the Kingdom. Collectively they account for two-thirds of the total Saudi population (Ashwan et al., 2012) and have lower house ownership compared to less urban regions (Salam et al., 2014). However, the urbanisation level of these three regions is significantly associated with better job opportunities, education, and health services.

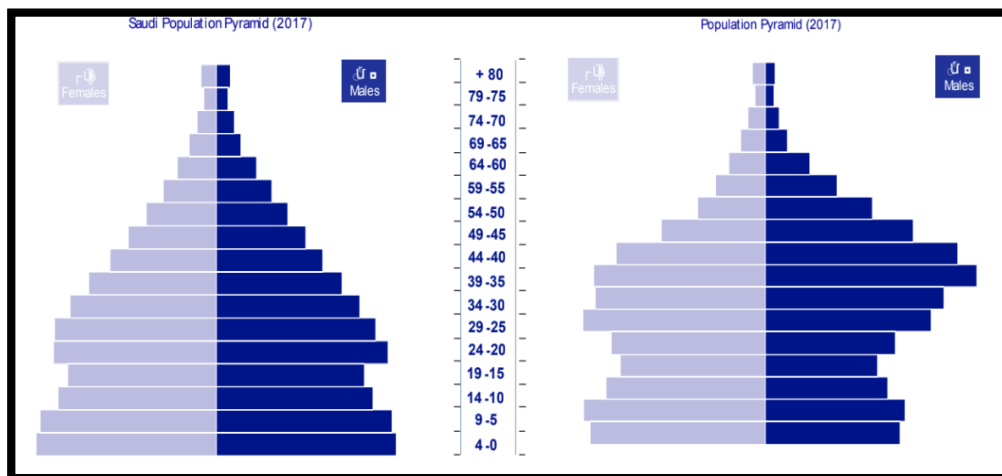


Figure 3.2 Age-sex population pyramid in 2017 ¹³

¹¹ General Authority for Statistics, Kingdom of Saudi Arabia <http://www.stats.gov.sa/en/node>

¹² GCC countries: The Cooperation Council for the Arab States of the Gulf is a cooperation established in 1981 between Saudi Arabia, Oman, Bahrain, Qatar, Kuwait and the UAE (GCC, 1981).

¹³ The General Authority of Statistics (<https://www.stats.gov.sa/en>)

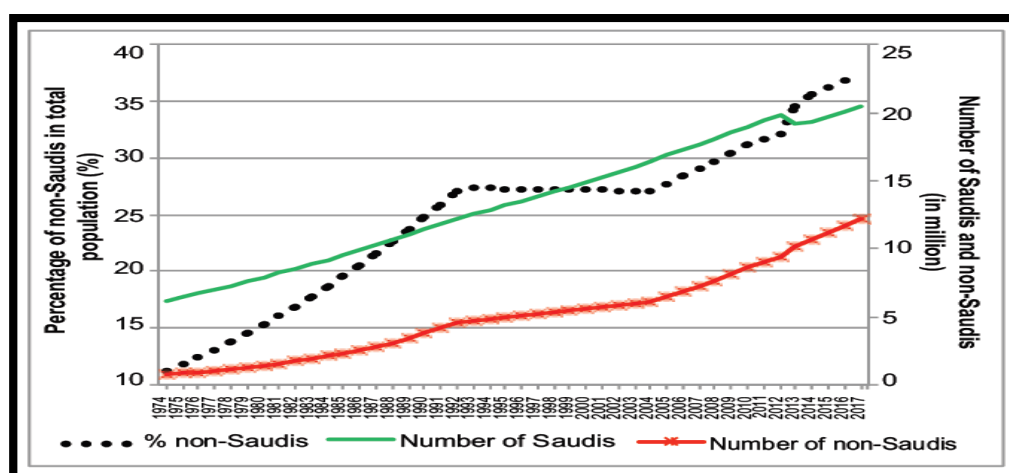


Figure 3.3 Total numbers of Saudis and non-Saudis resident population (1974 - 2017)¹⁴

The General Authority of Statistics claimed that Saudi Arabia become one of top five destination countries for migrants in 2017, adding to an estimated 12.2 million non-national population, with the majority coming from South-Asian countries (De Bel-Air, 2014). This expatriate population reached 37% of the total population by the mid of 2018 (De Bel-Air, 2018)¹⁵, however the expatriate population has since declined by about 800,000, largely due to the introduction of controls on dependents, fees and new taxation measures.

There are two main reasons for the large expatriate community in Saudi Arabia. The first is the influx of Muslims from around the world on pilgrimages to the holy places of Makkah and Medina (2.4 million in 2018). The second reason is more structural, resulting from immigration in response to oil exploration in the Kingdom from the 1930s onwards. The consequent large increase in job opportunities led to an influx of semi-skilled and skilled workers from poorer/low-income neighbouring states and Western countries, respectively. The five most frequent countries of origin for these economic migrants were India, Pakistan, Bangladesh, Egypt, and Syria. However, the number of non-national workers also increased in the private sector (De Bel-Air, 2014), affecting job opportunities for well-educated Saudi nationals. This was largely due to immigration from Egypt, Jordan, and Lebanon, all comparatively wealthy countries. The effect of this high unemployment rate for the young national population will be discussed later in this chapter.

¹⁴ Source: Central Department for Statistics and Information (CDSI)/ General Authority for Statistics (GASTAT).

¹⁵ (http://gulfmigration.org/media/pubs/exno/GLMM_EN_2018_05.pdf)

The Kingdom has seen significant advances in terms of life expectancy at birth and mortality rates matching those of other developed economies. Over recent decades, life expectancy has increased in Saudi Arabia (74.7 years for females and 72.5 years for males) due to a combination of socio-economic development and reductions in communicable diseases (Almalki et al., 2011). The Health Statistics Yearbook (2015)¹⁶ stated that life expectancy in the Kingdom is now four years higher than the global average. The proportion of the population under 15 years is 29.1%, which is higher than the global average of 27.0% and similar to the EMR average of 34.0%. Nonetheless, the effects of rapid urbanisation, economic growth and the growing expatriate population represent a significant demographic challenge.

3.3 Socio-economic drivers and behavioural patterns

The rapid urbanisation and socio-economic change described above has resulted in considerable variation between regions in terms of the quality of healthcare services and health. In this section, I discuss the association between the socio-economic status of the regions and health outcomes, with particular emphasis on diabetes mellitus. I will examine how urbanisation and rapid socio-economic change contributes to this through changes in lifestyle and behavioural patterns. Finally, I discuss how socio-economic characteristics differ between both national and non-national population and the effect of high non-national population on the employment status for Saudi nationals.

3.3.1 Administrative regions

Most previous literature examining the association between socio-economic factors and T2DM has focused at the level of individuals, with few studies exploring the impact of urbanisation on the prevalence of T2DM. Gassasse et al. (2017) investigated the association between urbanisation and T2DM by country income level, concluding that obesity and physical inactivity were the most significant risk factors in high-income countries. Urbanisation tends to result in an increase in the cost of healthy food (fruit and vegetables), but a reduction in the cost of processed and high fat food (Patil, 2014, Vlahov et al., 2007).

¹⁶ Saudi Ministry of Health Statistics Book. Available: <http://www.moh.gov.sa/en/Ministry/Statistics/book/Pages/default.aspx>. 2015.
Ref Type: Electronic Citation.

These shifts are due to industrialisation and a shift from an agricultural economy (Moore et al., 2003). The shift in diet, resulting in an increased consumption fat and carbohydrate, has led to an increase in the prevalence of obesity and the likelihood of become diabetic. This trend is most pronounced in urban regions, contributing to their high prevalence of T2DM (Moore et al., 2003). As previously discussed, two-thirds of Saudi population live in urban regions and are exposed to these changes in eating habits. The reduction in the cost of fast-food cost and concomitant increase in that of healthy food is associated with other changes in lifestyle which I will discuss in section 3.3.2. The following studies highlight the differences in obesity and consanguinity between regions.

El Mouzan et al. (2012) investigated the regional variation of bodyweight and obesity among children and adolescents in the Kingdom to inform design of a prevention program based on regional needs. They found that the Southwestern region has a lower prevalence of obesity and individuals who are overweight compared to Central and Northern regions and that regional variation is an important contextual factor in developing appropriate prevention programs. In a previous report, the same authors investigated the prevalence of consanguinity in rural and urban areas by region in Saudi Arabia (El-Mouzan et al., 2007). They found that overall, the prevalence of consanguinity in the Kingdom is high, but that it is subject to significant variation both between region and between rural and urban areas. The prevalence of consanguinity was consistently higher in rural areas. These two studies underlined the socio-economic diversity between and within Saudi regions and identified obesity and consanguinity as significant risk factors for T2DM in the country. Together they demonstrate that regional variation is an important contextual factor for health policy recommendations to government.

In a recent study, AlOmar et al. (2018) generated two socio-economic indices to measure inequality in health care access and health outcomes and support better health policy interventions. The Eastern region was the most affluent region as expected, since it houses the largest oil companies in the world such as Arabian-American Oil Company (ARAMCO). Outside the Eastern region, Riyadh and Jeddah were the most affluent cities. Key characteristics of the affluent regions are that they host the best universities, have access to specialised healthcare services, and better job opportunities. The study provides further evidence that the diversity of socio-economic status between regions is a significant factor influencing health outcomes - particularly for T2DM - and that this is in part related to differences in quality and efficiency of healthcare services between and within different regions.

This final point was examined in a study by Mousa and Aldehayyat (2018), who investigated the efficiency of healthcare services across the country. These investigators found significant variation in healthcare efficiency across regions and between the public and private sectors. Overall, less-urban regions showed higher healthcare efficiency compared to other developed regions such as the Eastern Province and Makkah in both sectors. However, the private sector was less efficient in some regions, other regions showed the opposite trend, while healthcare facilities for both public and private sectors were efficient in the rural regions such as Northern Borders and Al-Jouf. Overall, 69% of public healthcare facilities were classified as efficient as opposed to 54% for the private sector. Hence, the prevalence of T2DM could be higher in some regions due to variation in healthcare efficiencies.

The study concluded that health policymakers needed to take account of these regional variations in order to improve the efficiency and quality of healthcare services. This might extend to a requirement to reallocate resources between regions. Another report by the Cooperation Country Strategy for WHO and Saudi Arabia showed that, although there had been a huge effort to develop the efficiency of public and private healthcare sectors, non-national workers made up the majority of the healthcare workforce (WHO, 2013). To sum up, provision of health care facility varies between regions, however this does not always reflect differences in urbanisation level and socio-economic status. However, where a person lives in Saudi Arabia unquestionably has an impact on their health, affecting their chances of receiving an early diagnosis of disease and managing it effectively.

3.3.2 Lifestyle patterns

As mentioned in the previous section, **urbanisation** is an important determinant affecting exposure to risk factors for non-communicable diseases (NCDs); in many cases this operates through lifestyle changes such as smoking habits, unhealthy food consumption and lack of physical activity (Al-Rubeaan et al., 2014b). The GCC countries experienced a rapid socio-economic transition that resulted in lifestyle changes and concomitant changes in disease patterns (Al-Rasheedi, 2015). A striking example of this is the high prevalence of T2DM in the Arabian Gulf region (Bener et al., 2009). The GCC is now ranked among the countries with the highest prevalence of diabetes worldwide (Bennett, 2007), exceeding 23.0% among adults in Saudi Arabia, Bahrain and the UAE (Mehio Sibai et al., 2010). Much of this increase is can be explained by the shift to physical inactivity and unhealthy diets, both of which are well-established risk factors for T2DM (Davies et al., 2008, Al-Rasheedi, 2015).

In addition to socio-economic and political factors, these lifestyle and behavioural changes are influenced by the demographic transformation in the region. These changes have contributed to an increase in the intake of food high in fat and carbohydrate coupled with a reduction in physical activity (Aljoudi and Taha, 2009, Midhet et al., 2010, Al-Hazzaa et al., 2011, Awadalla et al., 2014, Al-Rubeaan et al., 2015, Al Bassam, 2012). Traditionally, Saudi food is already high in fat and carbohydrate with dishes typically containing meat, rice, dates, and bread. Saudi studies indicate that the majority of the population now consume unhealthy food, resulting in a decrease in the intake of vegetables and fruit. This dietary shift is most pronounced in adults (AL-Aboudi et al., 2015, AlQuaiz and Tayel, 2009, Al-Rubeaan et al., 2015, Aljoudi and Taha, 2009, Bani, 2015) and is likely to be contributing to the increase in obesity and T2DM.

A recent study by Moradi based on a national survey found that the proportion of Saudi adults consuming an unhealthy diet (high in processed foods, meat and soft drinks; low in fruit and vegetables) was particularly high in those aged 15 – 24 years (Moradi-Lakeh et al., 2017). With respect to age and gender, the study observed that older adults consume more fruit and vegetables and fewer processed foods¹⁷ compared to younger adults who are more influenced by the nutritional transition. Male more likely to consume meat and soft drinks compared to females, while a higher proportion of females consume vegetables (Moradi-Lakeh et al., 2017). Moreover, high socio-economic individuals are more likely to consume fruit, vegetables and pure juices and less soft drinks compared with low socio-economic individuals (Moradi-Lakeh et al., 2017). The findings of this study underline the poor nutritional habits among young Saudis from low socio-economic groups, a likely factor behind the high T2DM prevalence in SA.

The prevalence of obesity is high and ranged from 29.2% up to 45.3% among women in the GCC region (Alshaikh et al., 2017). The prevalence of **obesity** is higher among females than males; however, more men than women were classified as being overweight in the GCC countries. The highest prevalence of obesity was observed in Qatar, Kuwait and KSA (45.3%, 35.2% and 38.4%). Several factors could contribute to this high prevalence. High levels of carbohydrate and fat consumption (WHO, 2016b), multiple pregnancies (Daryani et al., 2005) combined with a sedentary lifestyle contribute to a significant increase in obesity among women in Arab countries and particularly in the GCC region. A large number of households have housemaids, further increasing physical inactivity, particularly in Saudi Arabia and Kuwait (Mabry et al., 2010). The increase in the number of housemaids reflect an increase in Saudi women with careers, large houses resulting from increased living standard, and socio-cultural reasons that reflect expectations of families with high socio-economic status.

¹⁷ Processed food: pastrami, salami, sausages, hot dogs, fast foods, canned/fried foods

Physical inactivity is a global public health issue, particularly in the Middle East and the GCC countries, where studies reveal that physical inactivity is prevalent among all ages and sexes. This is particularly acute among women due to gender inequalities, resulting in a further increase in lifestyle-related disease (Alshaikh et al., 2017). It is therefore important to introduce additional physical education programmes to promote active lifestyles throughout the GCC region, particularly among women. Although this will require political and socio-cultural change, a diabetes prevention programme for diabetes based on physical activity and diet could save 54% of the total direct costs and 31% of total medication costs while simultaneously improving quality of life (Mabry et al., 2016). Overall, most of the Saudi population are physically inactive, with one Saudi study in Asir (Southern region) reporting that 99.5% of Saudi females and 81% of Saudi males in Riyadh city could be classified as inactive. In general, two main types of factor contribute to the high prevalence of physical inactivity in the GCC countries.

The first are environmental factors such as a lack of public transportation and the warm climate. The urban environment in some countries actively discourages walking and cycling (Giles-Corti et al., 2016), encouraging people to use cars even for short journeys. Furthermore, increased urbanisation leads to a rapid increase in office-based administration and services that are often associated with physical inactivity (van der Berg et al., 2016). The second are behavioural factors resulting from social and cultural norms (AlQuaiz and Tayel, 2009). Many of the reasons for physical inactivity discussed above disproportionately affect womens' health due to Saudi cultural practices (Al-Alfi et al., 2004, Al-Nozha et al., 2004, Midhet et al., 2010, Aljoudi and Taha, 2009, Awadalla et al., 2014).

An example of this is the lack of female sport facilities. Recently, the Kingdom announced plans to introduce a Quality of Life Program by 2020¹⁸ as part of its overall Saudi Arabia 2030 Vision. This programme aims to promote healthier lifestyles in the Kingdom by increasing the quality and number of sport facilities, developing education programs and curricular activities for students, thereby reducing the prevalence of diabetes and obesity.

¹⁸ Arab Health Market Series (2018) *Diabesity, the Pulse*: ninth Edition. Available at: <https://www2.colliers.com/en-AE/Research/Dubai/Diabesity-Arab-Health-Market-Series-The-Pulse-9th-Edition> (Accessed 20 December 2018)

Another important lifestyle factor is the prevalence of smoking in the region. The prevalence is higher among males than females (Alshaikh et al., 2017), primarily because smoking by females is regarded as culturally unacceptable (Dar-Odeh and Abu-Hammad, 2011), however water pipe smoking has increased among women in the GCC region. The prevalence of smoking among women in Saudi Arabia is 9%, 7.9% in Kuwait, 20.7% in Bahrain, 0.8% in UAE and 0.5% in Oman (Aljefree and Ahmed, 2015), however these figures are likely to be underestimates. Water pipe smoking is highly prevalent in the region, as there is a (mistaken) perception that water pipe smoking is less damaging than smoking cigarettes (Dar-Odeh and Abu-Hammad, 2011).

Overall, considering socio-economic factors in relation to lifestyle factors in Saudi Arabia will enable policy makers to identify suitable interventions specific to high-risk groups. The high burden of T2DM has become a concern of national high priority on the Saudi public health agenda. As mentioned in the introduction, most diabetes epidemiology studies in Saudi Arabia have focused on lifestyle and health factors without considering the interaction between socio-economic factors and both lifestyle factors and diabetes mellitus. Furthermore, few studies have investigated the relationship between environmental factors and T2DM in Saudi Arabia. To summarise, community support is an essential requirement in modifying individuals' lifestyles to manage diabetes in Saudi Arabia (AlQuaiz and Tayel, 2009, Bani, 2015). An obvious example of this are the cultural barriers resulting in limited access for females to sport facilities. This combined with the unequal distribution of facilities across and within regions means that there are very limited opportunities for women to exercise, particularly in rural areas.

3.3.3 Socio-economic characteristics

In the previous section, I discussed the importance of socio-economic factors in relation to regional variation in lifestyle patterns based on regional and urbanisation levels. However, an individuals' socio-economic characteristics are also related directly to health. A good example of this is **education level**. This is recognised as one of the main socio-economic factors and there has been a decline in gender inequality in terms of educational attainment and literacy which has resulted from a significant demographic transition in the GCC countries (Shah, 2012). This transition, coupled with the Kingdom's focus on reducing the gender gap in education and literacy since the 1980s, has significantly improved literacy and educational levels in recent decades. However, although literacy has improved, the proportion of females who are illiterate (33%) is still high compared to males (13.6%) (Shah, 2012).

The outlook for Saudi women has been transformed by the rapid socio-economic changes driven by urbanisation. There has been a large increase in the number of public schools and universities across the regions, increasing access for both men and women (Hamdan, 2005). Based on the most recent education statistics and reports, there is now high educational attainment for both Saudi males and females. This is echoed by the WHO (2013)¹⁹, who observed that there has been a significant increase in literacy rates among adult Saudi females (15 years and above), from 79.7% in 2004 to 85.0% in 2010, when the overall literacy rate of the adult population was 88.0%. The proportion of people who had completed secondary education as of 2010 was 50.3% for females and 57.9% for males. The increase in the number of educational institutions in the Kingdom over this period is reflected in the significant increase in the number of university degrees completed by Saudi women.

Significantly, this increase in educational levels is mirrored by more diverse range of roles for women in the workplace after university. The Kingdom made huge investments in tertiary education (Masters and PhD degrees) between 2000 and 2010 by supporting nationals to study abroad in highly developed countries²⁰. Nonetheless, over the period of this investment **poverty** and **adult unemployment** for Saudi nationals increased significantly, reaching 12.2% (33.4% for women against 7.6% for men) overall and 28.4 % for young nationals aged 15 – 29 (60.3% for women against 17.5% for men) (De Bel-Air, 2014). The high proportion of non-national workers in the country also contributed to the higher rate of unemployment among national females compared to males (Shah, 2012). In 2008, however the overall unemployment rate among young Saudi nationals (age 15-24) was 28.2%, for females in this group it was 45.8% (WHO, 2013).

¹⁹ World Health WHO (2013) Country Cooperation Strategy for WHO and Saudi Arabia 2012–2016. *World Health WHO, Regional Office for the Eastern Mediterranean, Cairo*. Available at: [\[http://apps.who.int/iris/bitstream/10665/113227/1/CCS_Saudia_2013_EN_14914.pdf\]](http://apps.who.int/iris/bitstream/10665/113227/1/CCS_Saudia_2013_EN_14914.pdf) (accessed 3 March 2016)

²⁰ The King Abdullah Scholarship Programme launched in 2005, then restructured in 2015 under the title “Your Job First and Then Your Scholarship Program” (<https://www.moe.gov.sa/en/studyabroad/The%20CustodianOfTheTwoHolyMosquesOverseasScholarshipProgram/Pages/YourJobFirstAndThenYourScholarshipProgram.aspx>).

These nationalisation policies seek to reduce the high unemployment rate among the national population by increasing the **employment opportunities** for nationals and reducing dependency on foreign labour. Another key aim of the labour nationalisation policies is to make the private sector more attractive for Saudi nationals. One of three main **Saudi Vision 2030**²¹ pillars is to achieve a “thriving economy” by increasing job opportunities for Saudi Citizens, increasing women’s participation in the labour force and improving the robustness of the economy of the Kingdom by a shifting from a dependency on oil resources. Nonetheless, unemployment of nationals is still significant issue in Saudi Arabia.

To conclude, the characteristics of the substantial non-national population in Saudi Arabia suggest that inequality of healthcare services could be related to high inflows of non-national workers in the three main regions. As discussed in the beginning of this chapter, the Riyadh, Western and Eastern regions have the highest level of socio-economic development and urbanisation, implying that they host better healthcare services, have better access to healthy food and increased availability of public places for physical activity and sports facilities. As discussed, most non-national workers live and are employed in these regions. It is therefore possible that non-nationals could indirectly affect the availability of better diabetes screening and management in these regions compared to the North and South regions.

Some tables show the socio-economic and demographic characteristics of the administrative regions based on the Saudi Health Interview Survey (SHIS, 2013), drawing on the SMOH and National Diabetes for Control and Prevention Program database (see appendix A). With respect to household monthly income, the recent Household Income and Expenditure Survey (HIES, 2018)²² show that families who live in the Eastern region have the highest average monthly income per household followed by Jazan and Riyadh regions, while families who live in Najran region have the lowest. The results highlight that there is an obvious difference between regions based on monthly household income for the total population (national and non-national). Considering the average monthly household income for Saudi households only, the Eastern and Riyadh regions are the highest, while Tabuk region is the lowest. The following **Table 3.1** and **Figure 3.4** highlights the average household monthly income for Saudi and total population.

²¹ https://www.saudiembassy.net/sites/default/files/u66/Saudi_Vision2030_EN.pdf

²² Household Income and Expenditure Survey (HIES): <https://www.stats.gov.sa/en/37>

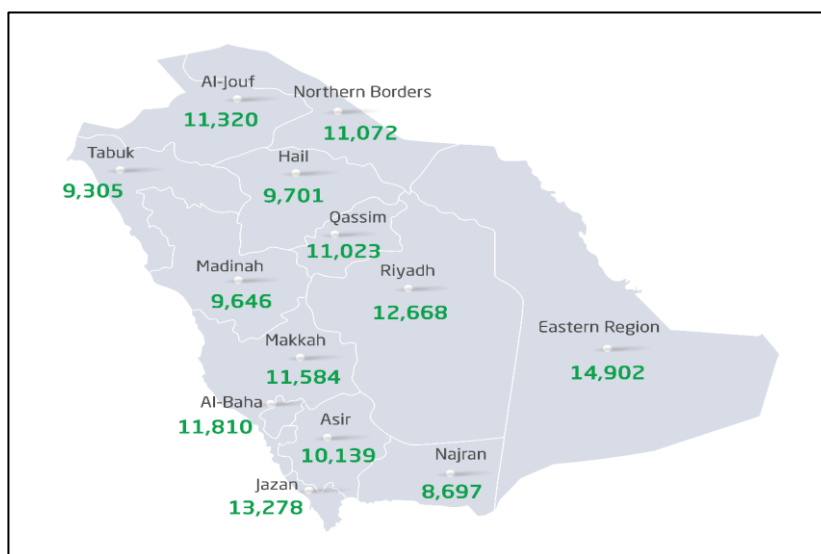


Figure 3.4 Average Monthly Household Income for the total Saudi population by administrative regions using Household Income and Expenditure Survey

Table 3.1 Average Household Monthly Income by Administrative Region and Nationality of Household Head²³

Administrative region	Saudi population (Saudi Riyals)
Eastern Region	17,872
Riyadh	16,011
Qassim	15,322
Jazan	15,199
Makkah	14,648
Al-Jouf	14,193
Al-Baha	13,728
Northern Borders	12,051
Madinah	12,016
Asir	11,817
Hail	11,571
Najran	11,388
Tabuk	11,024
Total	14,823

²³ 1000 Saudi Riyal (SR) $\approx$ £214 $\approx$ 267\$

3.4 Saudi health care services

In this section I discuss the structure of the Saudi health system structure, the development of health care services and its response to the challenge presented by T2DM. Due to rapid economic growth, the GCC governments have been in a position able to tackle the prevalence of communicable diseases and improve health care services, both contributing to an increase in life expectancy (WHO, 2006b). However, in recent years, there has been a parallel increase in the prevalence of NCDs such as diabetes mellitus, obesity, and hypertension in GCC countries, driven in large part by the results of economic growth. These diseases are complex and multifactorial, and mounting an effective health care response is further complicated by the huge geographical area of the Kingdom which effectively impedes the efficient distribution of health care services and the workforce.

3.4.1 Health system structure

The Saudi health system is divided into two main sectors: public and private (**Figure 3.5**). In the public sector, there are three main levels of healthcare provision: primary, secondary, and tertiary (MOH, 2009)²⁴. These are supported and funded by the Ministry of Health. Primary healthcare provides prevention and therapeutic services; while the secondary level is responsible for more intensive health care requiring hospitalisation or specialist clinics; the tertiary level deals with further referrals to specialist centres for more complex cases. The Primary healthcare level is the main provider of healthcare services for Saudi nationals (Albejaidi, 2010). These centres deal with a high proportion of non-communicable disease such as cases of hypertension, diabetes, and obesity; these are usually referred to secondary healthcare centres.

The Government is also responsible for the main referral hospitals such as The King Faisal Specialist Hospital and Research Centre, the Security Forces Medical Services, the Army Forces Medical Services, the National Guard Health Affairs, ARAMCO hospitals, The Royal Commission for Jubail and Yanbu health services, the Red Crescent Society and the Ministry of Education hospitals (School health units, and University hospitals) (Almalki et al., 2011). In contrast, the private sector represents only 20% of healthcare facilities (MOH, 2015). Secondary healthcare is provided at the district level and organised under regional supervision. The tertiary healthcare level is responsible for chronic and complex health issues requiring highly experienced medical staff and the most advanced equipment.

²⁴ Ministry of Health. (2009). *Health Statistical Yearbook*. Riyadh: Saudi Ministry of Health Available at: <https://www.moh.gov.sa> (Accessed 6 July 2018)

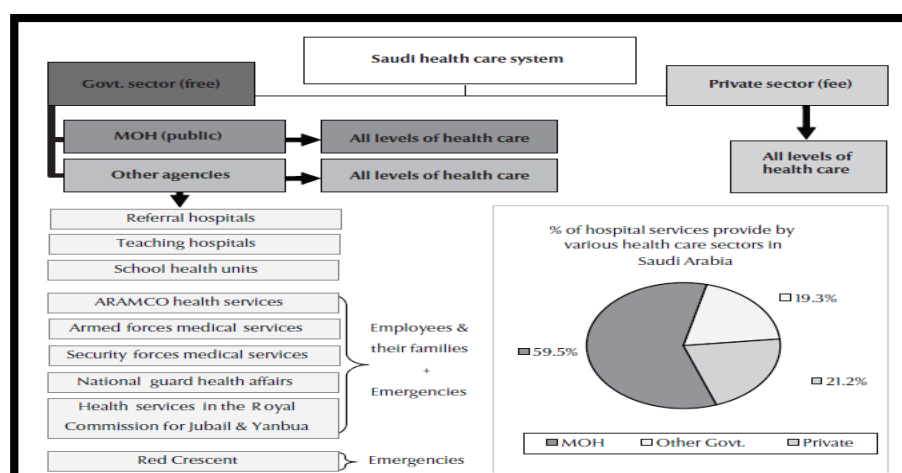


Figure 3.5 Current structure of the health care sectors in Saudi Arabia (Almalki et al., 2011)

In addition to these organisational issues, the increase in NCDs such as T2DM also present a major challenge to the Saudi MOH (Khoja, 2009). In addition to provision of the basic medical infrastructure at all three levels of care, the MOH is responsible for setting healthcare policies, implementing programs focused on early detection and treatment of disease (Almalki et al., 2011, Al-Sharif and Al-Khalidi, 2003) (MOH, 2013c; MOH, 2015)²⁵.

3.4.2 Development and challenges of healthcare services

The efforts of GCC governments to tackle the prevalence of communicable disease has resulted in an increase in life expectancy (WHO, 2006). These large-scale public health initiatives and shifts in demography have benefited from and contributed to rapid socio-economic changes in the region (Salam, 2019). The improvement in living conditions is leading to a change in the GCC population age-structure. Thus a reduction in fertility rate combined with increases in life expectancy is predicted to increase the proportion of the adults and overall ageing of the population in the region (Khraif et al., 2015). A Saudi study based on data from the Saudi National Survey for Elderly Health (NSEH) provided evidence that there is an association between income status and health status for those aged over 60, with high-income individuals more likely to have "very well" health status compared to low-income individuals (Khoja et al., 2018).

²⁵ Ministry of Health. (2013c). *National Executive Plan of Diabetes Control (2010-2020)*. Available at: <https://www.moh.gov.sahttp://www.pitt.edu/> (Accessed 20 August 2018)

GCC government health facilities at all levels are still not well prepared for the increases in the prevalence of NCD or the ageing population. These dual challenges require large-scale interventions to develop health care services appropriate to an ageing population. A particular focus should be on improving primary health care centre (PHCC) services and facilities with a focus on NCD prevention and early diagnosis in order to increase mobility and extend healthy life spans. The Saudi Ministry of Health (SMOH) has a national development plan to address public health concerns. This has resulted in an increase in nursing and physician staff and new hospitals, with concomitant reduction in infant and maternal mortality rates (WHO, 2013). One element of this national improvement plan was service delivery for primary health care patients. This service has recently been improved by the MOH in cooperation with national and international organisations (Walston et al., 2008). However, an analysis of primary care services in Saudi Arabia found that access for older adults remains poor. Health care services also face a major challenge from the increase in prevalence in NCDs with a lack of effective interventions (Mourshed et al., 2006). Although health care services have improved in the Kingdom during recent decades, these improvements have failed to reach all socio-economic groups (Almalki et al., 2011).

The government offers free health care at public hospitals and clinics for the Saudi national population. However, a number of citizens additionally receive health insurance from the employer that provides advantageous healthcare access. A recent Household Health Survey (HHS, 2017)²⁶ show that 37.1% of Saudi total population are covered by health insurance. The Eastern region has the highest healthcare insurance coverage rate for Saudi nationals (29%) while the lowest is observed in Al-Baha region (1.6%). Makkah region recorded the highest health insurance coverage rate for non-nationals (86.3%) with the lowest in Jazan region (26.3%). Although health insurance is now provided to Saudi employees and their families along with access to private health care centres (Mourshed et al., 2006), it could be argued that the GCC governments should provide health insurance for all nationals, regardless of their employment status. The diversity of healthcare insurance status between individuals with different working status or in different occupations inevitably leads to inequity in health status. This has led to a recognition by the GCC governments that major health system policy changes will be needed, particularly with respect to healthcare quality in public health care centres (Khoja et al., 2017).

²⁶ Family/ Household Health Survey (FHS, 2017): <https://www.stats.gov.sa/en/965>

Health insurance partnership were established in 1999 for expatriate workers in the Kingdom, however this only covers Saudis and foreigners who are working in the private sector (Almalki et al., 2011) and does not extend to unemployed individuals or public-sector workers. Health insurance is now provided for all nationals and non-national workers in the government sector (Alsharif, 2008), however non-Saudi and Saudi employers in the private sector have to pay for advanced health services. The MOH has made a concerted effort to improve the quality of health care in both public and private health care facilities. Public health care services are provided free of charge for all Saudi citizens and employees within the public sector (Almalki et al., 2011). However, health care costs in the private sector are very expensive and are not covered by health insurance and out-of-pocket costs for employees in the public sector (Walston et al., 2008).

Furthermore, most private clinics and hospitals are in urban areas within the main cities (Albejaidi, 2010). As a result, rural areas often have only limited access to the better health care services provided by private facilities. Moreover, foreigner workers are only allowed to use public health facilities for emergencies. The health insurance system therefore faces challenges in terms of health care access and its funding in the Kingdom²⁷. These considerations make it clear that there are issues with poor quality of public health care services, lack of treatment and prevention care services and inequities in the health insurance system. This inevitably contributes to the increase in the prevalence of non-communicable diseases such as T2DM, obesity and hypertension.

These issues are further compounded by variation in the geographic distribution of human resources and limited access patient information documents. Such problems are driven in part by the strain of increasing health care costs associated with the high prevalence of chronic disease and the introduction of new health and biomedical technologies. Whatever the cause, the majority of nationals are not completely satisfied with the quality of health care in government hospitals (Mourshed et al., 2006). Key concerns reported by patients were access to health care centres and long waiting times, a lack of public transportation and limited parking facilities (Alhamdan et al., 2015). One consequence of the fact that improvements to health care services have not provided the high quality of services appropriate to national income is that some patients still rely on traditional treatments (Almalki et al., 2011). These have significant influence in the Saudi Arabian context, as the use of traditional herbal remedies is common practice among the poorly educated and older generations, in both rural and urban areas in the Kingdom (Al Saeedi et al., 2003, Bakhotmah and Alzahrani, 2010, Sabra et al., 2010).

²⁷ <https://www.moh.gov.sa/endepts/Health-Insurance/Pages/Challenges.aspx>

3.5 Diabetes Management and Screening: policy, implementation and prevention programmes

3.5.1 Diabetes policies and plans

The Saudi Charitable Association of Diabetes (SCAD)²⁸ was launched in 2007 with Ministry of Social Affairs approval and aims to increase community awareness of diabetes and promote health education. The association includes charitable, scientific, social, and financial activities and provides medical support for diabetes. However, it is the community association considered responsible for diabetes care (SCAD, 2017), the role of the community in supporting individuals with T2DM and wider healthcare policy implementation has not been studied. With regard to national policy, the Saudi MOH has launched three national initiatives relating to the prevention and management of T2DM and other chronic diseases. The first of these was in 2001, when the Gulf Cooperation Council States prepared an action plan to address the high burden of diseases in general and T2DM in particular.

The Ministries of GCC countries²⁹ assembled a committee focused on developing and updating healthcare programmes that agreed a number of key resolutions (Khoja, 2009) (Khoja, 2010³⁰). The *Second national initiative* was a T2DM specific challenge developed as a Country Cooperation Strategy by the WHO and KSA working in partnership between 2006 and 2011 (WHO, 2013). One key recommendation of this program was that the Saudi Healthcare system needed to raise awareness of the need to adopt a healthy lifestyle, particularly among the young. Additional recommendations included the need for wider health education programmes and more research into the current situation in the country. An important outcome of this was the development of the third initiative - *The National Executive Plan of Diabetes Control 2010 – 2020*; this was the first major program focused exclusively on the control of diabetes, with a dual emphasis on prevention and minimising the development of comorbidities.

²⁸ Saudi's Charitable Association of Diabetes: <http://alsukkary.org.sa/index.aspx>

²⁹ <http://www.bsac.org.uk/antimicrobialstewardshipebook/Chapter%2015/National-AMR-Plan-2017.pdf>

³⁰ Khoja, T. (2010). *Strategic Approaches in Combating Diabetes Mellitus among GCC Countries*. Available at: <http://www.pitt.edu/> (Accessed 15 August 2018)

National Executive Plan of Diabetes Control (2010-2020)³¹:

Objectives:

Primary:

- To reduce incidence and prevalence of T2DM and its risk factors.

Secondary:

- To reduce T2DM complications and comorbidity through early detection of disease.

Tertiary:

- To develop and improve the quality of health services delivered to diabetic patients.
- To improve diagnosis, follow up, and assessment of diabetic patients through a Diabetics Registration Program.
- To improve the research tools available and both the number and quality of studies on diabetes.
- To enable diabetic patients and their families to contribute to the control diabetes and its complications.
- To support and encourage diabetes management within community.

Programmes:

- Anti-diabetes Education National Program.
- National Campaign for Diabetes Control.
- World Diabetes Day.

Although these objectives were consistent with global action plans for the control and prevention of non-communicable diseases in 2013 (WHO, 2013), the plan failed to provide any specific policies or strategies for prevention or enhanced monitoring of T2DM. In addition, there was no explicit statement regarding the need for cooperation between levels of healthcare and between the private and public sectors (MOH, 2013a)³². The WHO recommended that all governments and Ministries of Health should revisit how their policies addressed the existing “Health for All” strategy developed in 1998 to provide health policy guidelines at national and global levels (WHO, 1998). In particular, they noted that policy makers in Saudi Arabia should design effective healthcare policies that take into account broader characteristics such as environment, culture and healthcare services. An effective “Health for All” plan is in effect a nation’s public health strategy, and should have at its core three main principles: equity in terms of access to healthcare resources; a recognition of the individual’s needs and their right to participate in decisions relating to their health; and the importance of collaboration between sectors in the delivery of care.

³¹ **National Executive Plan of Diabetes Control:** [<https://www.moh.gov.sa/en/Ministry/MediaCenter/Publications/Pages/Publications-2013-11-12-001.aspx>]

³² Ministry of Health. (2013a). Guideline for Diabetes. Available at: <https://www.moh.gov.sahttp://www.pitt.edu/> (Accessed 15 August 2018)

Arguably, the current Saudi healthcare policy for control and management of T2DM does not fully address these three core tenets of the “Health for All” vision (Al-Sharif and Al-Khaldi, 2003, Midhet et al., 2010). For example, key contextual factors such as economic status and sociocultural considerations are effectively ignored in the development of Saudi health policies (Sutton, 1999, Walt and Gilson, 1994, Gilson et al., 2008). In the last two decades, Saudi Arabia has faced significant and rapid political and economic transition. In an attempt to update the third principle of the “Health for All” strategy, in 2012 the Saudi MOH launched the Health Programs and Chronic Diseases Control General Department. The Chronic Disease department is responsible for improving community and education projects on NCDs (MOH, 2013b³³ ; MOH, 2013c). This includes administration of the Diabetes Unit, covering 22 specialised centres and 270 diabetes units within hospitals throughout the regions (MOH, 2013c).

As discussed in the previous section, much of the literature assessing diabetes policies and plans has concluded that there is lack of an effective health programme for dealing with T2DM. In particular, deficiencies in the health information system are undermining attempts to improve the referral system (Al-Sharif and Al-Khaldi, 2003, Al-Alfi et al., 2004, Al-Ghamdi et al., 2007, Al-Alfi et al., 2007, Abdelwahid et al., 2010, Al Wadaani and Balaha, 2012). The referral system for T2DM is critically dependent on efficient transfer of diabetic patients’ data from primary healthcare centres to secondary and tertiary healthcare centres. This problem is exacerbated by the fact that referrals are usually between primary healthcare centres to specialised diabetes units that are often in the private sector. In addition to a poor referral system, there are very few dieticians and diabetes educators and primary healthcare clinics often suffer from a lack of equipment and access to the most appropriate drugs. Doctors in General Practice are focused on treatment and are unwilling or unable to provide advice on the control and management of diabetes, resulting in a high referral rate to diabetic units (secondary level) (Al-Alfi et al., 2004, Al-Sharif and Al-Khaldi, 2003).

³³ Ministry of Health. (2013b). *Media Report on the MOH Efforts of Educating on Diabetes*. Available at: <https://www.moh.gov.sa> (Accessed 20 August 2018)

In 2013, the MOH acknowledged the importance of improving referral system policies and issued guidelines for the management of diabetes (Al Shahrani and Baraja, 2014). An additional factor affecting the efficiency of diabetes management is that most nurses (70%) working in primary health care centres are expatriates from non-Arabic speaking countries. This inevitably compromises communication between patients and nurses. Health education and disease prevention efforts are further compromised by the fact that many expatriate nurses lack appropriate training for public health (Al Wadaani and Balaha, 2012, Midhet et al., 2010, Aldossary et al., 2008, Sabra et al., 2010).

In view of these considerations, it is not surprising that most studies observed that there is insufficient diabetes awareness in the population. This even extended to diabetics themselves, most of whom believe that the disease can be managed effectively through medication alone, with little awareness of the importance of diet and physical activity. The inevitable conclusion is that knowledge about diabetes in the Kingdom falls far short of WHO recommendations (Sabra et al., 2010, Al Saeedi et al., 2003, Sharaf, 2013, Aljoudi and Taha, 2009, Al Shahrani and Baraja, 2014, Al-Elq, 2009). In addition, the lack of optimal nutrition is another important issue of increasing the prevalence of NCD's including T2DM in the country.

A review of the quality of T2DM management in the GCC countries by Alhyas et al. (2012) highlighted the lack of standardisation in the outcome measures for T2DM, and how this together with the large amount of missing data makes meaningful comparison difficult. For example, assessment of glycaemic control requires a full documentation of fasting blood glucose or HbA_{1c} values over the last few months. Another key issue identified by these authors was the poor management of high-risk diabetes groups in the GCC region due to the diversity of health system type and lack of management guidance within primary healthcare centres.

3.5.2 Public health concerns of T2DM

One of main concerns relating to the prevalence of T2DM in the country is its huge economic burden (Alhowaish, 2013, Alharbi et al., 2014, Mokdad et al., 2014). Individuals with diagnosed diabetes incur health care costs some ten times higher than non-diabetic individuals (\$3,686 vs. \$380). In 2014, the MOH estimated that management of diabetes cost \$4.5 billion for Saudi citizens, projected to rise to \$7 billion. About 72% of the total budget spent on insulin was for treatment of T2DM at an annual cost of insulin per patient of \$375 (Alhowaish, 2013, Al-Rubeaan et al., 2013). These estimates do not consider indirect costs and or reflect loss in quality of life and it is clear that T2DM places huge pressure on both the healthcare system and society as a whole.

A report by AL-Aboudi et al. (2015) investigated the health-related quality of life of T2DM patients who attended diabetes units in Riyadh city hospital. The authors observed that Saudi T2DM patients had low health-related life quality, with females more affected than males. The Saudi government has now planned to increase total healthcare expenditure by \$9.4b; of this \$0.9b will be dedicated to management of diabetes. The use of evidence-based guidelines for the management and prevention of T2DM was recommended in the 2016 WHO Global Report on Diabetes (WHO, 2016a). It also highlighted the need to tackle the high prevalence of diabetes through programmes to encourage healthy diets and physical activity, particularly directed at individuals from lower socio-economic groups. Current Saudi policy is based on the national reference for diabetes mellitus guidelines and national policy³⁴ for primary health care centres (PHCC) issued in 2014 by the national diabetes prevention and control programme.

3.5.3 Screening and management guidelines for T2DM

The Saudi guidelines for the management and control of diabetes focus on three main elements: weight loss, diet, and physical activity. These factors are key to effective self-management of diabetes. As discussed previously, there has been significant effort and investment made in an attempt to reduce the barriers to physical activity by increasing the quality and quantity of public sports facilities. However, there is an urgent need to go further in modifying lifestyles. This should involve legislation to reduce the cost of healthy food for all socio-economic groups, reduce the amount of sugar in food and drinks and consumption of unhealthy food through the introduction of taxes on processed food high in sugar and fats.

Additionally, both individuals and communities should work together to prevent and control diabetes. As an example of this approach, the UAE Ministry of Health have put in place an education programme to promote self-management of diabetes alongside the introduction of new prevention strategies including:³⁵

- Enforcement of a 50% 'sugar tax' on soft drinks and a 100% tax on energy drinks and tobacco
- Research into reducing the sugar content of soft drinks
- A blanket ban on marketing unhealthy food to children

³⁴ Saudi National Reference for Diabetes Mellitus Guidelines in Primary Health care, KSA, MOH, 2014: This reference has requested from the national diabetes program for diabetes.

³⁵ Arab Health Market Series (2018) Diabetes, the Pulse: ninth Edition. Available at: <https://www2.colliers.com/en-AE/Research/Dubai/Diabetes-Arab-Health-Market-Series-The-Pulse-9th-Edition> (Accessed 20 December 2018)

These regulations are more effective than just concentrating on education and publicity for individuals and diabetic patients and could be readily adopted by the Kingdom. Such measures would reduce the risk of progression from pre-diabetes to diabetes. The Saudi Diabetes guidelines are similar to those of the American Diabetes Association (ADA) guidelines in the US³⁶ that encourage diabetic patients to follow healthy lifestyle choices, take their medication and undergo regular blood sugar testing. Similar guidelines are followed by other high-income countries:

- **United Kingdom:** The most recent updated guidelines for the National Institute for Health and Care Excellence (NICE)³⁷ covered: adults aged 18 years old and more, patient education, dietary advice, managing blood glucose level, long-term complications and cardiovascular risk.
- **Finland:** the Finnish National Institute for Diabetes Management considered: lifestyle changes increased physical activity, weight loss, educated high-risk groups for lifestyle modifications, and the need to “generate regional and local models and program for prevention diabetes” (Saaristo et al., 2007).
- **Sweden:** Swedish Diabetes Care³⁸ included the importance of health care follow-up between the patient and the care unit, follow-up with those who are working in these units and educated them with the diabetes guidelines, and the region (the quality of healthcare units within and between regions and their location).

In addition to the new regulations listed above, the UAE have a comprehensive diabetes strategy that includes education, long term follow-up of individuals, region specific models of care and a focus on comparing the quality of diabetes healthcare units between regions. These are all meaningful additional strategies that could be usefully adopted for the management of diabetes in the Saudi context. The latest update of these guidelines listed below including the screening and standard care for T2DM based on the Saudi National Reference for Diabetes Mellitus Guidelines in Primary Health Care (MOH, 2014)³⁹ (see Appendix B).

³⁶ American Diabetes Association (2019) Diabetes Care and Treatment Available at: <https://www.diabetes.org/diabetes/treatment-care>

³⁷ National Institute for Health and Care Excellence (2015) Type 2 diabetes in adults: management. Available at <https://www.nice.org.uk/guidance/ng28> (Accessed 10 August 2019)

³⁸ Swedish Diabetes Care: Available at <https://webbutik.skl.se/bilder/artiklar/pdf/5388.pdf> (Accessed 10 August 2019)

³⁹ Ministry of Health. (2014). *Saudi National Reference of Clinical Guidelines for Diabetes Care*. Available at: <https://www.moh.gov.sa> (Accessed 21 November 2016)

3.6 Conclusion

In this chapter I have described the Saudi Arabian context in terms of demographic changes in life expectancy and socio-economic factors such as urbanisation and economic growth. The increase of urbanisation has been associated with job opportunities, increases in educational attainment and improved health care services. Over the same period, due to the oil boom, the country has faced a rapid increase in the migrant expatriate population to 12.2 million in 2017. These socio-economic improvements were linked to an increase in educational attainment amongst Saudi nationals.

However, these transformations also had significant negative consequences, as they were associated with lifestyle changes that resulted in a nutritional transition from a healthy diet to one characterised by overconsumption of fats and carbohydrate. This led in turn to a rapid increase in the prevalence of obesity and diabetes over last thirty years. Other environmental and community factors have also contributed to changes in individual lifestyles, with dependence on cars for mobility and a lack of sports facilities and public places reducing physical activity levels dramatically.

I also reviewed the Saudi health care services in detail, and the pressure on the system caused by the increasing prevalence of T2DM. I considered public health concerns in general, particularly as they apply to the diagnosis and management of T2DM and concluded by examining the screening guidelines for diabetes and how they compared to those of other high-income countries. Although I have touched on the national intervention programmes for T2DM in this chapter, these have so far been limited in scope and have not yet shown any effect on the burden of the disease and its complications. In the next chapter, I describe the characteristics of the dataset and the analytical methods I used.

Chapter 4 Data and methods

4.1 Introduction

In this chapter, I present a description of the study population and data sources used to investigate the association between education level and outcomes related to T2DM in Saudi Arabia. **Section 4.2** describes the data sources, while the data analysis plan and structure, weighting methodology and main statistical modelling approach are described in **Section 4.3**. This will be supplemented with a more detailed review of the relevant statistical methods within each of the main analytical Chapters 5 – 8.

4.2 Data sources

4.2.1 Saudi Health Interview Survey (SHIS)

The Saudi Health Interview Survey (SHIS) is the main source of data used throughout this thesis. This survey is a multistage nationally representative sample of Saudi individuals aged 15 years and over. It covers all Saudi Arabian regions and was conducted in 2013. The main objective of the survey was to collect data relating to health, demographic factors, and socio-economic risk factors. The SMOH hoped that this would identify health inequalities in Saudi Arabia and how these related to the major risk factors for common chronic diseases, informing policy and leading to better health for the Saudi population.

4.2.1.1 Survey design and flow

The survey used a multistage stratified sample designed to ensure that the sample collected was appropriate for each stratum. The stratification was based on division into the 13 regions of the Kingdom: (Al Riyadh; Makkah Al Moukarrama; Eastern Region; Northern Borders; Madinah; Jazan; Aseer; Najran; Qaseem; Tabuk; Hail; Al-Jouf; Al-Baha). The Saudi Arabian Census Bureau has generated small clusters of households that represent the primary sampling 'units' for the SHIS. The number of households per unit depended on population size, density, and geographical location. Interviewers visited households in pairs - one male and one female - to enable participants to be interviewed by a member of the same sex.

The SHIS survey consisted of two phases. The first was a household questionnaire filled out by the interviewing teams. In the second, respondents from the first stage were invited to complete a health examination at a local clinic. The two phases were further divided into four modules (three modules in phase one, one module in phase two). In the **first module (household roster module)**, the interviewer randomly selected an adult aged 15 years or above to complete the subsequent modules. If the head of the household was away or unavailable, any other household member could provide the information. At the end of this module, the DatStat software randomly selected an adult to complete the following survey stages. The household questionnaire included questions on the following socio-demographic factors: Sex, age, marital status, education level, employment status, and monthly household income estimate (MHIE) status. Individuals who did not answer the household income question were classified either as unknown or declined to respond.

The **second module (questionnaire and physical measurements module)** was divided into two parts: an adult questionnaire and collection of physical measurements. After obtaining informed consent from the randomly selected adult, they were asked a series of socio-demographic and health related questions. After completing the questionnaire, physical measurements such as weight and height were recorded (SHIS variables listed in the Appendix A). At the end, all selected participants were invited to visit their local health clinic for a series of further biomedical tests. The **third module** involved the final disposition code to compute the response rate when the survey was completed. In phase two, the **final module (biomedical measurements)**, was based on examination in the local clinic. All blood samples were shipped to King Fahd Medical City (KFMC) lab in Riyadh, where they were tested for fasting lipid profile, HbA_{1c} and vitamin D. Use of a single central laboratory ensured that all measurements were standardised. All values were entered and then transferred to the MOH. Questions relating to medication and insulin usage were also asked during the phase two clinic visit.

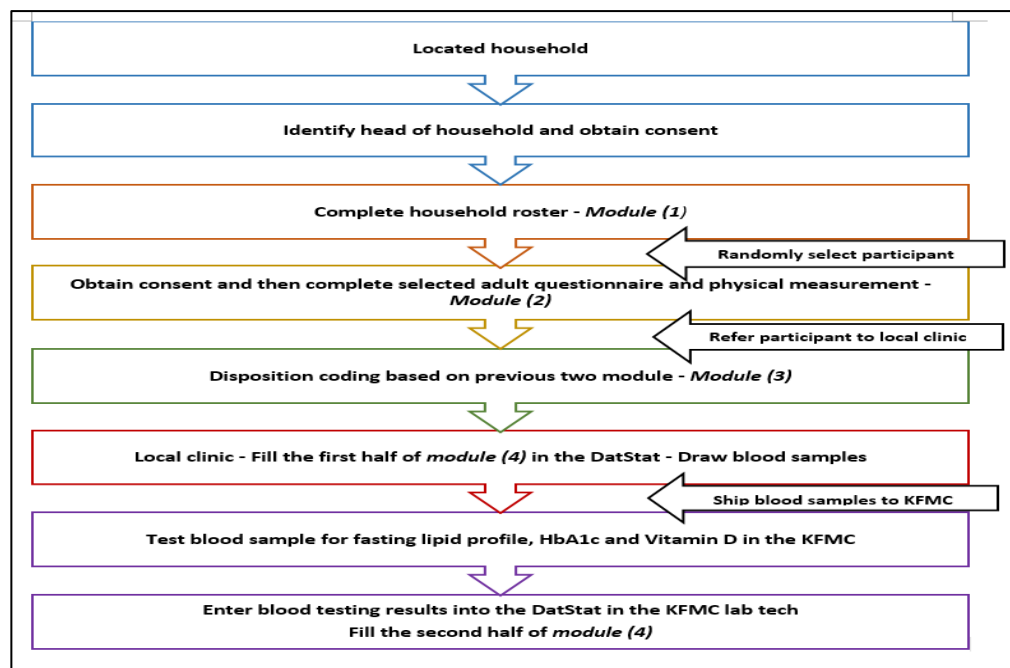


Figure 4.1 Saudi Health Interview Survey flow

4.2.1.2 Weighting methodology

As discussed previously, the SHIS data was provided with two sampling weights for data analysis. The **household questionnaire weight** (first sampling weight) was an individual sampling weight for all participants who completed the household questionnaire. This weight considered the selection probability of an eligible participant within a household, the selection probability of the household within a stratum and the post-stratification differences in the age and sex distribution between the SHIS sample and the Saudi population⁴⁰. This was calculated by the following process. Firstly, an individual's probability of selection was calculated by taking the total number of sampled individuals in their strata divided by the total number of households in that strata and the number of individuals eligible for selection in their household.

individual probability of selection

$$= \frac{\text{total number of sampled individuals in the strata}}{\text{total number of households in the strata} + \text{the number of individuals eligible for selection in the household}}$$

Secondly, an individual's selection *weight* was taken as the inverse of the probability of selection.

$$w_{person} = \frac{1}{p_{selection}}$$

⁴⁰ <http://www.healthdata.org/ksa/projects/saudi-health-interview-survey>

Finally, the household questionnaire weight was calculated as follows:

$$w_{first} = w_{person} \times \frac{\text{the percentage of individuals in that selected individual's age and sex group in Saudi population}}{\text{the percentage of individuals in that selected individual's age and sex group in the sample}}$$

The lab weight (second sampling weight) was for participants who gave a blood sample that was analysed at a King Fahd Medical City (KFMC) lab and used for subsequent data analysis. This weight took into account the household questionnaire weight, the probability of visiting the clinic, the differences between participants who went to the clinic or not based on socio-demographic, health and behavioural factors; and the post-stratification differences in age and sex distribution between participants who visited the clinic and the Saudi population. This weight is defined as the ratio of the total number of sampled individuals in that individual's strata to the number of individuals in that individual's strata from whom laboratory data were collected; multiplied by the individual's selection weight computed as described above.

$$\text{Lab weight} = \frac{\text{total number of sampled individuals in the strata}}{\text{the number of individuals in that individual's strata from lab data}} \times w_{person}$$

The lab weight was calculated by the following formula:

$$w_{second} = \text{Lab weight} \times \frac{\%pop(\text{household sample variables})}{\%pop(\text{module4 variables})}$$

The following flow diagram explains the additional weight to the SHIS weights:

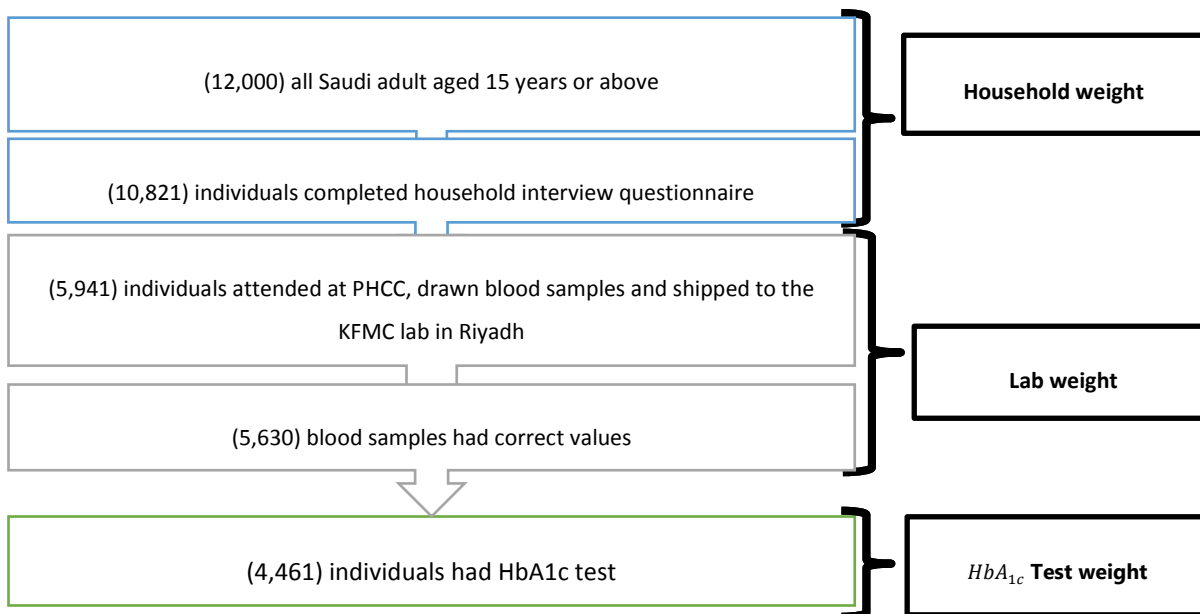


Figure 4.2 Flow diagram for SHIS stages and sampling weights

4.2.1.3 Ethical considerations

Each participant was informed of the purpose of the study, what participation would involve and how long it would take to complete. Consent could be written or verbal, and participants had the right to withdraw from the study at any time. Participant information and consent sheets were translated into Arabic. All interviews and physical measurements were taken one-to-one by someone of the same gender. Each interview was carried out by two interviewers, one who asked the questions and the other who took the physical measurements. No other person was present at the interview other than the participant and each participant was assigned a unique ID number. I can confirm that I have had permission to analyse the SHIS data from the Director of the National Diabetes Prevention and Control Program in Saudi Arabia. Appendix A gives a fuller description of the SHIS protocol and questions.

4.2.2 Saudi Arabia censuses and population estimates

I obtained contextual population and demographic data for Saudi Arabia from The General Authority for Statistics (GASTAT) in Saudi Arabia and the United Nations (UN). The Saudi population structure extracted from the GASTAT website is the main source of official population statistics in Saudi Arabia⁴¹. This comprises six Saudi censuses that are nationally representative for both Saudi and non-Saudi ex-patriates by age and sex (for 1992, 2004, 2010, 2016, 2017, and 2018). I also extracted Saudi total population life tables from the WHO website (2000 – 2016)⁴² to make projections for the Saudi population. Four early censuses were conducted to provide Saudi policy and decision makers with official data and statistics to support national development (1974, 1992, 2004, and 2010). The population structure was banded by gender into 5-year age groups (0-4, 5-9, 10-14 up to 80+) in each census⁴³. By combining historical population estimates covering the period 1992–2018 with population projections up to 2030, I could generate a complete sequence of national population estimates in 5-year steps constructed for 5-year age/sex bands.

⁴¹ **Population structure:** Central Department of Statistics and Information, Kingdom of Saudi Arabia (Official website). Available from: (<http://ghdx.healthdata.org/WHOs/central-department-statistics-and-information-saudi-arabia>).

⁴² **WHO Lifetables:** <http://apps.who.int/gho/data/?theme=main&vid=61440>

⁴³ **Saudi census:** <https://saudicensus.sa/en>

The GASTAT Saudi population estimates⁴⁴ were developed in stages. The numbers for 2017-2018 were generated based on last census conducted by the Saudi authorities in 2016. This provided interim data to bridge between the last full census in 2010 and the next census in 2020. This interim census provided important data and indicators with which to monitor changes in demographic characteristics. It used estimated data based on previous censuses coupled with recent administrative records covering births, deaths, and migration. These estimates started in the middle of the census year, termed the base year.

This census together with other sectorial projections, population and housing census results in the GAS allowed baseline estimates of the total population size by age-sex distribution. The numbers for the following years were projected based on assumptions relating to fertility, migration, and mortality rates. The GASTAT used a synthetic method to estimate population and provide a long-term projection of changes in population growth components (births, deaths, and immigration). This method was also used to generate separate Saudi and non-Saudi population projections in 2016 census by considering all available data on the non-Saudi population.

After finalising the projections, both data and indicators were reviewed in two stages. In the first stage, the projected population results were evaluated against other data sources, such as the UN and other relevant population projections sources. In the second stage, the projections were further evaluated by comparison against specialised demographic studies such as published articles to ensure the numbers were consistent. The following tables present the six Saudi censuses extracted from the General Authority of Statistics (see **Tables 4.1** and **4.2**). All these censuses can be used for projecting population growth and distribution for social and economic studies. 'Virtual' censuses for 2016 – 2018 censuses can be estimated based on previous censuses from GASTAT.

⁴⁴ Population Projections Methodology <https://www.stats.gov.sa/en/43>

4.2.3 United Nations population projections

Additional to previous censuses, I obtained additional data for 2015 and 2020 from the United Nations 'World Population Prospects'. These datasets are provided as Excel spreadsheets using different indicators and estimation methods through to 2100⁴⁵. The United Nations data on the Saudi population is based on probabilistic projections. This method reflects the historical variability of projections by considering previous experience for each country. The current study extracted the Saudi population structure to predict T2DM prevalence trends based on available datasets. The Saudi Ministry of Health (SMOH) relies on health statistics and estimates from the United Nations website as reliable data source (Al-Quwaidhi et al., 2014b). United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision⁴⁶.

⁴⁵ **Projections in population structure:** The United Nations; Department of Economic and Social Affairs. Population Division, Population Estimates and Projections Section. Detailed indicators. Available from: (<https://population.un.org/wpp/Download/Standard/Interpolated/>), (<https://population.un.org/wpp/>), (<https://population.un.org/wpp/Download/Probabilistic/Population/>), (<https://population.un.org/wpp/DataSources/>); **Demographic profiles:** (<https://population.un.org/wpp/Graphs/DemographicProfiles/>); Projections data sources - standard: (<https://population.un.org/wpp/Download/Standard/Population/>).

⁴⁶ United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects: The 2017 Revision, Online Demographic Profiles. Available from <https://population.un.org/wpp/Graphs/DemographicProfiles/>, Accessed on [March 2019].

Table 4.1 The number of Saudi national population 1992 – 2018⁴⁷

Sex	MALE						FEMALE					
Age	1992	2004	2010	2016	2017	2018	1992	2004	2010	2016	2017	2018
0 - 4	1,096,513	1,066,010	830,349	1,088,308	1,105,414	1,123,261	1,059,504	1,054,525	760,805	1,048,877	1,065,750	1,084,529
5 - 9	1,061,429	1,127,253	969,754	1,063,173	1,079,790	1,097,810	1,046,203	1,112,582	913,064	1,025,746	1,041,985	1,060,664
10 - 14	913,956	1,081,884	1,070,045	948,277	963,593	979,511	883,052	1,155,728	1,006,100	921,048	935,929	952,700
15 - 19	660,372	948,707	1,149,837	896,157	910,426	926,156	674,584	938,982	1,044,838	864,870	878,743	894,846
20 - 24	525,534	760,146	1,065,450	1,037,332	1,053,973	1,072,129	516,615	786,510	1,044,838	949,321	964,084	980,185
25 - 29	415,443	725,413	943,734	963,289	978,880	996,017	455,358	701,326	1,024,784	941,694	957,520	975,252
30 - 34	326,933	569,152	822,276	867,516	881,634	896,790	320,143	575,126	827,947	851,823	866,098	881,961
35 - 39	264,540	492,543	678,777	760,653	773,290	786,779	283,492	498,170	678,773	741,778	754,229	768,617
40 - 44	180,207	411,890	500,125	643,959	654,497	665,841	186,672	372,981	515,210	619,628	629,836	641,244
45 - 49	153,968	313,340	441,748	541,114	549,902	559,539	161,226	277,511	447,945	511,696	520,252	529,550
50 - 54	143,948	222,166	336,626	431,579	438,652	446,271	147,998	199,415	321,506	407,708	414,429	421,928
55 - 59	117,397	146,079	228,771	336,326	341,824	348,081	92,678	153,194	203,626	308,905	314,017	319,715
60 - 64	127,584	126,594	188,776	243,696	247,647	252,157	92,843	134,390	152,209	227,960	231,687	235,932
65 +	227,969	296,193	348,989	409,985	417,392	425,553	173,892	279,530	247,928	429,164	436,889	445,609
Total	6,215,793	8,287,370	9,575,257	10,231,364	10,396,914	10,575,895	6,094,260	8,239,970	9,189,573	9,850,218	10,011,448	10,192,732

⁴⁷ General Authority of Statistics, the general population and housing census for 1992, 2004, 2010 accessed on March 2019 [<https://www.stats.gov.sa/en/13>], population estimates 2016 - 2018 accessed on March 2019 [<https://www.stats.gov.sa/en/43>]

Table 4.2 Total number of Saudi population (National and Non-national resident)⁴⁸

	MALE						FEMALE					
Age	1992	2004	2010	2016	2017	2018	1992	2004	2010	2016	2017	2018
0 - 4	1,299,829	1,289,391	1,120,415	1,364,535	1,391,364	1,421,387	1,254,396	1,270,272	1,017,471	1,310,397	1,338,056	1,367,544
5 - 9	1,243,974	1,334,260	1,290,134	1,412,610	1,441,596	1,474,952	1,225,681	1,310,483	1,170,116	1,358,422	1,388,283	1,420,685
10 - 14	1,043,533	1,258,452	1,366,851	1,236,014	1,261,617	1,290,059	1,008,717	1,323,387	1,223,326	1,192,305	1,218,464	1,246,253
15 - 19	740,526	1,081,182	1,400,540	1,132,009	1,154,648	1,180,718	755,279	1,099,005	1,229,522	1,084,041	1,107,108	1,132,037
20 - 24	784,713	1,058,397	1,514,914	1,315,835	1,342,577	1,372,725	618,359	946,430	1,251,078	1,155,922	1,179,394	1,203,773
25 - 29	1,028,036	1,352,119	1,780,159	1,672,065	1,712,597	1,761,010	611,759	941,482	1,338,173	1,360,343	1,393,713	1,428,320
30 - 34	960,572	1,278,882	1,813,164	1,795,177	1,842,356	1,898,027	546,507	846,219	1,157,665	1,268,056	1,299,730	1,332,414
35 - 39	764,240	1,167,277	1,548,167	2,123,161	2,184,487	2,257,350	426,764	688,273	956,393	1,276,827	1,311,434	1,347,654
40 - 44	500,822	923,873	1,154,033	1,930,607	1,986,896	2,054,536	259,012	494,888	696,611	1,107,577	1,137,967	1,169,310
45 - 49	325,684	651,488	889,798	1,487,446	1,529,717	1,580,928	197,845	350,286	562,841	773,673	793,049	813,067
50 - 54	236,052	415,299	622,604	1,075,976	1,105,834	1,141,779	171,009	242,931	396,836	506,199	517,043	528,518
55 - 59	157,287	230,959	371,607	722,149	741,308	764,508	102,991	174,838	242,618	373,285	381,069	389,390
60 - 64	148,295	166,149	257,485	450,237	461,487	475,084	102,991	149,363	175,904	274,019	279,657	285,780
65 +	246,410	331,689	401,600	541,898	553,989	567,893	187,765	300,988	277,131	486,795	496,896	507,959
Total	9,479,973	12,539,417	15,531,471	18,259,719	18,710,473	19,240,956	7,469,075	10,138,845	11,695,685	13,527,861	13,841,863	14,172,704

⁴⁸ General Authority of Statistics, the general population and housing census for 1992, 2004, 2010 accessed on March 2019 [<https://www.stats.gov.sa/en/13>], population estimates 2016 - 2018 accessed on March 2019 [<https://www.stats.gov.sa/en/43>]

4.3 Statistical methods

4.3.1 Data analysis plan and structure

In my analysis, I have used three main diabetes outcomes, each examining a different aspect of affected status, namely: **1)** Prevalence of diabetes; **2)** Diagnosis of diabetes and **3)** Control of diabetes. I have used each of these to investigate the role of education in understanding the development of diabetes; improving detection of the disease once developed; and the control of the disease once diagnosed. This will contribute to the ongoing consideration of policies and interventions that target individuals who are more likely to develop the disease, remain undiagnosed, or fail to control the disease when they are aware of it. Data analysis was performed using STATA v.14 software and Microsoft Excel (2013). Most of study variables were included as they were needed to address the research aims and research questions, particular for the main research outcomes.

The data was first checked for inconsistencies - for example in nomenclature - and corrected where appropriate. Missing data patterns were identified, and their impact considered before analysing the dataset (see Chapter 5 for details). The choice of appropriate statistical method depended on the nature of the study outcomes and the specific research question addressed. Descriptive analysis was performed using Chi-square goodness-of-fit test to assess the bivariate relationships between the categorical outcome variable and exposure variables. The results are reported by appropriate cross-tabulation presented as proportions and frequencies. Binary logistic regression modelling was used throughout all the analytical chapters to test for association between the three main outcomes (prevalence, diagnosed and controlled diabetes) and study exposure variables. Additional methods are discussed within each of the four main analytical chapters as appropriate.

In **Chapter 5**, I investigated the association between education level and prevalence of T2DM and then predicted the future trends of diabetes prevalence. The outcome used here considers whether an individual has diabetes or not, irrespective of whether they have been formally diagnosed. This will be variously termed as ‘all diabetes’ or ‘prevalence’⁴⁹, and indicates if someone has either previously been diagnosed with diabetes or has been identified retrospectively through the HbA_{1c} test. This outcome therefore differentiates between non-diabetics and diabetics. The projection methods used are discussed in more detail in the chapter. **Chapters six and seven** focus on the diagnosed and undiagnosed population. In **Chapter 6**, I first investigate the association between education level and diabetes based only on the population with known diabetes. This outcome is variously termed ‘diagnosed diabetes’ or ‘previously known diabetes’, indicating that individuals have received a formal diagnosis of disease.

The majority of previous studies on socio-economic factors in relation to diabetes are based on such definitions, however with some variation in precise diagnostic thresholds (Bener et al., 2009, Corsi and Subramanian, 2012, Su et al., 2016). Based on the results of Chapter 6, in **Chapter 7** I focus on developing a prediction model for ‘undiagnosed diabetes’ or ‘previously unknown diabetes’, i.e. those individuals who did not self-report as having diabetes, but whose blood tests indicated that they are likely to have the disease. Additional methods used in this analysis are discussed, as is a comparison between the findings for diagnosed and undiagnosed disease in the two chapters.

In **Chapter 8**, I concentrate on the association between education level and controlled diabetes. This outcome is defined as those individuals who are aware of their diabetic status and yet had a normal HbA_{1c} test, and few studies to date have investigated its association with socio-economic factors (Khattab et al., 2010; Stringhini et al., 2012). In this chapter I investigate the association between the controlled diabetes outcome and interventions based on diet and physical activity as two of main approaches to management of diabetes. In this chapter, as in the previous analytical chapters, I include administrative region in the analysis as a potentially important external socio-economic factor. The following diagram **Figure 4.3** shows the survey flow and how it relates to the main outcomes of this thesis. Further discussion on each outcome and the most significant explanatory variables, chosen based on the previous literature and study conceptual framework, are included in the following chapters.

⁴⁹ Prevalence or all diabetes included diagnosed and undiagnosed diabetes cases

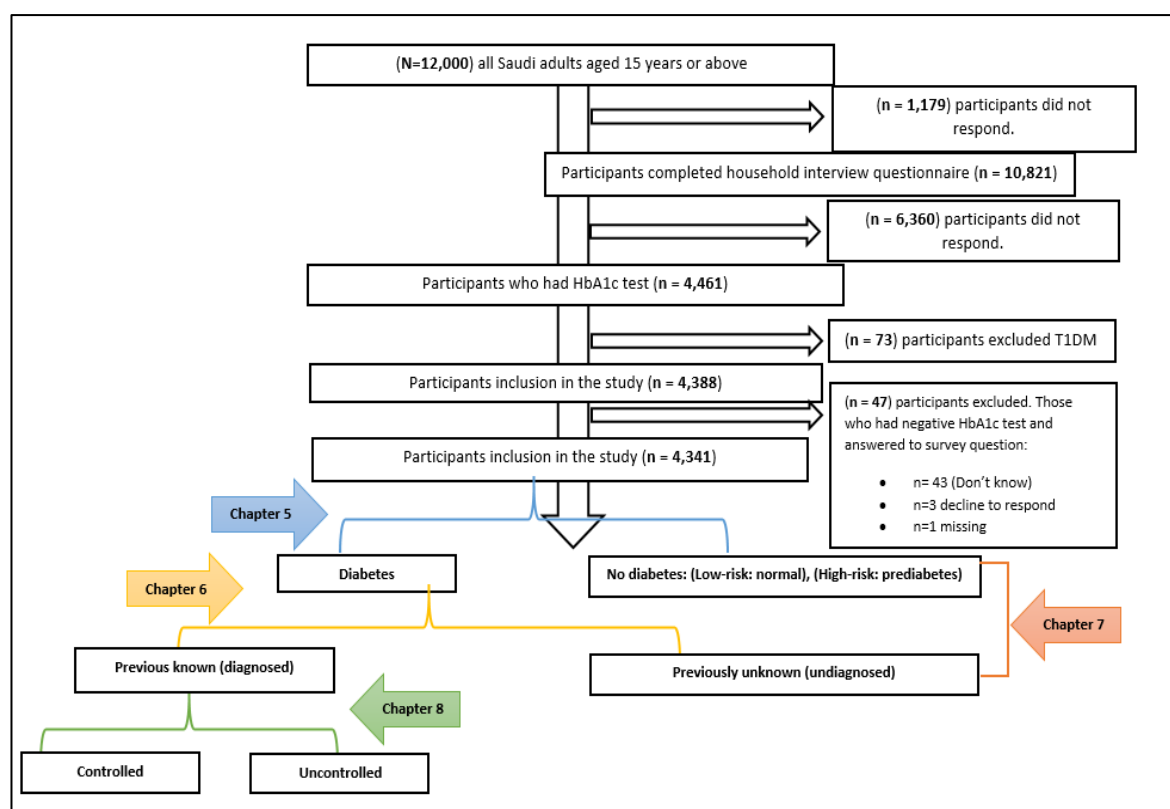


Figure 4.3 Survey flow and diabetes outcomes in this thesis⁵⁰

4.3.2 Definition of thesis outcomes

Table 4.3 presents the classification of diabetes outcomes based on the HbA_{1c} test and the self-reported question about diabetes asked within the individual questionnaire. The HbA_{1c} test classified individuals into two groups: positive if $HbA_{1c} \geq 6.5\%$ and negative if $HbA_{1c} < 6.5\%$. The relevant survey question was ‘Has a doctor or other health professional ever told you that you had diabetes mellitus?’. The possible answers to this question were yes, no, yes - but during pregnancy, do not know or decline to respond. As shown in the table, individuals were categorised based on the three main outcomes used in this study: all diabetes, previously known diabetes, and controlled diabetes. The first “Survey Question” column presents the answers to the self-reported survey question.

⁵⁰ **Blue colour:** represents Chapter 5 outcome; **Yellow colour:** represents Chapter 6 outcome; **Orange colour:** is follow up from previous chapter and represents Chapter 7 outcome; **Green colour:** represents Chapter 8 outcome.

These answers are further classified based on the HbA_{1c} test with the interaction between HbA_{1c} test and survey question presented in the total column. There are three main outcomes considered in this thesis as mentioned in previous section. The **first outcome “all diabetes”/ prevalence** includes all individuals who self-reported as well as those who had a positive $HbA_{1c} \geq 6.5\%$ test even if not-self reported. Thus, the non-diabetic group only includes those who answered: no; or yes, but during pregnancy to the survey question **and** had a negative test. Individuals who answered don't know or where the response was missing and had a negative HbA_{1c} test were excluded as I do not whether they were unaffected or had controlled diabetes.

The **second outcome, “previously known diabetes”/diagnosis** includes individuals who answered yes to the survey-question irrespective of their test status but excludes individuals who had a positive HbA_{1c} test but had answered that they did not have diabetes in the survey question. For inclusion in the **third outcome group “controlled diabetes”** individuals must have self-reported diabetic status but have had a negative HbA_{1c} test ($HbA_{1c} < 6.5\%$). The corresponding uncontrolled diabetes group will include only individuals who previously knew they had diabetes, but still had a positive HbA_{1c} test. As mentioned in the introduction section, the second and third outcomes will be investigated in the following chapters (Chapter 6 - 8).

Table 4.3 summarizes the number of individuals that were excluded from the study. The sample size for the study was 4,461 individuals who received a HbA_{1c} test and were aged 15 years and above. A total of 6,360 individuals who did not have a recorded HbA_{1c} test were excluded from the analysis. Some additional participants were excluded from the study if their response to the survey question was ‘Don't know’, ‘Decline to respond’ or missing and their HbA_{1c} test was negative. The study used the HbA_{1c} test to determine diabetes status following the criteria used in the National Health and Nutrition Examination Survey (NHANES, 2009), with individuals who had a HbA_{1c} test result of less than 6.5% classified as non-diabetic.

In the household questionnaire, individuals with known diabetes were also asked “What type of diabetes do you have?” 154 (12.3%) individuals answered T1DM; 808 (64.6%) indicated T2DM; and 288 individuals answered that they did not know or declined to respond. As the main focus of this study is T2DM, individuals with T1DM were excluded from the study. In total 40.6% of the original sample were included, giving a final sample size for this study of 4,388 individuals. An important aspect to note is that the Tabuk region was excluded from the study due to high rates of missing data for the HbA_{1c} test (only two successful HbA_{1c} test were reported), precluding any useful estimate of diabetes in the region.

Table 4.3 Classification of diabetes outcome based on HbA1c test and survey question

Diabetes identification			Outcome measures			
<i>HbA_{1c}</i> test		Survey-question	Total	All diabetes/ prevalence	Previously known/ diagnosis diabetes	Controlled diabetes
Sample sizes			4,388	4,341	937	538
1	Positive (+ve)	Yes	301	Diabetes (919)	Yes (538)	No (301)
2	Negative (-ve)	Yes	237			Yes (237)
3	Positive (+ve)	No	381		No (381)	excluded
4	Negative (-ve)	No	3,331	No diabetes	excluded	excluded
5	Positive (+ve)	Yes, but during pregnancy	9	Diabetes (9)	Included to No (9)	excluded
	Negative (-ve)		73	No diabetes	excluded	excluded
6	Positive (+ve)	Don't know	8	Diabetes (8)	Included to No (8)	excluded
	Negative (-ve)	Don't know	43	excluded	excluded	excluded
7	Positive (+ve)	Missing	1	Diabetes (1)	Included to No (1)	excluded
	Negative (-ve)	Missing	1	excluded	excluded	excluded
8	Positive (+ve)	Decline to respond	0			
	Negative (-ve)	Decline to respond	3	excluded	excluded	excluded

Positive (+ve): $HbA_{1c} \geq 6.5\%$; **Negative (-ve):** $HbA_{1c} < 6.5\%$; **Survey-question:** Has a doctor or other health professional ever told you that you had diabetes mellitus?

4.3.3 Inverse probability weight (IPW)

On initial analysis of the data, I noted that there was missing data for the HbA_{1c} test that was not accounted for in the lab weight. Although all blood samples taken as part of the SHIS were shipped and tested at the King Fahad Medical City (KFMC) lab, in a proportion of cases an HbA_{1c} test was not reported. As a result, I calculated an **additional weight** to allow for participants with no HbA_{1c} result despite attending the clinic and providing a blood sample. The aim of developing this new weight is:

- To investigate if the non-measurement of HbA_{1c} test in the SHIS affected the estimates of diabetes prevalence.
- To derive weights which could be used in the analysis of the association between socio-economic factors and diabetes mellitus in Chapter 5.

The new weight would also be expected to improve the accuracy of the estimate of diabetes prevalence used to address first research question in **Chapter 5**. The new weight was calculated using the inverse probability weight method to avoid selection bias in the analysis. This method corrects unequal selection probabilities in the sample by combining the sample selection weight with the missing data weight (Little and Rubin, 2014). The new weight is applied to each individual, taking into consideration both non-observed data and sample selection. In order to obtain the new weight a binary logistic regression model was estimated in which Y is a binary outcome with observed cases coded as 1, and those with missing data coded as 0:

$$y = \begin{cases} 1, & \text{if } y = y_o \\ 0, & \text{if } y = y_m \end{cases}$$

Where: y is the outcome variable; y_o is observed data and y_m is missing data. Thus 1 refers to individuals who had HbA_{1c} test ($N = 4,461$), and 0 refers to those who did not have a HbA_{1c} test result ($N = 1,169$). I fitted the logistic regression model based on the result of using backward selection of variables to avoid bias in the final estimates. From the model, the probability of an individual having an observed HbA_{1c} can be derived.

The probability sampling is unit selected from target population with specific probability π_i and then the weight is representing π_i^{-1} . Hence, for example, if we selected a unit in stratum j that represents N_j/n_j in a stratified random sample, it can be estimated that the population total by the weight sum is as follows:

$t_{HT} = \sum_{i=1}^n y_i \pi_i^{-1}$, which called the Horvitz-Thompson estimator, and the stratified mean can be written as follows is approximate unbiased.

$$\bar{y}_{st} = \bar{y}_w = \frac{1}{n} \sum_{i=1}^n y_i w_i, \text{ where the sampling weight is } w_i = n\pi_i^{-1} / \sum_{k=1}^n \pi_k^{-1}$$

We extend this approach to obtain non-response probability:

$$pr(\text{selection and response}) = pr(\text{selection}) \times pr(\text{response}|\text{selection}) = \pi_i \phi_i$$

Based on the available information for respondents and non-respondents, we estimated the non-response probability weight.

$$IPW_{HbA1c} = 1/\text{probability}$$

I used the household weight to generate the new weight, as the household weight is an individual sample weight and we have taken the inverse probability of selection. Hence, the new weight can be written:

$$\text{New weight} = \text{household weight} \times IPW_{HbA1c}$$

Accordingly, both descriptive and analytical analyses were presented using four different weighting scenarios: unweighted; using the household weight; using the lab weight; and using the newly generated weight correcting for missing HbA_{1c} data. However, after comparing the resultant values in Chapter 5, I found that household weight appeared the most appropriate weight to apply in subsequent analytical chapters.

4.3.4 Binary logistic regression model

The binary logistic regression model was used throughout to investigate the association between education level and T2DM outcomes. The logistic model is a generalised linear model (GLM) with binary outcome variables. GLMs are applied in cases where the outcome measure is a binary response and are appropriate here given the nature of the study outcome variables (Agresti and Kateri, 2011). The binary outcome variable is denoted as y_i for all individuals i , where $y_i = 0$ or 1 and π_i is the probability of y_i .

It assumed that there are J independent variables denoted as x_{ji} for each subject in the data. The logistic model in this analysis is then described by the following equation:

$$f(\pi_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} \dots + \beta_j x_{ji} \quad (4.1)$$

$f(\pi_i)$ = The transformation function of π_i , which requires a range for $\pi_i = (0, 1)$. The transformation function needs to find the link function in order to identify logit to avoid misinterpretation of the results, following the function is the logit transformation:

$$f(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) \quad (4.2)$$

Where the odds are $\left(\frac{\pi_i}{1-\pi_i}\right)$ of $y_i = 1$

Then, the binary logistic model by using the logit link function is:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_j x_{ji} \quad (4.3)$$

The expression is required to obtain the odds of y_i

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_j x_{ji})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_j x_{ji})} \quad (4.4)$$

Where i are the survey participants and $x_{1i}, x_{2i}, \dots, x_{ji}$ represents the study explanatory variables. After fitting the model, the interpretation of each β_j in the model is presented as an odds ratios (OR) with 95% confidence intervals. The odds ratio is calculated to identify the most important groups related to DM outcomes for targeted policy interventions.

4.4 Conclusion

This chapter provides a detailed description of the research data sources and main statistical method used in this thesis, along with the ethical considerations relating to the study. The key feature of the SHIS database that underpins this study is that it was based on a cross-sectional multi-stage research design. Finally, I have presented the sampling method, data collection, weighting approach, along with a rationale for the choice of analytical methods. The next four chapters detail the analysis and present the findings for the four main strands of this study.

Chapter 5 The Prevalence of Type Two Diabetes Mellitus in Saudi Arabia: Quantifying the Association of Socio- economic Factors and Predicting Future Trends

5.1 Introduction

Worldwide, the prevalence of diabetes has increased dramatically over the last 40 years. According to a recent diabetes report by the World Health Organisation (WHO, 2016a), the EMR has the highest prevalence of DM (13.4%), and the prevalence of diabetes in Saudi Arabia is one of the highest in the world (Al-Rasheedi, 2015). As mentioned in Chapter 3, the GCC countries have faced a rapid socio-economic transition that has resulted in lifestyle changes with consequent impact on patterns of disease (Al-Rasheedi, 2015), in particular the high prevalence of DM in the Arabian Gulf region (Bener et al., 2009). As a consequence, Saudi Arabia is ranked seventh in the list of diabetes prevalence among countries worldwide (Guariguata et al., 2014). Past diabetes research in Saudi Arabia and the GCC countries investigated the association between T2DM and demographic factors such as gender, age, and health and behavioural factors such as physical inactivity and obesity (see Chapter 2).

However, few studies in these areas have explored the relationship between socio-economic factors and T2DM. This is a missed opportunity, as exploring the relationship between education and DM in Saudi Arabia will help policy makers to identify interventions that effectively target high-risk groups. The need for such evidence-based guidelines for the management and prevention of T2DM was recommended in the 2016 WHO Global Report on Diabetes (WHO, 2016a). Healthy diet and physical activity were identified as key areas of focus, particularly for lower socio-economic groups. Education is another critical factor and is perhaps the strongest socio-economic predictor of health outcomes (Dahlgren and Whitehead, 1991, Marmot and Health, 2007). It has unique potential as a population intervention, as education remains stable over lifetime compared with income and employment status, acting as a socio-economic ‘vaccination’ against DM and other related diseases. Education improves an individual’s chances of employment, increases income and is related to a higher standard of living (Murad et al., 2014, Ajlouni et al., 2008, Hwang and Shon, 2014).

In some cases, this might extend to better access to the health care services, and healthier diet and activity levels, however the relationship between education and diet might be more complex during economic transition, as increased income allows access to more and a wider range of food. Most studies of the association between socio-economic factors and diabetes have focused on education level. Highly educated individuals are more likely to be aware of diabetes and more likely to check their health status routinely. If diagnosed, they tend to monitor their blood glucose better than less well-educated individuals. Several studies have shown that as a group, less well-educated individuals are both more likely to be diabetic and also to be unaware that they have the disease (the previously unknown diabetes group). They are less likely to have knowledge of diabetes mellitus symptoms and how they can control it well (Majeed et al., 2014, Bener et al., 2009, Ajlouni et al., 2008).

Any epidemiological study of diabetes requires a clear definition of disease status. Critical to this study is the recognition that all diabetic patients undergo a journey from unaffected status, through pre-diabetic, undiagnosed, and diagnosed phases to living with the disease and its comorbidities. The transition between phases and eventual outcomes are influenced both by genetic and socio-economic factors. There are several outcomes that will be investigated in this thesis, each looking at a different definition of affected status, as mentioned in Chapter 4. This chapter considers whether an individual has diabetes or not, irrespective of whether they have been formally diagnosed. This will be termed **‘all diabetes’ or ‘prevalence’**⁵¹, indicating if someone has previously been diagnosed with diabetes or has been identified retrospectively to have the disease through the HbA_{1c} test. Therefore, this outcome essentially differentiates between non-diabetics and all diabetics.

Additionally, this chapter will provide projections for T2DM prevalence and identify the most important factors to address in reducing the high burden of T2DM in the Kingdom. Specifically, I investigate the association between educational levels and the prevalence / development of T2DM amongst the Saudi adult population and explore how these associations are modified by gender. The study considers a set of risk factors including socio-economic factors (e.g. employment status), demographic factors (e.g. age), lifestyle factors (e.g. obesity), and administrative region and their effect on the association between the prevalence of T2DM and education level.

⁵¹ Prevalence or all diabetes included diagnosed and undiagnosed diabetes cases

One complication encountered in this analysis concerned the weighting factors applied to the data to reflect missing information. In the SHIS database, the weighting factor applied to lab data only addressed the attendance at the clinic and the taking of a blood sample, and not any subsequent missing HbA_{1c} results if a blood sample was taken. I therefore generated a new weight to allow for participants who did not have a HbA_{1c} result despite attending the clinic and providing a blood sample as discussed in Chapter 4. In this chapter, I compare the results of the association between education level and diabetes based on four different weighting methods: unweighted, using lab weight, the household weight and the new weight to examine to what extent the results are influenced by the new weight. This chapter has three main aims: the **first** is to assess the prevalence estimates of diabetes using weighted and unweighted analyses. The **second aim** is to investigate the association between the socio-economic factors and the prevalence of T2DM and identify the interaction effects of education-sex and education-household income on this relationship. These two aims address two research questions:

1. How does the prevalence of T2DM vary with and without survey weights?
2. What is the association between education level and the **prevalence** of T2DM in Saudi Arabia? More specifically, is this association modified by sex? How do this association effects when controlled for other factors including demographic, administrative region, socio-economic and lifestyle factors?

The **third aim** is to predict and quantify future projections and estimates of T2DM prevalence and stratify these estimates by age-sex groups to address the need to help reduce the high burden of T2DM in the Kingdom. The third research question is therefore:

3. What are the future estimates for T2DM total prevalence in Saudi Arabia? How do these estimates vary by age and sex of the individuals?

Research hypothesis: I hypothesise that less well-educated individuals have a higher likelihood of being diabetic than well-educated individuals and that there is an education-gender effect on the prevalence of T2DM. Additionally, the prevalence of T2DM is expected to increase over the next five to ten years.

This chapter is divided into five main sections: **Section 5.2** reviews previous studies of diabetes projection models. **Section 5.3** addresses the methodology of the approach and explains the definition of the study variables and statistical methods that will address the chapter's aims. The results are presented in **Section 5.4** including a sensitivity analysis, descriptive analysis, statistical analysis of disease prevalence and future projections of T2DM. The chapter concludes with a discussion and summary of conclusions in **Section 5.5**.

5.2 Background and rationale

This section focuses on studies that have explicitly addressed future projections of diabetes prevalence as opposed to risk factors for the disease. In Chapter 2, I reviewed global estimates of diabetes prevalence, both historic and current. Additionally, a number of country specific studies have attempted to project future trends in diabetes prevalence. To identify such studies, I carried out a systematic literature search using Medline, Science Direct, and Google Scholar covering the period 1990-2018 using the search terms: ‘diabetes’, ‘type two diabetes’, ‘model’, ‘diabetes model’, ‘prevalence’, ‘project’, ‘predict’, ‘future’, ‘forecast’ in article titles, abstract and keywords. Articles included in this study used projection models for diabetes prevalence that were based on total population or large cohort datasets of T2DM (see **Table 5.1**). I excluded any irrelevant articles after reviewing and skimming title and abstract of each article. I excluded global predictions of diabetes prevalence as these were discussed in Chapter 2. Note that the final list of studies differs markedly with respect to data sets, projection model type, and the predictors used to generate the diabetes model.

The majority of studies used demographic factors, incidence, migration, mortality, and high-risk diabetes proportion trends as predictors used to perform the projection model. Others used diabetes risk factors such as population trends in the proportion of overweight individuals, obesity, physical activity, and smoking. Some studies only predicted total diabetes prevalence (diagnosed and undiagnosed) for the population; others focused on more specific aspects such as diagnosed diabetes, diabetes mortality, pre-diabetes as predictor of high-risk of diabetes, or healthcare costs for diabetic patients. Overall, projections for most MENA region countries predicted higher T2DM prevalence than Western countries.

One of the limited number of projections for Saudi Arabia - the “Saudi Diabetes Impact Model” by Al-Quwaidhi et al. (2014a) is of particular relevance, and I will consider it here in more detail. The main data sources used in this study were the Saudi population structure by age group and sex obtained from the Central Department of Statistics and Information (CDSI), United Nations population estimates, and trends of diabetes prevalence based on obesity and smoking. The population structure in each census was divided by sex and 5-year age groups (0-4, 5-9, 10-4 up to 80+). The Al-Quwaidhi study estimated the total diabetes prevalence for the adult Saudi population (age ≥ 25) for the years 1992 – 2022 by filling the gaps in available census data (1993-2003, 2005-2009, 2011-2014, 2016-2019, and 2021-2022) using a linear interpolation approach (Al Quwaidhi, 2013).

This approach estimates unknown values between two time points based on a linear increase assumption. For example, to estimate population in 1999 where the available flanking values are in 1997 and 2005, the required formula is $(7 * \alpha_1 + \alpha_2) / 8$, where α_1 is the population size in 1997 and α_2 is the population size in 2005, and 7 is the gap of years between 1997 -2005. This formula is applied for each age group at each single year. Other data for 2010, 2015, and 2020 were obtained from the World Population Prospects, published by the United Nations Department of Economic and Social Affairs. These datasets are available as Excel spreadsheets using different indicators and estimation methods up to 2100, with data for the Saudi population based on population probabilistic projections. The prevalence of T2DM should be provided as age-sex specific rates; this assumption is necessary in order to obtain standardised age-sex prevalence rates for each year and for all available studies. Similarly, the age group categories need to be similar for included studies.

The author used data from neighbouring countries (e.g. Oman) to deal with missing age groups using the linear extrapolation approach, however this assumes that age-specific prevalence rates were similar in these countries (Al-Lawati and Jousilahti, 2004). This approach is probably reasonable given that the GCC countries share similar environment, socio-economic status and lifestyle characteristics. (Al-Quwaidhi et al., 2014a) developed their T2DM prediction model using 1992 as starting year and diabetes prevalence estimates extracted from a published nationwide study (Warsy and El Hazmi, 1999). Trends in obesity were obtained from the same study and two additional Saudi datasets (Warsy and El Hazmi, 1999, Al-Nozha et al., 2005, Al-Hamdan et al., 2005).

For analysis of trends in obesity, (Al-Quwaidhi et al., 2014a) used a body mass index of greater than or equal to 30 kg/m^2 to classify obese individuals as this was the value used in most similar studies. Missing data were estimated, again using linear interpolation. Future estimates were based on linear extrapolation to predict values outside the range of the dataset. The study used sensitivity analysis to estimate the uncertainty introduced by assuming a linear increase in obesity and T2DM. The results suggested that total diabetes prevalence would increase from 8.5% in 1992, 29.2% in 2011, 31.4% in 2013 to a projected 44.1% in 2022. Correspondingly, the number of people with diabetes is projected to increase from 555,000 in 1992 to 7.4 million in 2022. In conclusion, the Al-Quwaidhi study provides evidence that the prevalence of diabetes in Saudi Arabia will continue to increase based on trends in obesity and smoking status through to 2022.

In this chapter, I will apply similar methods to extend these projections to 2030. Some studies have projected diabetes prevalence based on previous surveys. For example, the diabetes prevalence projection model for England (Holman et al., 2011) predicted both diagnosed and undiagnosed diabetes (total diabetes prevalence) in those aged 16 or over through until 2030. The main data set used for input into this model was the current English population data set by age, sex, ethnicity, and deprivation from the Office for National Statistics (ONS) for 2002 – 2007. The authors extrapolated the age-sex specific trends for 2002 – 2007 using linear regression. To construct their model, the authors extrapolated these trends, including ethnicity, from 2008 to 2030. The second dataset was the age-sex specific prevalence rates for diagnosed diabetes based on physician reported data, while undiagnosed diabetes was defined using HbA1c diagnostic criteria based on data extracted from the Health Survey of England (HSE, 2006). Overweight and obesity prevalence trends were also used to generate projections of diabetes prevalence projections. Overall, the model estimated that the prevalence of diabetes would rise from 7.4% to 9.5% (2010 – 2030). The prevalence estimates were highest for men, and for individuals with a South Asian ethnic origin.

Other projections of T2DM prevalence have relied on multi-state or series models. For example, a US study (Boyle et al., 2010) projected trends of diabetes to 2050 among adults aged 18 -79 years. This was a four-state model that included diagnosed diabetes, undiagnosed diabetes and no diabetes divided into two states (low-risk: normal, and high-risk: prediabetes). The dataset used for input was from the US Census Bureau. Diagnosed diabetes estimates were obtained from the Centre for the Disease Control and Prevention (1980 – 2007) and projections were estimated using logistic regression. The results suggested that overall diabetes prevalence would rise from 14% to 21% over the period 2010 – 2050.

Table 5.1 Summary of diabetes prevalence projected models

Study	Setting and Year of survey	Age group	Data output	Data input	predictors used to perform the projection	Projection model type	Sensitivity and uncertainty analysis	Validation	Key findings
Narayan et al. (2006)	United States - National Health Interview Survey (1984–2004)	US total population	Projected prevalence of diagnosed diabetes (2005 - 2050)	the age-, sex-, and race/ethnicity-specific prevalence and incidence of diabetes	Incidence rate of diabetes in 2004	Markov model three states (no diagnosed diabetes, diagnosed diabetes, and death)	Model transition rates using 95%CI	Not mentioned	- The prevalence of diabetes increased from 5.62 to 12.0% (2005 – 2050). - The number of people with diabetes (men: 7.59 million to 20.81 million; women: 8.59 million to 27.47 million). - The highest projections for those aged 65 – 74 and 75 + (women 75+: 1.72 to 9.80 million; men 75+: 1.15 to 6.01 million).
Mainous et al. (2007)	United States - National Health and Nutrition Examination Survey (NHANES) II mortality survey (1976–1992), the NHANES III (1988–1994) and the NHANES 1999–2002.	US population aged > 20 years	Future diabetes (diagnosed and undiagnosed) burden for the adult US Population in 2011, 2021, and 2031.	- Prevalence of total diabetes (diagnosed and undiagnosed) and prevalence of diagnosed diabetes projection. - Diabetes mortality estimates (all-cause mortality for diagnosed and undiagnosed diabetes).	Demographic factors, incidence, migration, mortality, and high-risk diabetes proportion trends	Multi-state Markov model	increases in the proportion of persons at high risk by 10, 20 and 30%, as well as estimates based on decreases in the proportion of persons at high risk for results.	Not mentioned	- The prevalence of diabetes expected to rise from 11.5% to 13.5% (2011 – 2021). - The total number of people with diabetes predicted to increase 25.4 million in 2011, 32.6 million in 2021, and 37.3 million in 2031.
Magliano et al. (2008)	Australia - Australian Diabetes, Obesity and Lifestyle population-based Survey (1999 – 2000), and national mortality data	Australian adult population aged ≥ 25 years	Lifetime risk of diabetes Number of years lived with and without diabetes Projected prevalence of diabetes to 2025	- Age-sex specific incidence rates of diabetes. - All-cause mortality in those with or without diabetes.	Demographic changes Incidence and mortality rates	Multi-state lifetables	Age-specific incidence rates of diabetes, prevalence diabetes, and RR mortality	Not mentioned	The prevalence of diabetes expected to rise from 7.6% to 11.4% (2000 – 2025).
Lau et al. (2011)	Canada - Four Alberta Health and Wellness databases (1995 - 2007)	Canadian population aged ≥ 20	Projected prevalence (diagnosed) diabetes. Projected health care costs of diabetes. (2008 – 2035)	- Age-sex specific diabetes prevalence. - Historical age-sex specific incidence and mortality rates. - Future projections of prevalence, incidence and mortality. - The current and future projected estimates of population structure. - Health care costs data.	Demographic changes Incidence and mortality changes	Life table model	Not mentioned	Not mentioned	- The prevalence of diabetes projected to increase from 4.5% to 11.1% (2007 – 2035). - The highest projection observed among individuals who aged 75+.

Table 5.1 Summary of diabetes prevalence projected models

Study	Setting and Year of survey	Age group	Data output	Data input	predictors used to perform the projection	Projection model type	Sensitivity and uncertainty analysis	Validation	Key findings
Huang et al. (2009)	United States - the Centres for Disease Control and Prevention's National Health and Nutrition Examination Survey and National Health Interview Survey.	Diabetic male and female aged 24 – 85 years	Projected prevalence and incidence of obesity and diabetes Diabetes care and complications costs (2009 - 2034)	- Prevalence and incidence of diabetes. - Total population structure and projections (age-sex-ethnicity). - Population mortality data. - Estimated of future total health care costs. - Distribution of BMI categories. - Yearly transitions across BMI categories.	Demographic changes Trends of population rates of overweight and obesity.	Markov model	Not mentioned	Not mentioned	The number of people with diabetes expected to rise from 23.7 million to 44.1 million (2009 – 2034).
Honeycutt et al. (2003)	United States - National Health Interview Survey (NHIS) 1984–2000	Diagnosed diabetes (US men and women) US total population	Projected the number of people with diagnosed diabetes and diagnosed diabetes prevalence (2000 - 2050) three states (diagnosed, not diagnosed, and death)	- Estimated diagnosed diabetes prevalence and incidence. - The relative risk of mortality from diabetes compared with no diabetes. - Estimates of current population structure (age-sex-ethnicity), live births, net migration. - The mortality rate of the general population.	Demographic changes	Markov model	Model run for prevalence's, RR, and incidences 2000 using Lower and Upper bounds of 95%	Compared the results with previous projected model	- The diagnosed diabetes prevalence expected to rise from 4.4% to 9.7% (2000 – 2050). - The number of individuals with diabetes predicted to increase from 12 million to 39 million.
Holman et al. (2011)	United Kingdom - Health Survey for England 2006	England population aged ≥ 16 (diagnosed and undiagnosed)	Projected diabetes prevalence (2010 - 2030)	- The current projected estimates of population structure (age-sex-ethnicity). - Age- and sex-specific prevalence rates for self-reported physician diagnosed diabetes. - Glycated haemoglobin (HbA1c) results used to identify undiagnosed diabetes prevalence. - Future projections and prevalence of overweight and obesity.	Demographic changes. Trends of population rates of overweight and obesity.	Not specified	the Prevalence estimates using multivariable sensitivity analysis.	Model results compared with results of the model from the general practitioner diabetes Registers and the results of the PBS Diabetes Prevalence Model.	- The prevalence of diabetes predicted to rise from 7.4% to 9.5% (2010 – 2030). - The total number of people with diabetes expected to rise from 3.1 million to 4.6 million.
Helms (1992)	United States	US total population	Projected the number and prevalence of diabetes (1990 – 2050)	- US population data and projections - Age- prevalence of diabetes 1987 - Age-incidence rate of diabetes 1987	Estimated incidence of diabetes	Not specified	Not mentioned	Not mentioned	The number of people with diabetes increased from 6.5 million to 11.6 million (1987 – 2030).

Table 5.1 Summary of diabetes prevalence projected models

Study	Setting and Year of survey	Age group	Data output	Data input	Predictors used to perform the projection	Projection model type	Sensitivity and uncertainty analysis	Validation	Key findings
Boyle et al. (2010)	United States - US Census Bureau	US population aged 18-79 years	Projected annual diagnosed diabetes (2008 - 2050); projected annual diabetes prevalence (2000 - 2050)	- Current and projected estimates of diabetes incidence, and prevalence of diabetes and prediabetes. - Current and projected estimates of mortality rates, net migration, and births.	Incidence and mortality of diabetes trends.	Multiple states dynamic models	Different values based on mortality risk and Incidence.	Not mentioned	- The diabetes prevalence projected to rise from 14% to 21% (2010 – 2050). - The number with diagnosed diabetes projected to increase from 11 million to 29 million.
Bagust et al. (2002)	United Kingdom - combining historical population estimates for the UK covering the period 1931–96	UK population aged 20–59	Projected incidence and prevalence of T2DM. The trends in the levels of diabetes-related complications. The annual health care costs of T2DM and its complications. (2000 – 2060)	- Projected total population structure (age-sex). - Prevalence of diabetes complications. - Annual costs of T2DM population	Demographic changes.	Markov model	Results presented used standard error of proportional reduction in cardiovascular mortality.	Not mentioned	The total diabetes prevalence will rise by 11% maximum in 2030 for those aged 30+, then decrease to 8.5% by 2060.
(Ruwaard et al., 1993)	Netherlands	Dutch total population	Projected number of diabetic individuals and prevalence of diabetes (1980 - 2005)	Estimates and future projections of prevalence and incidence of diabetes, and life expectancy.	Trends of diabetes incidence and population life expectancy	Not specified	Results based on diabetes prevalence, incidence, and life expectancy reduction.	Compared the results with previous prevalence estimates	The number of diabetic patients expected to rise from 244,000 in 1990 to 355,000 in 2005.
(Martinsen et al., 2006)	Greenland - Greenland Health Survey (1993 - 1994) Health Survey (1999 - 2001)	Greenlandic population aged ≥ 18	Projected prevalence of T2DM (1999 - 2014); and Projected prevalence of complications (1999 – 2014)	- Current and projected estimates of population structure aged ≥ 35 years. - Current and projected estimates of obesity prevalence. - Current and projected estimates of diabetes complications.	Demographic changes Obesity trends changes	Linear function	Based on two scenarios of BMI distribution trends	Not reported	- The number of people with diabetes predicted to rise from 3,100 to 4,600 (1999 - 2014). - The prevalence of diabetes expected increase for women: 10% to 23%; but decrease for men: 12% to 11%.
(Andersson et al., 2015)	Sweden	National population aged ≥ 20	Projected diabetes prevalence (2013 – 2050)	- Prevalence diabetes estimates aged ≥ 20 years. - Previous data on relative mortality with diabetes, and demographic projections.	Annual forecasts of Sweden population	Not specified	Based on different assumptions	Not reported	The prevalence of diabetes expected increase from 6.8% in 2013 to 10.4% in 2050.

Table 5.1 Summary of diabetes prevalence projected models

Study	Setting and Year of survey	Age group	Data output	Data input	Predictors used to perform the projection	Projection model type	Sensitivity and uncertainty analysis	Validation	Key findings
Al-Quwaidhi et al. (2014a)	Saudi Arabia - Central Department of Statistics and Information and the United Nations population estimates	Saudis aged ≥ 25 years old	Projected prevalence of T2DM (1992 - 2022)	- Saudi population structure by age group and sex. - The trends in the prevalence of two risk factors obesity and smoking.	Demographic changes Obesity and smoking trends changes	Discrete-state Markov model SAUDI IMPACT	Analysis of extremes methods and Scenario analysis	Compared the results with previous prevalence estimates	- The total diabetes prevalence estimated to rise from 8.5% in 1992, 29.2% in 2011, 31.4% in 2013 and 44.1% in 2022. - The number of people with diabetes expected to increase from 555,000 to 7.4 million (1992 – 2022).
Awad et al. (2018)	Qatar - Epidemiological and demographic history data	Qataris aged 15 – 64 years old	- Predict trends in T2DM prevalence and incidence (2012 - 2050). - Estimate health expenditures directly attributed to T2DM. - Evaluate the key risk factors of T2DM.	- Prevalence and incidence of diabetes - Total population structure and projections (age-sex). - Projected of future total health expenditures for T2DM. - Distribution of BMI, smoking and physical activity categories.	Demographic changes Obesity, physical activity and smoking trends changes	Mathematical model based on sex, age, risk factors status and T2DM status.	Relative risk (RR) of mortality in T2DM compared to the general population and Monte Carlo sampling for CI ranges.	--	- The prevalence of diabetes expected to rise from 16.2% to 24.0% (2012 – 2050); women: 15.9% to 22.9%; men: 17.5% to 24.9%. - The highest projected group was 45 – 54 years old. - The total number of people with diabetes predicted to rise 20,953 to 51,921.
Al Ali et al. (2013)	Syria - The Syrian Bureau of Statistics and United Nations	Syrians aged ≥ 25	Projected the prevalence of T2DM (2003- 2022)	- Syrian population structure by age group and sex. - The trends in the prevalence of two risk factors obesity and smoking.	Demographic changes Obesity and smoking trends changes	Discrete-state Markov model the IMPACT Diabetes Model	--	--	- The prevalence of diabetes projected to rise double from 10% to 21% (2003 – 2022); women: 12.2% to 23.6%; men: 7.7% to 19%. - The total number of people with diabetes expected to rise (686,195) to (2,944,814). - In women (423,694) to (1,611,729). - In men (262,500) to (1,333,084).
Saidi et al. (2015)	Tunisia - The Tunisian National Nutrition Survey 1996/97	Tunisian population aged ≥ 25 years old	Projected the prevalence of T2DM 1997 and 2027	- Tunisian population structure by age group and sex. - The trends in the prevalence of two risk factors obesity and smoking.	Demographic changes Obesity and smoking trends changes	Markov model the IMPACT Diabetes Model	Analysis of extremes methods	Compared the results with t2DM prevalence survey 2005	- The prevalence of diabetes expected to increase from 12.0% in 1997, 15.1% in 2005, and to 26.6% in 2027. - The number of people with diabetes projected to rise from (516,000) to (2,125,000). - In females: 12.9% in 1997, 14.1% in 2005 and 24.7% in 2027. - In males: 13.5% in 1997, 16.1% in 2005 and 28.6%.

5.3 Methods

5.3.1 Weighting method and sensitivity analysis

As noted above, I generated a new weight to provide unbiased and representative estimates of diabetes prevalence. The HbA_{1c} test has a high level of non-observed values, yet the SHIS weights only accounted for an individual failing to attend clinic for a blood test and did not account for a failure in analysis. As a result, this weight did not truly reflect missing HbA_{1c} test results. I therefore calculated an additional weight taking this into account to ensure that results are representative of the population. The new weight is based on the inverse probability weight (IPW) method of an observed HbA_{1c} test out of the total sample size (see Chapter 4).

This weight was generated after model selection based on sex, age, socio-economic factors (education level, employment status and household income), marital status, previously diagnosed diabetes, and region. Age, sex, and region factors were related to missing data and used to generate the new weight based on the selection procedure results. The HbA_{1c} test weight is defined as the IPW multiplied by the SHIS household weight. The aim of the calculation of this new HbA_{1c} test weight is to investigate if the absence of HbA_{1c} in the SHIS affects the estimates of diabetes prevalence and to derive a weight which could be used in the analysis of the association between education level and T2DM prevalence. The new weight is compared with the SHIS weights and unweighted results to assess its influence on the results.

5.3.2 The association between education level and prevalence of diabetes

In this chapter, I focus on two main analyses addressing the research questions defined in **Section 5.1**. The first investigates the association between education level and prevalence of diabetes based on whether individuals were previously diagnosed diabetes or not. This outcome variable is the result of cross tabulation between the HbA_{1c} test result: ($HbA_{1c} < 6.5\%$ or $HbA_{1c} \geq 6.5\%$) and the survey question: “Have you ever been diagnosed with diabetes by a doctor or health professional?” according to the classification table discussed in the previous section. Binary logistic regression was selected for the first analysis due to the dichotomous nature of our study outcome with the objective of understanding the relationship between the outcome variable and possible explanatory variables (see Chapter 4).

The analytical framework for this chapter assumes that socio-economic, demographic and lifestyle factors increase the likelihood of an individual being diabetic. This is based on the conceptual framework for the prevalence of T2DM presented in Chapter 2 and past literature studies on the relationship between the prevalence of T2DM and education level. For example, such factors might influence the likelihood of an individual developing diabetes through an effect on obesity. Individual demographic characteristics such as age, marital status and gender could affect an individual's socio-economic profile and the likelihood of being obese and diabetic.

In addition to individual level characteristics, previous studies indicate that regional factors also affect an individual's likelihood of being diabetic. All factors considered in the review chapter and available in the dataset were included in the analytical framework. Before applying the logistic regression model, I applied bivariate analysis to assess the significance of association between the prevalence of diabetes and explanatory variables using the Chi-square test with a significance threshold level of $p=0.05$. After checking for significant relationships, six logistic models were fitted for prevalence of diabetes, controlled for different socio-economic, demographic, lifestyle factors, and administrative region.

The following table presents the analytical framework and six logistic regression models. To address the research questions presented in **section 5.1**, model one will study the relationship between education level and prevalence of T2DM. In successive models, I then added other factors to the logistic model to see if they affected the strength of this association. **Model two** adds demographic factors; **model three** tested the effect of adding the region variable; **model four** added other socio-economic factors; **model five** added lifestyle factors; finally, **model six** added an interaction term to test for a gender effect. As mentioned above, the population eligible for the outcome used in this analysis contained 937 individuals with diabetes and 3,404 individuals who did not have the disease. A summary of the explanatory variables is presented in **Table 5.3**.

Table 5.2 Analytical frameworks for the prevalence of T2DM outcome

Main predictor	Model adjustment	Variables list based on SHIS database
Education level	Model 1: Unadjusted	
	Model 2: Demographic factors	Age, sex, and marital status
	Model 3: Geographical location	Administrative region
	Model 4: Socio-economic factors	Employment status and household income
	Model 5: Lifestyle factors	BMI, smoking, diet, and physical activity
	Model 6: Full adjusted + gender differences	Model 5 + interaction between education level and sex

Socio-economic factors

- **Education level** was categorised into six groups based on the years of education an individual had completed in or referenced to the Saudi Arabian education system. I classified those without formal education into two groups: those who cannot read or write, and those that can read or write but did not have formal education. The six categories for educational level were therefore: **Low education:** (1) unable to read and write; (2) no formal education; **Middle education:** (3) primary school education; (4) intermediate school education; **High education:** (5) high school education; (6) university degree and above (which included completed university, postgraduate degree, and technical training).
- **Employment status** was categorised into six groups: (1) employed; (2) unemployed; (3) student; (4) homemaker; (5) retired; and (6) self-employed. “Employed” status groups employees who worked in either the governmental or non-government sectors. However, in the next chapter this category will be further divided into public and private sectors employees. The unemployed group included those able to work but unemployed at the time of the survey; those who are unable to work and individuals who are in unpaid jobs such as volunteers. As in previous studies, I chose to exclude students, homemakers and retired so that they could be analysed separately.
- **Monthly household income estimate (MHIE)** status was collected in nine groups within the survey itself, with respondents identifying from a list about the category of income they belonged to. This was reclassified into five groups. **Lowest :** (1) less than 3,000 Riyals; (2) 3,000 Riyals to less than 5,000 Riyals; (3) 5,000 Riyals to less than 7,000 Riyals; **Low:** (4) 7,000 Riyals to less than 10,000 Riyals; **Middle:** (5) 10,000 Riyals to less than 15,000 Riyals; **High:** (6) 15,000 Riyals to less than 20,000 Riyals; **Highest:** (7) 20,000 Riyals to less than 30,000 Riyals; (8) 30,000 Riyals or more; **Missing:** (9) Unknown or declined to respond. Individuals who did not answer the household income question were classified as unknown or declined to respond to reduce the effect of the high level of missing data observed for the question (18.0% of responses to the income question were missing and classified into this group). Based on previous research into the reporting of income, this group is more likely to be made up from households from the highest or lowest income group (Kim et al., 2007). However, research on this topic has not been conducted for Saudi Arabia, and hence this assumption cannot be assured. Due to the potential heterogeneity in this category, interpretation of results for this group is difficult.

Demographic factors

Gender and age were considered as demographic factors, and marital status as a social factor. Age is one of main risk factors for diabetes. I treated age as a categorical variable split into categories to reflect previous observations that the middle-aged group had the highest diabetes prevalence (El Hazmi, 1998, Connolly et al., 2000). Age was therefore divided into six 10-year age groups. Marital status was classified into three main groups based on the marital status answer in the SHIS household questionnaire. The 'separated', 'divorced' and 'widowed' categories were merged into one group 'currently not married'.

Lifestyle factors

Obesity is one of main lifestyle indicators that is modified by physical activity and diet. Body mass index (BMI) was used as the measure of obesity and calculated as weight (kg)/ height (m^2). Participants were categorised into four main groups as per standard recommendations: (1) Underweight (BMI less than 18.5); (2) normal weight (BMI between 18.5 and 24.9) – the reference category; (3) overweight (BMI between 25.0 and 29.9); (4) obese (BMI more than or equal to 30.0). Furthermore, a measure for **diet** was calculated based on the daily servings of fruit, vegetables, chicken, and red meat. Servings per day were estimated from the dietary questionnaire as the sum of the average daily consumption of fruit, vegetables, red meat, and chicken. However, there are four dietary questions available in the SHIS dataset ("How many servings of fruit/vegetables/chicken/meat do you eat on one of those days?"), responses were incomplete with some 20% data missing. These respondents were classified into missing group. The level of physical activity⁵² (High or Moderate, Low and None) was used as a dependent variable. I used the International Physical Activity questionnaire to classify respondents into four groups of physical activity (Craig et al. 2003):

- Vigorous physical activity: A high level of physical activity is considered as a total of more than 450 minutes of moderate to intense activity per week.
- Moderate physical activity: A moderate level of physical activity is considered as a total of 150–450 minutes of moderate to intense activity per week.

⁵² **Survey questions:** Do you do any vigorous/moderate-intensity sports, fitness or recreational (leisure) activities that cause large increases in breathing or heart rate for at least 10 minutes continuously? In a typical week, on how many days do you do vigorous/moderate-intensity sports, fitness, or recreational activities? How much time do you spend doing vigorous/moderate-intensity sports, fitness, or recreational activities on a typical day?

- Some physical activity but insufficient to meet vigorous or moderate levels: A low level of physical activity is considered as a total of less than 150 minutes of moderate to intense activity per week.
- No physical activity.

Smoking is one of main behavioral factors that increases the likelihood of developing diabetes. Smoking status was classified into three categories: (1) never smoked; (2) ex-smoker; and (3) current smoker.

Community factors

The study used the **region** variable to explore regional difference in the three diabetes outcomes. Recent studies highlighted regional variation in the prevalence of diabetes within the same population (Al-Moosa et al., 2006, Al-Nuaim et al., 2012), where region was categorised based on the 13 administrative regions: (1) Al Riyadh; (2) Makkah; (3) Eastern Region; (4) Northern Borders; (5) Madinah; (6) Jazan; (7) Aseer; (8) Najran; (9) Qaseem; (10) Tabuk; (11) Hail; (12) Al-Jouf; (13) Al-Baha. In this study, the region variable was re-classified based on geographical and urbanisation criteria: **(1) Central region** (Al Riyadh and Qaseem); **(2) Eastern region** (reference category); **(3) Western region** (Makkah and Madinah); **(4) Northern region** (Northern borders, Al-Jouf, and Hail); **(5) Southern region** (Jazan, Aseer, Najran, Al-Baha). As mentioned in Chapter 4, the Tabuk region was excluded due to limited number of respondents who completed the lab stage.

All analyses were carried out using STATA software version 14 (Stata Corp. Inc, TX, USA), and Microsoft Excel 2013. Descriptive statistics were used to describe general characteristics. I used Chi-squared tests to assess the relationship between the “all diabetes” outcome measure (prevalence) and the study variables. The association between education level and the prevalence of diabetes was investigated based on a generalised binary logistic regression model and investigated before and after adjusting for demographic, lifestyle, and socio-economic factors (employment status and household income). Both unadjusted and adjusted odds ratios were estimated for education level.

Chapter 5

Most of the reference categories were chosen to be related to my primary interested based on the previous literature. As this chapter is focused on better understanding the risk factors of diabetic individuals, I selected as reference categories those groups expected to have a low risk of diabetes. With respect to the age variable, for example, I chose the youngest age group. Another important factor in choosing the reference category is based on sample size. In the case of the body mass index variable, for example, although both the underweight and normal groups were less likely to be diabetic, I chose the normal weight category as the reference as it provided the largest sample size.

Table 5.3 Study explanatory variables list

Study variables		Categories
Demographic factors	Sex	Dichotomous variable: 1= Male, 2= Female; Reference category: Female
	Age	Categorical variable: (1: 15 - 24, 2: 25 – 34, 3: 35 – 44, 4: 45 – 54, 5: 55 – 64, 6: 65+) <i>Reference category: (15 - 24)</i>
	Marital status	This was recoded into three categories: 1 = Currently Married 2 = Never married. 3 = Not currently married: Separated/divorced, and Widowed. <i>Reference category: Never married</i>
Socio-economic factors	Education level	Recoded into three categories: 1 = Low education level: cannot read or write, and no formal education. 2 = Middle education level: Primary and intermediate school education 3 = High education level: High school, university degree and above. <i>Reference category: High education level</i>
	Employment status	Coded into seven categories: 1 = Employee: Government and Non-government employee; 2 = Unemployed; 3 = Retired; 4 = Homeworker. 5 = Self-employed; 6 = Student. <i>Reference category: Employee</i>
	Monthly household income	Recoded into five categories: 1 = Lowest: 1= Less than 3,000 Riyals; 2=3,000 Riyals to less than 5,000 3=5,000 Riyals to less than 7,000; 2 = Low: 4=7,000 Riyals to less than 10,000 3 = Middle: 5= 10,000 Riyals to less than 15,000; 4 = High: 6=15,000 Riyals to less than 20,000 5 = Highest: 7=20,000 Riyals to less than 30,000; 8= 30,000 Riyals or more <i>Reference category: Lowest</i>
Lifestyle factors	Body Mass Index	Recoded into two categories: 1 = Underweight; 2 = Normal; 3 = Overweight; 4 = Obese <i>Reference category: Normal</i>
	Smoking status	Recoded into two categories: 1 = Currently smoker; 2 = Ex-smoker 3 = Never smoked <i>Reference category: Never smoked</i>
	Physical activity levels	Recoded into three categories: 1 = None; 2 = Low; 3 = Moderate; 4 = High <i>Reference category: High</i>
Administrative region	Region location	Recoded into five categories based on geographical location: 1 = Central region; 2 = Eastern region; 3 = Western region; 4 = Northern region 5 = Southern region <i>Reference category: Western region</i>

5.3.3 Projection of the prevalence of T2DM

In the second analysis, I predict the future trends of diabetes prevalence by age and gender through to 2030. This projection required reliable Saudi population estimates covering the relevant years. I therefore used demographic data from two independent sources - the United Nations population projections and the Cohort Component Projection (Hinde, 2014) - to construct a robust Saudi national population projection (for more details about data sources see Chapter 4). Combined with the historic prevalence data and my modelling of socio-economic factors, this allowed me to produce future estimates of diabetes prevalence using two independent approaches to increase confidence in the projections.

5.3.3.1 Method 1: United Nations population projections

In the first method I calculated percentages based on the United Nations data and the available Saudi census. I first calculated the percentage of the Saudi national population for each age-sex group in all available years. I then extended the calculation of these percentages for each of the projected years (2020, 2025, and 2030) calculated as follows. For example, to calculate the percentage of Saudi male nationals aged 5 – 9 years in 1992:

$$\frac{\text{the number of Saudi male aged 5–9 in 1992}}{\text{total number of males aged 5–9 in 1992}} = \frac{1,061,429}{1,299,829} \times 100 = 84.36\%$$

For my projections, therefore, there are two different possibilities. First, the percentage of Saudis remain constant at 2018 levels - the conservative assumption. A second possibility is that the any trend observed between 2016 and 2018 continues at the same rate. In order to calculate the age-sex group percentages for the projected years, I calculated the average difference between 2016/17 and 2017/18. The formula for average growth from 2016 – 2018 is given by:

$$\text{the percentage of Saudis in 2018} - \text{the percentage of Saudis in 2016}.$$

For example, for Saudi males aged 0 - 4 the average increase per year is 0.365. I applied this increase to estimate the percentage for this age group in 2020, 2025 and 2030. For this age group, the percentage of Saudis will be 74.64 in 2030. This calculation applied to all projected population years. The percentage for Saudis aged (0 – 4) in 2030:

$$\text{percentage of Saudis male aged (5 – 9)} - 12 \times \text{average growth of Saudis males aged (5 – 9)}.$$

After finishing this step, I then extracted the total Saudi population by age and sex from the United Nations website for these years. As the UN population projections include both Saudi and non-Saudi nationals, I applied the percentages calculated above to derive estimates for the age-sex stratified Saudi national population in 2020, 2025 and 2030. I applied these calculations twice to derive projections based on the conservative and non-conservative growth assumptions described above. Finally, I calculated the total prevalence of each year by divided the total number of diabetic patients over the total number of Saudi populations in specific year.

5.3.3.2 Method 2: Cohort component projection

In this analysis I used the Cohort Component Projection method to estimate the Saudi national population without considering fertility and migration (Hinde, 2014). This method starts with application of the survivorship ratio for the whole population obtained from the WHO life tables. The survivorship ratio formula is based on the L and T columns from the life table where L represents “*person years lived between x and x+n*” and T represents “*person years lived above age x*”. Two main assumptions are considered in this method for projecting population to 2023 and 2028. Firstly, those individuals who were aged 20-24 in 2018 will be 25-29 in 2023, so I used the number of individuals in 2018. I then applied the survivorship ratio I obtained from the WHO life tables in 2016 (see Chapter 4). For the survivorship ratio, I used the ratios from the L column. I then calculated the survivorship ratio from the L_x and T_x columns extracted from Saudi life table mentioned above. The survivorship ratio formula:

$$\text{Survivorship ratio} = S_{x-5,x}^F = \frac{L_x}{L_{x-5}}$$

However, as the oldest age group has a different survivorship ratio formula (65+) I used T rather than L from the life table):

$$\text{survivorship ratio (65 +)} = [N_{(60-64)} + N_{65+}] \frac{L_{65+}}{[L_{(60-64)} + L_{65+}]}$$

Thus, to calculate the number of individuals to survive the next 5-years, then the formula will be:

$$N_x(t + 5) = N_{x-5}(t) \times S_{x-5,x}^F$$

For example, calculating the survivorship ratio between 20-24 and 25-29 gives: $484550.3/487084.3 = 0.994797615$. In other words, 0.99479 of those aged 20-24 survive to 25-29, giving an estimate of 1,066,551 Saudis males in 2023. However as mentioned before there is a complication in that the available life tables represent the Saudi total population including immigrants, when I have assumed no migration. I therefore removed the effect of migration by applying the same survivorship ratios to the Saudi national population (estimated as described in the previous section) instead of the total population.

Additionally, as I aimed to project the population of those aged 25 years and above, I need to include in the analysis all individuals aged 15 – 19 years old in 2018, as they will be aged over 25 years old in 2028. The third step in the analysis after calculation of survivorship ratios and applying these to population estimates corrected for Saudi Nationals is to project the prevalence of T2DM using 2013 as starting point including all individuals aged 25 years or over. I carried out the analysis using Microsoft Excel 2013. Finally, it should be stressed that I have made five main assumptions in this projection of T2DM prevalence.

- 1) The available prevalence rates in the study were applied to the midpoint of the period range; for example, if a study was conducted between 1996 – 2000, I assumed that the estimates applied to 1998
- 2) I adjusted age intervals to 5-years to be consistent with past prevalence rates studies
- 3) The distribution of the Saudi population remains constant from 2016
- 4) Trends in T2DM prevalence identified in the surveys will continue linearly until 2030
- 5) Mortality rates remain constant

5.4 Results

5.4.1 Weighted and unweighted estimates of T2DM prevalence

The overall prevalence of all diabetes varied between 11.2% and 16.5% depending on the weighting factor used (see section 5.3.2.1 for a discussion of weighting factors). The unweighted percentage gave the highest prevalence at 16.5%. The new HbA_{1c} weight developed in this study estimated diabetes prevalence at 11.5%, which is slightly higher than that estimated in the SHIS report and using the lab weight given in the dataset, which was 11.2%.

These are shown in **Table 5.4**, which shows the prevalence of diabetes using each of the four weights: unweighted, lab weight, household weight only and the HbA_{1c} test weight. The household weight is designed to ensure that the results are representative for the Saudi population; the lab weight mitigates the influence of missing data for visiting the clinic, and the HbA_{1c} test weight corrects for missing HbA_{1c} test data. The association between education level and the diabetes prevalence based on each of the four weighting method scenarios is presented in **Appendix C**. Before the relationship is fully adjusted (**Model 1**), all four weighting methods show similar patterns of association. However, after adjustment for demographic factors (**Model 2**), the unweighted and household weight models have higher impact on the association between education level and diabetes than other weights.

In the case of both the lab weight and the new HbA_{1c} weight, the models were no longer significant, even however all four methods gave a similar pattern. It is possible that the results using the lab weight and the new HbA_{1c} weight are affected by the variables used to generate the weights, resulting in an over-fitted model. When using both the lab weight and the new HbA_{1c} weight, education was no longer significantly associated with diabetes prevalence after controls were added to the model. In light of these results, I decided to apply the new HbA_{1c} weight only to estimates of diabetes prevalence and use the household sampling weight for analysis of the association between education level and diabetes mellitus outcomes. This weight best reflects the sampling approach used to select individuals for the survey and is therefore most likely to provide representative results. It avoids the risk of over-fitting the model and is less likely to be distorted by attempts to correct for missing lab data.

Table 5.4 Prevalence of diabetes status in the Saudi Health Interview Survey data, KSA, 2013

Weight	Threshold between		
	Non-diabetics < 5.7 %	Pre-diabetes 5.7 – 6.49 %	Diabetics ≥ 6.5%
Unweighted	60.7	22.8	16.54
Household weight ⁵³	64.9	22.5	12.6
Lab weight ⁵⁴	67.8	21.1	11.2
HbA_{1c} test weight ⁵⁵	67.3	21.2	11.5

⁵³ The **household questionnaire weight** (first sampling weight) was an individual sampling weight for all participants who completed the household questionnaire.

⁵⁴ The **lab weight** (second sampling weight) was for participants who gave a blood sample at a laboratory and attended clinic visits.

⁵⁵ The **HbA_{1c} test weight** is an additional weight of missing data for HbA_{1c} test.

5.4.2 Association between education level and the prevalence of diabetes

5.4.2.1 Descriptive results

Table 5.5 describes the characteristics of the study population as applied to the prevalence of T2DM. A total of 4,341 individuals were eligible for inclusion using the 'all diabetes' outcome. Low-educated Individuals were more likely to be diabetic than high-educated individuals (38.8% vs 13.6%). Considering demographic factors, a higher percentage of male participants were diabetic (25.0% versus 18.5% for females). In addition, the likelihood of having diabetes increases with age from 6.6% in the 15-24 age group to 54.5% in those aged 65 and over. Diabetes was more prevalent in the 'not currently married' group (39.3%) than the 'never married' (7.9%) and 'currently married' (24.5%) groups. In relation to employment status, the retired, self-employed, and unemployed individuals had the highest likelihood of being diabetic (55.5%, 31.0%, and 27.8% respectively). Regarding lifestyle factors, ex-smokers were more likely to be diabetic than current smokers (33.7% vs 20.5%). Additionally, physically active individuals were less likely to be diabetic than inactive individuals (85.7% versus 14.3%), whereas obese individuals were more likely to be diabetic than individuals of normal weight (27.6% versus 14.5%).

With respect to the 'region' factor, individuals who lived in the Western region were less likely to be diabetic than those who lived in the Southern region (16.7% versus 25.5%). In keeping with these descriptive results, the bivariate analysed the significance of these relationships. As expected, demographic factors (age, gender, and marital status) showed significant associations with the prevalence of T2DM. In general, the results shown in **Table 5.5** mirror the relationships reported in the literature, with low-educated individuals having the greatest likelihood of being diabetic. Smoking status, physical activity and BMI all showed significant association with the prevalence of T2DM. The region variable also showed significant association with the prevalence of diabetes. This result followed previous studies in SA and other countries showing that there were regional trends in the prevalence of diabetes. Interestingly, although socio-economic factors generally showed significant associations with the prevalence of diabetes, this was not the case for household income ($p=0.136$). An important consideration relating to the household income dataset is that the number of null responses (unknown or refuse to answer) undermines confidence in any relationship identified between income and diabetes prevalence for those who did answer this question. A high proportion of missing values reduces the statistical power of this relationship such that it becomes weak or undetectable leading to non-significant results (Suresh and Chandrashekar, 2012). As well as the impact on statistical power, a more insidious problem is that of non-response bias, as a lack of response is unlikely to be randomly distributed across all survey participants. Hence, the results of this variable cannot be representative.

Table 5.5 Characteristics of survey participants of prevalence of T2DM (unweighted percentages)

Variables	Categories N (%)	Total 4,341	Normal 3,404 (78.4)	Diabetes 937 (21.6)	p
Sex	Female	2,285	1,864 (81.5)	423 (18.5)	<0.001
	Male	2,056	1,542 (75.0)	514 (25.0)	
Age	15 - 24	888	829 (93.4)	59 (6.6)	<0.001
	25 - 34	1,013	927 (91.5)	86 (8.5)	
	35 - 44	954	803 (84.2)	151 (15.8)	
	45 - 54	672	456 (67.9)	216 (32.1)	
	55 - 64	365	172 (47.1)	193 (52.9)	
	65 +	420	191 (45.5)	229 (54.5)	
Education level	Low education	1,065	652 (61.2)	413 (38.8)	<0.001
	Middle education	1,159	922 (79.5)	237 (20.5)	
	High education	2,114	1,827 (86.4)	287 (13.6)	
Employment status	Employed	1,472	1,207 (82.0)	265 (18.0)	<0.001
	Unemployed	539	390 (72.2)	150 (27.8)	
	Student	707	665 (94.1)	42 (5.9)	
	Homeworker	1,113	875 (78.6)	238 (21.4)	
	Retired	353	157 (44.5)	196 (55.5)	
	Self-employed	141	98 (69.0)	44 (31.0)	
	other	16	14 (87.5)	2 (12.5)	
Monthly Household income	Lowest	1,794	1,388 (77.4)	406 (22.6)	0.136
	Low	686	556 (81.1)	130 (18.9)	
	Middle	575	468 (81.4)	107 (18.6)	
	High	276	212 (76.8)	64 (23.2)	
	Highest	211	165 (78.2)	46 (21.8)	
	Unknown or decline to respond	799	615 (77.0)	184 (23.0)	
Region location	Central region	1,519	1,212 (79.8)	307 (20.2)	<0.001
	Eastern region	153	115 (75.2)	38 (24.8)	
	Western region	826	688 (83.3)	138 (16.7)	
	Northern region	652	502 (77.0)	150 (23.0)	
	Southern region	1,191	887 (74.5)	304 (25.5)	

Table 5.5 Characteristics of survey participants of the prevalence of T2DM (cont.)

Variables	Categories N (%)	Total 4,341	Normal 3,404 (78.4)	Diabetes 937 (21.6)	p
Marital status	Currently married	2,863	2,163 (75.5)	701 (24.5)	<0.001
	Never married	1,090	1,004 (92.1)	86 (7.9)	
	Divorced, Separated, and widowed	374	227 (60.7)	147 (39.3)	
	Unknown / decline to respond	14	11 (78.6)	3 (21.4)	
Smoking status	Never smoked	3,700	2,924 (79.0)	776 (21.0)	<0.001
	Ex-smoker	205	136 (66.3)	69 (33.7)	
	Current smoker	429	341 (79.5)	88 (20.5)	
	Missing	9	5 (55.6)	4 (44.4)	
Daily servings of fruit and vegetable	0 – 2.9	3,275	2,554 (78.0)	721 (22.0)	0.116
	3 – 4.9	117	86 (73.5)	31 (26.5)	
	5+	17	15 (88.2)	2 (11.8)	
	Missing	932	749 (80.4)	183 (19.6)	
Daily servings of meat and chicken	0 – 0.9	1,414	1,142 (80.7)	273 (19.3)	0.097
	1 – 1.9	1,320	1,016 (76.9)	305 (23.1)	
	2 – 2.9	554	436 (78.7)	118 (21.3)	
	3+	220	174 (79.1)	46 (20.9)	
	Missing	833	638 (76.6)	195 (23.4)	
Levels of physical activity	High	133	114 (85.7)	19 (14.3)	0.006
	Moderate	221	182 (82.4)	39 (17.7)	
	Low	151	119 (78.8)	32 (21.2)	
	None	3,607	2,794 (77.5)	813 (22.5)	
	Missing	229	195 (85.2)	34 (14.8)	
Body mass index	Underweight (less than 18.5)	172	167 (97.1)	5 (2.9)	<0.001
	Normal weight (between 18.5 and 24.9)	1,097	938 (85.5)	159 (14.5)	
	Overweight (between 25.0 and 29.9)	1,364	1,058 (77.6)	306 (22.4)	
	Obese (greater than or equal 30.0)	1,669	1,209 (72.4)	460 (27.6)	
	Missing	41	34 (82.9)	7 (17.1)	

5.4.2.2 Analytical results

The results in **Table 5.6** show the odds ratios (ORs) and 95% confidence interval (CI) from weighted generalised binary logit regression models of the unadjusted and adjusted association between the prevalence of T2DM and education level. This analysis was fitted to six nested logistic models based on the analytical framework presented in **Table 5.2**. The first model explored the association between education level and the prevalence of T2DM. This relationship was then adjusted based on a series of logistic models by controlling for several demographic, socio-economic and lifestyle factors. The results of these logistic regression models are presented in **Table 5.6**. The unadjusted ORs from the first model indicate a significant inverse association between educational level and the prevalence of T2DM, with low-educated individuals more likely to be diabetic than better educated individuals (OR: 3.759, 95% CI: 2.936, 4.814, $p < 0.05$).

After adjusting for age group and sex and marital status in model two, all three of the demographic risk factors used in this model (individual age, sex and marital status) showed a significant relationship with the prevalence of T2DM. Individuals aged 55 – 64 years old were more likely to be diabetic than those who aged 15 – 24 years old (OR: 8.719, 95% CI: 4.94, 15.39, $p < 0.05$). The model also revealed that individuals in the ‘currently unmarried’ group were more likely to be diabetic than individuals in the ‘never married’ group (OR: 2.142, 95% CI: 1.234, 3.716, $p < 0.05$). The magnitude of OR estimates for educational level were less than those from model one, but there was an association between educational level and the prevalence of diabetes. This effect persisted after adjusting for region in model three. This model showed significant relationship between the ‘region’ variable and prevalence of T2DM, with individuals who lived in the Southern region more likely to have diabetes than those who lived in Western region (OR: 2.047, 95% CI: 1.515, 2.766, $p < 0.05$).

An inverse association between education level and the prevalence of T2DM was observed after controlling the third model using other socio-economic factors (model four). The odds of diabetes in low-educated individuals was 1.793 times higher than in those with a high-level of education (95%CI: 1.242, 2.588). Employment status also had significant a relationship with the prevalence of T2DM, however this only extended to the ‘students’ group who were less likely to be diabetic than employed individuals (OR: 0.389, 95%CI: 0.186, 0.812, $p < 0.05$). In contrast, monthly household income had no significant association with the prevalence of T2DM in any of the subsequent logistic regression models. Model five presents the regression analysis, which tests whether there is an association between the prevalence of T2DM and education level after correcting for lifestyle factors. The odds of being diabetic among low-educated individuals was 1.635 times higher than for well-educated individuals (95%CI: 1.008, 2.653).

Of the three lifestyle factors controlled in this model, two (smoking status and physical activity levels) had no significant association with the prevalence of T2DM. However, the body mass index variable did show significant association with the prevalence of T2DM, with obese individuals more likely to be diabetic than normal weight individuals (OR: 1.594, 95%CI: 1.150, 2.209, $p < 0.05$). The final model showed that there is no significant interaction between education level and gender or education and household income within the diabetes sample. Education level, age, region, employment status and body mass index showed a stable relationship across all the models tested. The high proportion of responses falling into the lowest income category is due to the way the household income variable is classified. There is no standard approach to the classification of household income in Saudi Arabia, but as described above the income categories were combined into five groups. Additionally, as the total sample size was reduced after restricting the analysis to those with laboratory reported diabetes, the three highest income groups were further combined to assess the relationship between household income and diabetes prevalence.

My results highlighted that the relationship between variables can become weak or non-significant after adjusting the models for other independent variables. However, additions to or adjusting the models could influence the statistical power and significance level of any relationships. For example, the marital status variable was significantly associated in the second, third and fourth models. However, the association became non-significant in the last two models due to the smaller sample size (resulting from declined, unknown or missing responses) leading to a reduction in the statistical power to detect a relationship. Thus, completed sample variables such as age and region tend to have stronger associations and appear significant in all adjusted models. In contrast, the inclusion of additional groups for missing data or those who decline to answer a specific question are likely to reduce power to detect association and the observed significance level.

There are several possible reasons for the changes in the estimates seen with the different models. First, they could result from changes in the sample size between models. However, this is unlikely to be the case given the sample size for this study. A second possible reason is that weighting can change the value of the coefficients. However, in my study, the most likely explanation is that adding new predictors to the model changed all the other coefficients. As some of these additional predictors are confounders, their introduction could reduce the strength of the relationship between education level and T2DM prevalence. This is in fact what I observed, as estimates of the effect of education level decreased in successive models due to addition of predictors such as age and gender, both of which are known confounders. My estimates for other predictors also showed small changes as the model was adjusted.

Table 5.6 Generalised binary logistic regression odds ratios and 95% confidence interval of the prevalence of T2DM (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Education level												
High education (reference)	1		1		1		1		1		1	
Low education	3.759 (2.936 - 4.814)	0.000	1.504 (1.082 - 2.090)	0.015	1.511 (1.087 - 2.099)	0.014	1.793 (1.242 - 2.588)	0.002	1.685 (1.179 - 2.407)	0.004	1.635 (1.008 - 2.653)	0.047
Middle education	1.501 (1.157 - 1.948)	0.002	1.295 (0.971 - 1.727)	0.079	1.318 (0.982 - 1.767)	0.065	1.465 (1.061 - 2.024)	0.020	1.457 (1.051 - 2.025)	0.024	1.231 (0.723 - 2.095)	0.443
Age group												
15 – 24 (reference)			1		1		1		1		1	
25 - 34			1.283 (0.758 - 2.170)	0.354	1.289 (0.771 - 2.155)	0.332	0.805 (0.438 - 1.481)	0.486	0.728 (0.393 - 1.347)	0.312	0.726 (0.392 - 1.344)	0.308
35 – 44			1.862 (1.104 - 3.141)	0.020	1.894 (1.133 - 3.167)	0.015	1.174 (0.644 - 2.137)	0.601	1.009 (0.552 - 1.846)	0.977	1.009 (0.552 - 1.847)	0.976
45 - 54			4.082 (2.422 - 6.882)	0.000	4.134 (2.468 - 6.925)	0.000	2.467 (1.359 - 4.478)	0.003	2.112 (1.151 - 3.873)	0.016	2.104 (1.145 - 3.867)	0.017
55 - 64			8.719 (4.940 - 15.39)	0.000	9.296 (5.297 - 16.31)	0.000	5.164 (2.729 - 9.771)	0.000	4.299 (2.249 - 8.219)	0.000	4.270 (2.228 - 8.185)	0.000
65 +			7.087 (4.054 - 12.39)	0.000	7.626 (4.358 - 13.34)	0.000	4.113 (2.201 - 7.685)	0.000	3.707 (1.966 - 6.991)	0.000	3.721 (1.970 - 7.029)	0.000
Sex												
Female (reference)			1		1		1		1		1	
Male			1.560 (1.210 - 2.010)	0.001	1.566 (1.216 - 2.017)	0.001	1.245 (0.885 - 1.753)	0.209	1.290 (0.907 - 1.847)	0.163	1.204 (0.788 - 1.840)	0.390
Marital status												
Never married (reference)			1		1		1		1		1	
Currently married			1.564 (1.017 - 2.405)	0.042	1.488 (0.976 - 2.268)	0.064	1.245 (0.812 - 1.910)	0.314	1.164 (0.755 - 1.795)	0.492	1.160 (0.752 - 1.790)	0.501
Divorced, Separated, and widowed			2.142 (1.234 - 3.716)	0.007	1.963 (1.130 - 3.410)	0.017	1.784 (1.026 - 3.101)	0.040	1.686 (0.961 - 2.947)	0.069	1.683 (0.954 - 2.968)	0.072
Decline to respond or unknown			1.340 (0.233 - 7.710)	0.743	1.022 (0.178 - 5.868)	0.981	0.883 (0.139 - 5.591)	0.895	0.824 (0.135 - 5.042)	0.834	0.820 (0.131 - 5.114)	0.832
Model 1: unadjusted; Model 2: adjusted by demographic factors (age and sex and marital status); Model 3: includes also administrative regions; Model 4: includes also socio-economic factors (employment status and monthly household income); Model 5: includes also Lifestyle factors (smoking, physical activity and BMI); Model 6: includes also education level and sex interaction term; P: p-value (significant in BOLD); OR: Odds Ratios; CI: Confidence interval.												

Table 5.6 Generalised binary logistic regression odds ratios and 95% confidence interval of the prevalence of T2DM (Household weighted models) (cont.)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Administrative Regions												
Western region (reference)					1		1		1		1	
Central region					1.889 (1.387 - 2.572)	0.000	1.767 (1.302 - 2.397)	0.000	1.727 (1.270 - 2.349)	0.000	1.731 (1.270 - 2.359)	0.001
Eastern region					1.394 (0.773 - 2.512)	0.270	1.344 (0.734 - 2.460)	0.338	1.317 (0.717 - 2.419)	0.375	1.311 (0.710 - 2.421)	0.386
Northern region					1.443 (0.998 - 2.089)	0.051	1.470 (1.006 - 2.147)	0.046	1.523 (1.040 - 2.231)	0.031	1.512 (1.033 - 2.215)	0.034
Southern region					2.047 (1.515 - 2.766)	0.000	1.996 (1.469 - 2.712)	0.000	2.060 (1.507 - 2.814)	0.000	2.050 (1.500 - 2.801)	0.000
Employment status												
Employee (reference)							1		1		1	
Unemployed							0.820 (0.537 - 1.252)	0.358	0.859 (0.559 - 1.320)	0.487	0.872 (0.566 - 1.343)	0.534
Student							0.389 (0.186 - 0.812)	0.012	0.443 (0.207 - 0.946)	0.035	0.440 (0.206 - 0.941)	0.034
Homeworker							0.663 (0.432 - 1.018)	0.060	0.688 (0.447 - 1.060)	0.090	0.706 (0.447 - 1.116)	0.136
Retired							1.530 (0.992 - 2.358)	0.054	1.492 (0.964 - 2.311)	0.073	1.492 (0.962 - 2.314)	0.074
Self-employed							1.055 (0.602 - 1.849)	0.852	1.027 (0.588 - 1.793)	0.926	1.043 (0.597 - 1.819)	0.883
Other							0.774 (0.120 - 5.003)	0.788	0.872 (0.136 - 5.586)	0.885	0.877 (0.140 - 5.490)	0.889
Monthly household income												
Highest (reference)							1		1		1	
Lowest							0.758 (0.466 - 1.235)	0.266	0.790 (0.485 - 1.288)	0.345	0.788 (0.484 - 1.281)	0.336
Low							0.886 (0.525 - 1.494)	0.649	0.889 (0.525 - 1.504)	0.661	0.888 (0.526 - 1.500)	0.657
Middle							0.852 (0.512 - 1.419)	0.539	0.849 (0.507 - 1.421)	0.534	0.856 (0.513 - 1.429)	0.552
High							0.940 (0.528 - 1.674)	0.833	0.952 (0.530 - 1.710)	0.870	0.972 (0.542 - 1.744)	0.924
Missing							0.985 (0.590 - 1.645)	0.954	0.979 (0.584 - 1.641)	0.937	0.979 (0.586 - 1.637)	0.936
Model 1: unadjusted; Model 2: adjusted by demographic factors (age and sex and marital status); Model 3: includes also administrative regions; Model 4: includes also socio-economic factors (employment status and monthly household income); Model 5: includes also Lifestyle factors (smoking, physical activity and BMI); Model 6: includes also education level and sex interaction term; P: p-value (significant in BOLD); OR: Odds Ratios; CI: Confidence interval.												

Table 5.6 Generalised binary logistic regression odds ratios and 95% confidence interval of the prevalence of T2DM (Household weighted models) (cont.)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Body mass index												
Normal (reference)									1		1	
Underweight									0.090 (0.029 - 0.270)	0.000	0.089 (0.029 - 0.270)	0.000
Overweight									1.352 (0.958 - 1.909)	0.086	1.351 (0.957 - 1.910)	0.087
Obese									1.594 (1.151 - 2.207)	0.005	1.594 (1.150 - 2.209)	0.005
Missing									1.141 (0.461 - 2.823)	0.775	1.128 (0.457 - 2.788)	0.793
Smoking status												
Never smoked (reference)									1		1	
Ex- smoker									1.196 (0.722 - 1.980)	0.488	1.184 (0.714 - 1.962)	0.513
Current smoker									1.104 (0.737 - 1.652)	0.632	1.105 (0.739 - 1.653)	0.627
Missing									1.751 (0.434 - 7.066)	0.431	1.748 (0.440 - 6.940)	0.427
Physical activity												
High (reference)									1		1	
Moderate									1.176 (0.516 – 2.681)	0.700	1.177 (0.517 – 2.679)	0.697
Low									1.113 (0.511 – 2.424)	0.788	1.129 (0.520 – 2.454)	0.758
None									1.421 (0.760 – 2.658)	0.271	1.436 (0.769 – 2.681)	0.257
Missing									1.240 (0.573 – 2.684)	0.585	1.253 (0.581 – 2.702)	0.566
sex × Education												
Male × low education											1.009 (0.550 - 1.850)	0.992
Male × Middle education											1.311 (0.695 - 2.472)	0.403

5.4.3 Projection of the prevalence of T2DM

In this section, I describe my projection of T2DM prevalence in the Saudi population based on the two main methods highlighted in the method section above (using the UN projection and the Cohort component projection). Overall, one might expect diabetes prevalence to increase along with age. However, in all my projection data (**Figures 5.1, 5.2, 5.3**) I found a significant drop in the absolute number of diabetics in the 60 – 64 age group, which then appeared to increase dramatically in the 65 years old and above group. This apparent anomaly can be accounted for by a combination of two factors. First, the drop in patient number in the 60-65 age group is due to a drop in the underlying prevalence of T2D in this group, suggesting that in Saudi Arabia, diabetes prevalence in fact peaks before 60 years of age (**Figure 5.4**). This might reflect an underlying demographic shift in for example lifestyle and dietary habits. Secondly, the jump in absolute number of cases seen in the above 65 group is simply a reflection of the large number of people in this composite age group, rather than a change in prevalence.

5.4.3.1 United Nations population projections

As mentioned above, I calculated the average difference between 2016/17 and 2017/18 to calculate the age-sex group percentages for the projected years. For example, for Saudi males aged 25 - 29 the average increase per year is 0.53. I applied this increase to estimate the percentage for this age group in 2020, 2025 and 2030. For this age group, the percentage of Saudis will be 50.25 in 2030. This is given by:

$$\text{percentage of Saudis male aged (25 – 29) in 2018} - 12 \times \text{average growth of Saudis males aged (25 – 29)} = 56.56 - 12 \times 0.53 = 50.25.$$

Based on the UN projection of the number of total Saudi population aged 25 - 29 and previous percentage, I calculated the number of Saudi male's population only aged 25 – 29 in 2030. Hence, the number of Saudi diabetics aged 25 – 29 in 2030:

$$\begin{aligned} \text{Diabetes prevalence of Saudi population aged (25 – 29) in 2013} &\times \frac{\text{total number of Saudis male aged (25 – 29) in 2030}}{100} \\ &= 8.47 \times \frac{794,972}{100} = 67334.13 \end{aligned}$$

$$\text{Total percentage of diabetes} = \frac{\text{total number of diabetic patients aged 25+ in 2030}}{\text{total number of Saudi population aged 25+ in 2030}} \times 100 = \frac{1,396,120}{5,994,983} \times 100 = 23.29\%$$

Such a linear function for estimating population growth cannot be fully accurate, however, the trend of Saudi national population appeared slow based on the available numbers. **Figure 5.1** displays the trends and future projections for T2DM prevalence by different age groups among males in SA over the period 2015 – 2030. The overall number of male diabetic cases is projected to increase from 848,504 in 2015 to 1,396,120 by 2030. The prevalence of T2DM is greatest amongst those aged 55 - 59 and 65+ in all four timeframes.

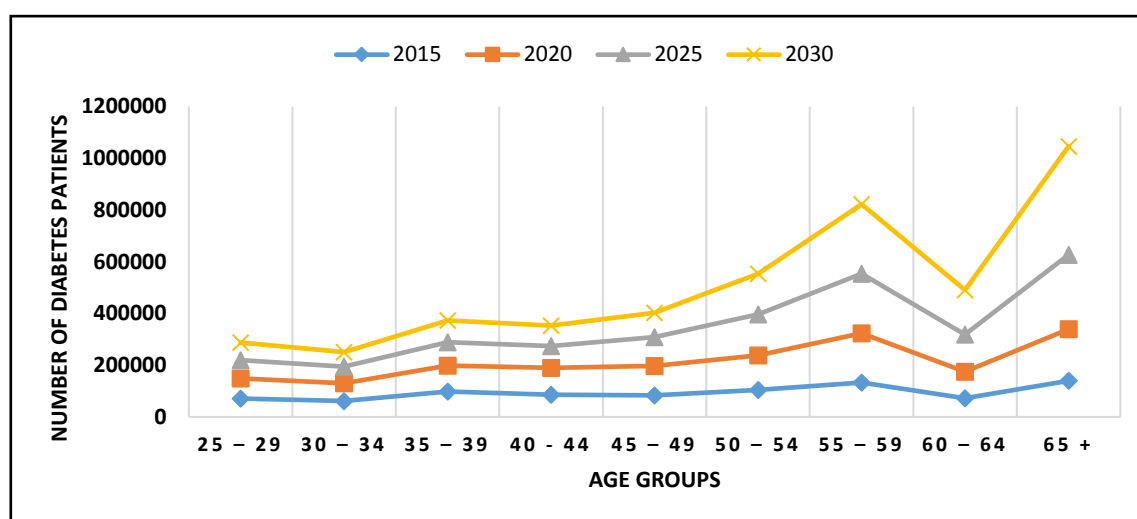


Figure 5.1 Projected number of people with T2DM for selected years among Saudi male adults 25 years old age and over in Saudi Arabia

Table 5.7 presents the projection for the overall prevalence of T2DM and the number of Saudi diabetic males from 2015 to 2030. The projection shows that over this period, the prevalence of T2DM is projected to increase rapidly in all age groups. The overall male prevalence of T2DM increases from around 18% in 2013 to 23.3% by 2030.

Table 5.7 Total number (cases) of Saudi diabetic males

Age	2013 (%)	2015	2020	2025	2030
25 – 29	8.47	71488	77881	70304	67334
30 – 34	7.5	61591	68282	64432	56276
35 – 39	14.15	98319	100163	90183	83907
40 – 44	15.12	85934	103985	84185	78364
45 – 49	17.95	82877	113754	111100	93752
50 – 54	27.21	103867	133591	159186	156950
55 – 59	45.63	132183	191041	230799	267746
60 – 64	32.95	72658	102367	143248	172676
65 +	37.14	139586	199476	287887	419116
Total number	N	848504	1090540	1242222	1396120
Total prevalence (%)	18	18.22	19.42	21.39	23.29

Table 5.8 and **Figure 5.2** presents the same analysis for Saudi females. Based on the current population size and prevalence, I estimate a total of 752,703 Saudi Females had T2DM in 2015. The projection suggests that this will increase almost two-fold to 1,340,408 by 2030. Stratification by age reveals a more complex picture, with peaks of prevalence in the 50-54 (258,336) and over 65 (284,423) age groups. Overall, males had higher projected prevalence rates than females. The greatest increase in the number of T2DM cases for both sexes was in those aged over 65, reaching 419,116 for men and 284,423 for women by 2030. In contrast, the lowest number of diabetic cases was in the 25-29 age group (67,334 for men and 42,865 for women by 2030).

Table 5.8 Total number of Saudi diabetic females

Age	2013 (%)	2015	2020	2025	2030
25 – 29	5.94	51,010	54,606	45,274	42,865
30 – 34	8.53	71,976	78,342	75,849	62,578
35 – 39	11	79,793	79,673	83,054	79,271
40 – 44	9.7	54,589	65,698	62,747	64,930
45 – 49	20.59	95,165	134,194	155,380	151,050
50 – 54	28.48	116,748	152,489	216,082	258,336
55 – 59	34.02	91,378	133,539	177,349	255,363
60 – 64	28.57	60,859	73,096	106,290	141,591
65 +	33.14	131,185	167,708	210,186	284,423
Total number	N	752,703	939,346	1,132,211	1,340,408
Total prevalence (%)	14.44	15.88	16.83	18.58	20.34

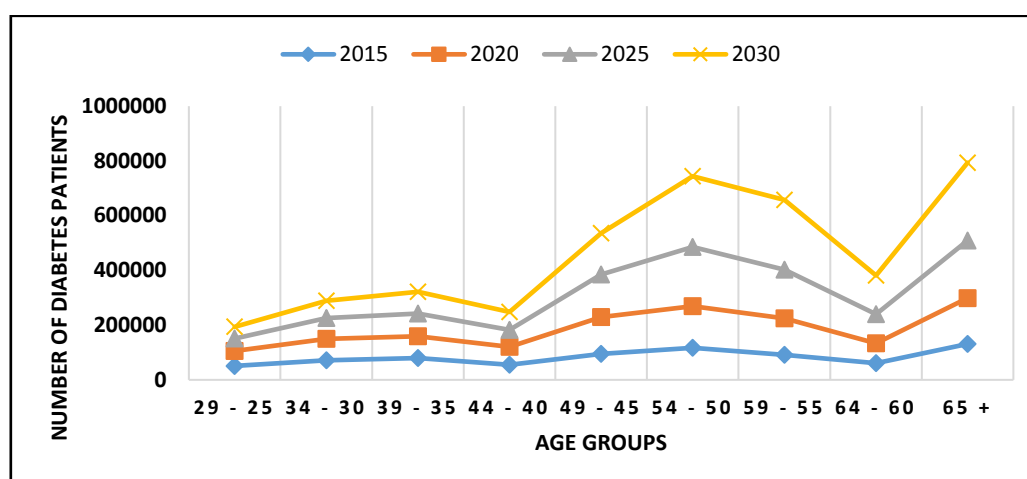


Figure 5.2 Trends and future projections of diabetic patient specific number among female adults in Saudi Arabia 2015 - 2030

Table 5.9 and **Figure 5.3** illustrates the predicted overall T2DM prevalence between 2015 and 2030 among Saudis aged 25 years and above. Reflecting the structural change in the population demography, the prevalence of T2DM is predicted to increase from 16.1% in 2013 to 21.8% in 2030. The rise in T2DM cases over this period is most prominent among in the 55 – 59 and 65+ age groups.

Table 5.9 Total number of Saudi diabetics

Age	2013	2015	2020	2025	2030
25 – 29	6.88	117,150	126,509	109,544	104,343
30 – 34	8.07	134,367	147,589	141,088	119,756
35 – 39	12.3	174,687	176,157	171,262	161,576
40 - 44	12.22	138,224	166,806	147,087	145,133
45 – 49	19.44	179,607	249,896	267,999	244,148
50 – 54	27.88	220,713	286,157	374,635	413,708
55 – 59	40	223,314	324,482	410,845	534,961
60 – 64	30.91	134,004	175,112	249,375	315,173
65 +	35.63	274,953	371,674	502,160	707,869
Total number		1,597,017	2,024,383	2,373,993	2,746,666
Prevalence (%)	16.13	16.99	18.08	19.94	21.82

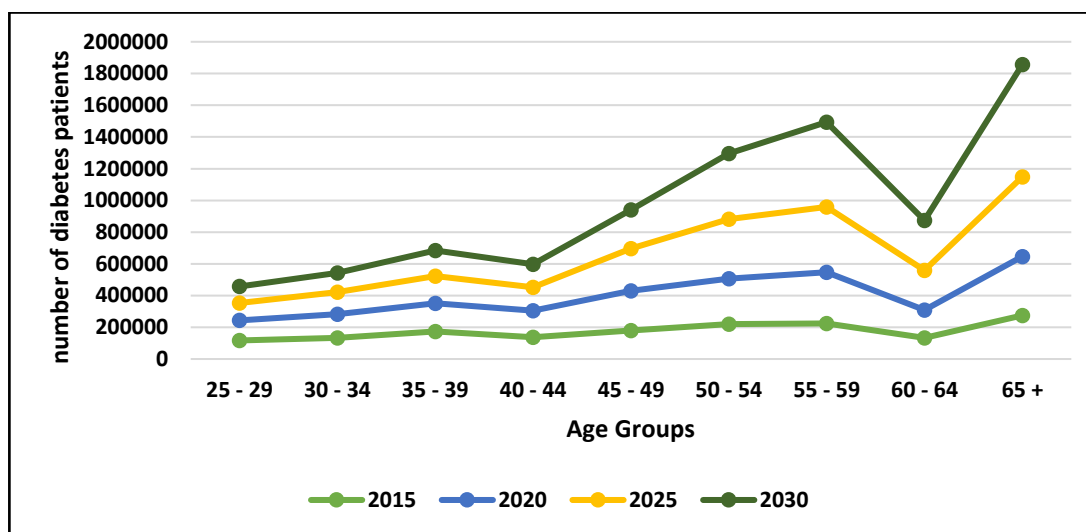


Figure 5.3 Projected total number of people with T2DM for selected years among adults 25 years old age and over in Saudi Arabia

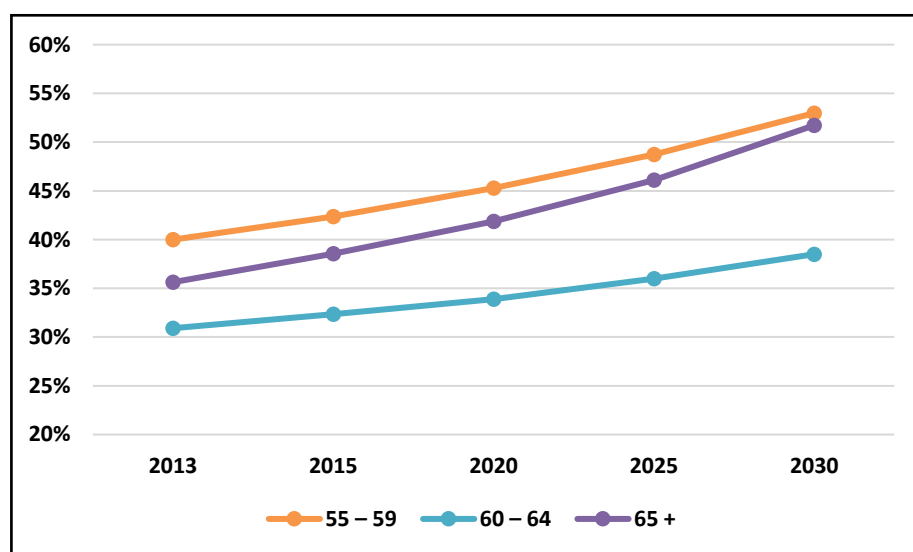


Figure 5.4 Prevalence of T2DM for selected years in three different age groups in Saudi Arabia

5.4.3.2 Cohort Component projections

Table 5.10 and **Table 5.11** presents the trends and future projections for age-specific prevalence of T2DM from 2015 – 2028. Over this period, the prevalence of T2DM is projected to grow rapidly in all age groups, with overall prevalence increasing from around 18.3% in 2015 to 19.8% by 2028 in males and from 15.9% to 16.9% in females. The highest number of diabetics is seen in men aged 55 - 59 (244,502) and for women aged 65 or over (206,484). In contrast, the smallest increase was in men aged 25 - 29 years (to 77,514) and for women in the same age group (to 52,883) in 2030. There is an increase in T2DM prevalence for both sexes over the period. Males had higher projected T2DM prevalence than females, particularly in the 55- 59 age group. **Figures 5.5** and **5.6** show these trends graphically, revealing similar trends for males and females in the highest age groups. However, males showed the peak incidence of T2DM in the 55 – 59 age group; in females the peak was in the 65+ age group. The lowest prevalence was observed amongst the 25 – 29 age group for both sexes in all three timeframes.

Table 5.10 Number of male diabetic patients in two timeframes 2023 and 2028

Age groups	2013	2018	2023	2028
25 – 29	8.47	84,363	90,337	77,514
30 – 34	7.5	67,259	74,456	79,729
35 – 39	14.15	111,329	126,399	139,923
40 - 44	15.12	100,675	118,152	134,145
45 – 49	17.95	100,437	118,306	138,843
50 – 54	27.21	121,430	149,785	176,433
55 – 59	45.63	158,829	198,217	244,502
60 – 64	32.95	83,086	108,039	134,831
65 +	37.14	158,050	183,536	222,629
Total number	N	985,459	1,167,226	1,348,548
Prevalence (%)	18	18.33	18.83	19.77

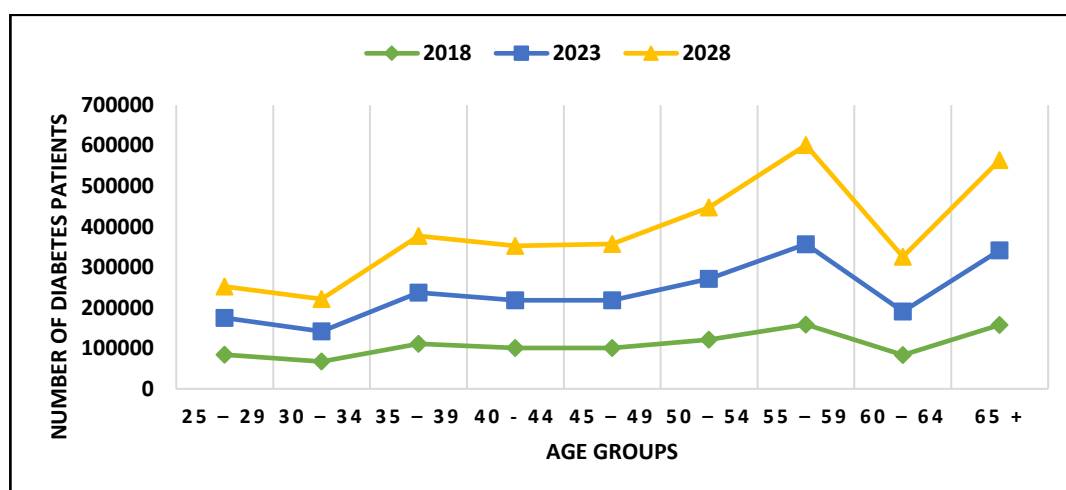


Figure 5.5 Future projection of (T2DM) patient number among Saudi males in three different timeframes; 2018, 2023, and 2028

Table 5.11 Number of female diabetic patients in three timeframes 2018, 2023 and 2028

Age groups	2013 (%)	2018	2023	2028
25 – 29	5.94	57,930	58,074	52,883
30 – 34	8.53	75,231	82,982	83,189
35 – 39	11	84,548	96,662	106,621
40 - 44	9.7	62,201	74,086	74,087
45 – 49	20.59	109,034	130,834	130,834
50 – 54	28.48	120,165	148,672	178,397
55 – 59	34.02	108,767	140,259	173,533
60 – 64	28.57	67,406	87,392	112,695
65 +	33.14	147,675	171,139	206,483
Total number	N	832,957	990,101	1,118,722
Prevalence (%)	14.44	15.96	16.54	16.90

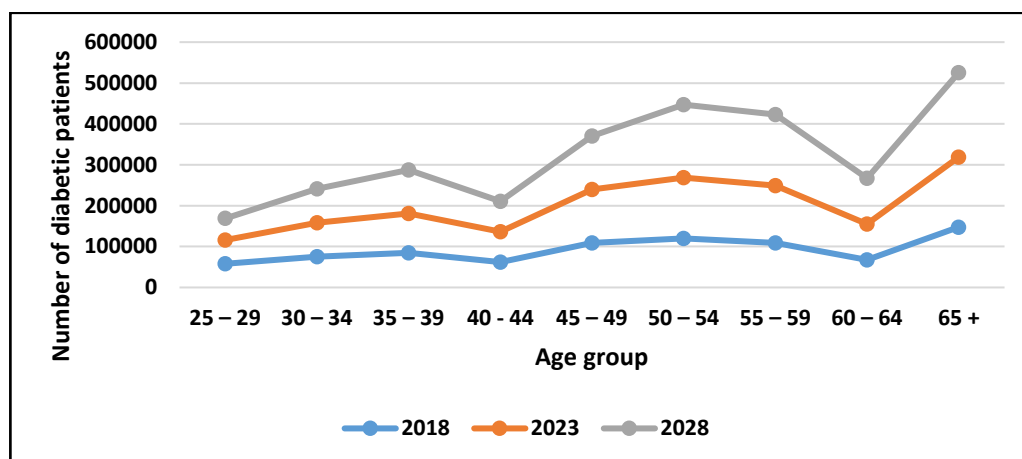


Figure 5.6 Trends and future projections of the diabetic number among female adult in Saudi Arabia 2018 - 2028

Table 5.12 and **Figure 5.7** illustrate the predicted overall T2DM prevalence rate between 2018 and 2028 among Saudis aged 25 years and over. The prevalence of T2DM is predicted to increase from 16.1% in 2013 to 18.6% in 2028. This is most marked in the oldest age groups, where diabetes is estimated to increase 3-fold in the 55-59 age group and among those aged over 64 between 2018 and 2028. In contrast, the lowest growth in prevalence is seen in the 25 - 29 and 30 – 34 age groups. Although the two T2DM prevalence models were based on different projections of population size, both predict a significant increase in the prevalence of the disease over the next twelve years. However, the absolute numbers of diabetic patients were not consistent between two methods due to the diversity of the underlying data and population projection method, there was surprisingly good agreement in the age groups showing highest prevalence of disease.

Table 5.12 Number of diabetic patients in two timeframes 2023 and 2028

Age	2013 (%)	2018	2023	2028
25 – 29	6.88	135,623	140,643	124,215
30 – 34	8.07	143,545	158,622	164,490
35 – 39	12.3	191,313	217,959	240,851
40 – 44	12.22	159,725	188,824	215,122
45 – 49	19.44	211,718	251,653	297,499
50 – 54	27.88	242,053	299,013	355,416
55 – 59	40	267,118	338,673	418,371
60 – 64	30.91	150,868	195,900	248,408
65 +	35.63	310,395	360,072	435,576
Total number	N	1,812,363	2,151,359	2,499,949
Prevalence (%)	16.13	17.10	17.66	18.60

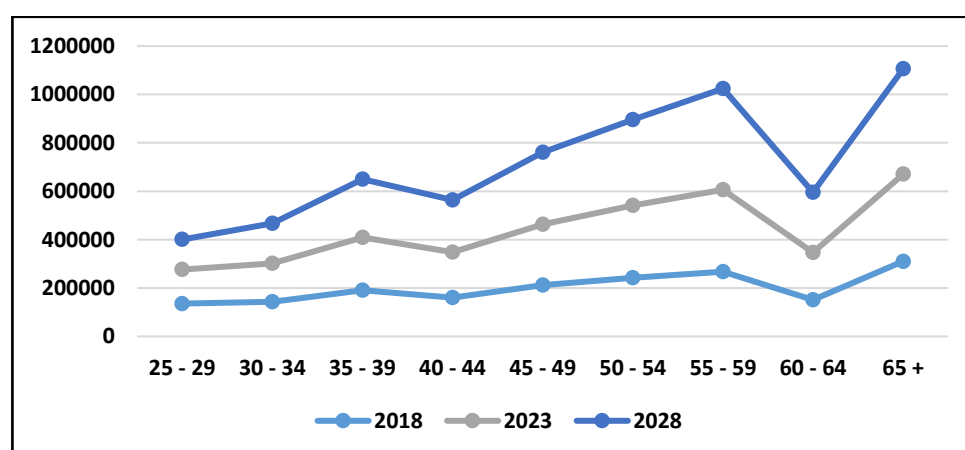


Figure 5.7 Projected number of people with T2DM for selected years among adults 25 years old age and over in Saudi Arabia

5.5 Discussion and conclusion

Addressing the association between education level and diabetes outcomes is important to understand better the high prevalence of diabetes in Saudi Arabia. In this chapter my first objective was to assess the estimates of diabetes prevalence using both weighted and unweighted survey samples. I then used binary logistic regression models to assess the relationship between education level and the prevalence of T2DM.

Finally, I applied the models I developed to estimate future trends of T2DM in the KSA. The results shed light on the importance of education level in relation to diabetes in Saudi Arabia and will contribute to meeting the huge challenges of reducing the high burden of diabetes prevalence. Overall, the evidence confirmed my hypothesis that there is a relationship between education level and the prevalence of T2DM. The results also show that the prevalence of diabetes will continue to increase until at least 2030 for both sexes unless there is significant policy intervention.

5.5.1 Summary of the key findings

I found that the choice of weighting factor significantly influenced the estimate of diabetes prevalence. Thus, estimates of the prevalence of diabetes were 16.5% (unweighted), 12.6% (household weight), 11.2% (lab weight), and 11.5% using the new weight. Using the new weight gives an estimate of disease prevalence that is 3% or 1% greater than using the lab weight or the household weight, respectively. I also observed an inverse association between education level and T2DM. This association remained significant after adjusting the model for other factors (e.g. demographic and regional). The prevalence of T2DM is greater among males, older age groups, the retired and the previously married. Prevalence is also high among ex-smokers, the physically inactive and the obese. An analysis of the effect other risk factors provides convincing evidence that regional, socio-economic and lifestyle factors lie on the causal pathway for developing diabetes, especially among low-educated individuals. Overall, the evidence confirmed my research hypothesis that the prevalence of diabetes is higher in low-educated individuals.

5.5.2 The association between education level and the prevalence of T2DM

With regard to education level, my findings indicated that education level is a significant predictor for having diabetes. This is in line with findings of previous studies showing an association between education level and the prevalence of T2DM (Corsi and Subramanian, 2012, Costanian et al., 2014, Al-Moosa et al., 2006, Bener et al., 2009, Hwang and Shon, 2014). Additionally, low-educated individuals have a higher risk of being diabetic compared to high-educated individuals (Al-Moosa et al., 2006, Murad et al., 2014, Abdella et al., 1998, Kim et al., 2015). A possible explanation of these observations is that better-educated individuals are more likely to be aware of the risks and symptoms of diabetes (Su et al., 2016). This might be compounded if low-educated individuals were predisposed to diabetes because of life-style factors such as diet and exercise.

In other words, better educated individuals are less likely to be affected by diabetes because they are more likely to lead healthier life-styles than less well-educated individuals (Silles, 2009). In order to explore this relationship, I examined the sensitivity of the association between education level and the prevalence of diabetes to correction for other factors. I found that the association remains significant after adjusting the model for demographic factors, administrative region, and marital status. However, this analysis did reveal additional factors that are associated with diabetes prevalence.

With regard to demographic factors, the results indicate that middle and older aged groups were more likely to be diabetic than younger aged groups. This result is consistent with previous investigations of the relationship between socio-economic factors and the prevalence of T2DM indicating that middle-aged groups are at higher risk of being diabetic (Connolly et al., 2000, Sicree and Shaw, 2007). In addition, I found that male participants had a higher likelihood of being diabetic. This finding is consistent with findings of past studies in Saudi Arabia (El Hazmi, 1998, Warsy and El Hazmi, 1999), and in the USA (Connolly et al., 2000).

In considering **marital status**, I found that belonging to the 'currently married' or 'not currently married' (as opposed to never married) groups showed a positive association with T2DM. However consistent with previous studies showing an association between marital status and the prevalence of diabetes, these are limited to a single study highlighting this relationship (Murad et al., 2014), and one other study that used marital status as a covariate (Corsi and Subramanian, 2012). One additional study reported that individuals who are married or divorced have a higher risk of being diabetic compared to individuals who never married (Al-Moosa et al., 2006). The result is also consistent with a UK study showing that never or not currently married individuals were more likely to have a poor health status than currently married individuals (Robards et al., 2012). A possible explanation of this complex interaction between marital status, other socio-economic factors and diabetes is that in Saudi Arabia, the social changes associated with marriage have an impact on living arrangements and overall health status that are more pronounced than in other societies. For example, in the Saudi and Gulf region social interaction greatly increases after marriage due to visiting family and relatives. This is often associated with extensive communal eating and potentially excessive food intake compared to individuals who have never married.

With regard to **administrative region**, individuals who live in the Southern region are more likely to be diabetic compared to those who live in the Western region, consistent with previous studies showing that regional differences and urbanisation level could influence prevalence of T2DM (Al-Nuaim, 1997, Al-Nozha et al., 2004, Al-Rubeaan et al., 2015). This effect could be related to differences in quality of and access to health care services and public sporting facilities between regions and provinces. For example, most of the provinces in the Western and Eastern regions have good health care services and other public facilities in marked contrast to the Southern region which is predominantly semi-urban and rural.

This difference in socio-economic level across regions was highlighted in a recent Saudi study showing that this diversity influenced health outcomes (AlOmar et al., 2018). Significantly, obesity rates vary between different regions (Alqahtani et al., 2014), and it is well-known that obesity is a key risk factor for T2DM. The possible implication of this finding is that it is demographic factors, marital status, and region that are more likely to explain the association between education and the prevalence of T2DM. In fact, these factors were more often used as confounders between education and diabetes mellitus. Another possible explanation of this finding is that these factors eliminate the effect of education on diabetes. This is probably due to the age profile of individuals who are not educated, with education level generally higher in the younger age groups.

With regard to the effect of **socio-economic factors**, the findings also confirm that low-educated individuals are also more likely to have a high risk of diabetes compared to high-educated individuals. The results showed a positive association between T2DM prevalence and **employment status**, with individuals in the 'unemployed', 'homeworker' and 'retired' groups showing significantly greater prevalence of T2DM. In the logistic model, only individuals in the 'student' group showed significant negative association with prevalence of T2DM. Retired individuals also showed an association, but this relationship was not significant. These results are inconsistent with a previous study that provided evidence that the unemployed and women who stayed at home are more likely to be diabetic compared to employed individuals (Murad et al., 2014). The most likely explanation for the lower risk of T2DM among students is that these individuals are predominantly in the youngest age group.

Given the strong relationship between age and socio-economic factors, adjusting for age might explain the loss of the relationship between some employment status groups and the prevalence of T2DM. Even after correcting for the influence of lifestyle factors on the prevalence of diabetes, the findings indicate that low-educated individuals still had higher odds of being diabetic compared to better-educated individuals. The results also showed there was no association between **monthly household income** and the prevalence of T2DM before and after adjusting the sample. This finding is at odds with past studies that suggesting an inverse association between income status and the prevalence of diabetes (Murad et al., 2014, Whiting et al., 2011, Rabi et al., 2006). One possible explanation for this discrepancy might be that the strong relationship between administrative region and monthly household income could influence the relationship between education level and the prevalence of T2DM.

Considering the impact of **lifestyle factors** on the prevalence of diabetes, the findings indicate that the relationship between lifestyle factors and subsequent development of T2DM is important in relation to our finding, as lifestyle factors and obesity are both strongly associated with low education in Saudi Arabia. It remains to be proved whether the association between education level and T2DM is mediated primarily via nutritional effects. However, my results are in line with previous studies that lifestyle factors are related to both education level and the prevalence of diabetes (Sicree and Shaw, 2007, Guariguata et al., 2014, Connolly et al., 2000, Hussain et al., 2007, Lidfeldt et al., 2007, Tamayo et al., 2010, Kumari et al., 2004, Espelt et al., 2011). This is particularly true in the GCC region and Saudi Arabia, where individuals tend to be less physically active (Al-Lawati and Jousilahti, 2004, Mehio Sibai et al., 2010), rely on a high carbohydrate, high fat diet, and as a result exhibit a high prevalence of obesity, all factors known to increase the risk of developing T2DM.

5.5.3 Projection of T2DM prevalence

In the second analysis, I modelled projections of T2DM prevalence over the next 8-10 years without considering the impact of other risk factors on this outcome. For this I used the United Nations projections and cohort component projection methods to predict the underlying growth in the Saudi population. Based on the UN population projections, my results predict an increase of diabetes prevalence from 16.9% to 21.8%, and in the number of individuals with diabetes from 1,597,017 to 2,746,666 between 2015 and 2030 among the Saudi population aged 25 years and older. Using the mortality-component projection method, I estimate the prevalence of diabetes will increase from 17.1% in 2018 to 18.6% in 2028 for Saudi adults aged over 25 years. Both methods suggest that the prevalence of diabetes will increase dramatically in older age groups over the period of the projection.

- **The future projection of T2DM and its risk factors**

The projected T2DM burden in the Saudi population is based on nationally representative population data. The analysis of prevalence highlights the urgent need for T2DM public health policy and intervention programs at the national level, and the identification of modifiable risk factors will help inform the development of appropriate strategies. As mentioned above, the model predicts a significant increase in the prevalence of T2DM over the next 5-10 years with age, education level, region, and obesity being the most significant risk factors. Another major driver of the increase in T2DM prevalence is demographic change resulting in a rapidly ageing population and other age-associated risk factors.

A further factor is the unique cultural context of the GCC countries that contributes to a lack of physical activity in the Saudi population, especially for women (Alshaikh et al., 2017). Overall, my results suggest that an intervention program to reduce the incidence of obesity has the greatest potential impact on the prevalence of T2DM. Additionally, the Saudi context requires further research to address the feasibility of intervention based on modifiable risk factors such as physical activity. Evidence from the Finnish Diabetes Prevention study underlined the importance of lifestyle interventions in preventing the progression of pre-diabetes, thereby mitigating increases in T2DM prevalence (Lindström et al., 2006). This study demonstrated that such preventive measures could lead to a 40% decrease in T2DM incidence. However, such approaches that are focused on behavioural modification in individuals, should probably be complemented by more population-based approaches such as taxation and marketing restrictions.

- **Comparison of the diabetes future projection and other diabetes estimates**

The International Diabetes Federation (IDF) projections predict an increase in diabetes prevalence from 16.2% to 20.8% among Saudi adults 20-79 years old between 2011 and 2030 (Whiting et al., 2011). Our estimates are somewhat lower than those reported by the IDF. This discrepancy presumably results from diversity in the underlying dataset, differences in the analytical methods, and definition of affected status (as their estimates include T1DM). The IDF estimates were based on combining data and surveys from several countries, taking into account geography, IDF region, and World Bank's income classification group. Our projections suggest that the prevalence and number of people with diabetes in Saudi Arabia will increase more than for other countries in MENA region (Shaw et al., 2010).

This observation is consistent with IDF estimates that the increase in the prevalence of diabetes in SA will place it within the top ten countries in 2030, along with other GCC region countries such as Kuwait (15.9% to 22.2%), Qatar (14.1% to 20.9%), Bahrain (15.3% to 24.6%), and the UAE (12.6% to 23.8%) from 2011 to 2030 (Whiting et al., 2011). Our projections are considerably higher than the IDF projections for other developed countries such as the US (from 10.9% to 11.8%), the UK (6.8% to 7.5%), and the Netherlands (7.3% to 8.8%) over the same period. However, our estimates are less than one recent projection for Saudi Arabia (Al-Quwaidhi et al., 2014b), which estimated the prevalence of diabetes in SA in adults aged 25 years and over at 29.2% in 2011, and projected growth to 44.1% by 2022. Again, these differences are likely due to methodological differences and the underlying survey dataset.

Interestingly, our findings predict that the greatest increase in diabetes prevalence and the number of affected individuals will be in older age groups (aged 50 years and older irrespective of gender). This is inconsistent with the Syrian diabetes projection study, in which the greatest increase in prevalence was projected for the younger age groups, presumably due to significant differences in demographics and income level. Apart from the Syrian study, my results are consistent with most global diabetes estimates in high-income countries, where the greater growth in T2DM is expected to be in those aged over 60. The picture is different again in low and middle-income countries, where the greatest increase in T2DM prevalence is expected in those of working age (40 – 60) (Whiting et al., 2011, Guariguata et al., 2011).

Despite these findings, my results show that the prevalence of T2DM amongst the middle-aged groups is also projected to grow significantly. This reflects diversity in socio-economic level within the Saudi population and highlights the sensitivity of these estimates to the rapid development of socio-economic status in the country expected in the next decade. Our predictions for these groups are lower than one analysis which projected that T2DM in Saudi Arabia will increase from 20.2% in 2013 to 27.1% in 2035 (Guariguata et al., 2014). Rapid urbanisation and rates and the reducing prevalence of infectious diseases are additional factors acting to increase the prevalence of diabetes in the country. Most global estimates of T2DM prevalence aggregate the total number of diabetes cases and apply logistic regression modelling to generate age-specific estimates using the United Nations population projections, taking into account ethnicity, economic development and geography (Guariguata et al., 2014). Overall, our finding that the prevalence of T2DM amongst the over 65s will be three times higher in high-income countries in comparison to low-income countries was in line with the most recent global estimates (Cho et al., 2018).

In conclusion, the results described in this chapter confirm the significant relationship between education level and diabetes outcomes. They also confirm that other socio-economic and lifestyle factors play a critical role in the pathway connecting education level and the prevalence of T2DM. The increasing prevalence of diabetes and the growing socio-economic gap is an essential issue in Saudi Arabia. My findings suggest that future studies should continue to focus on the role of socio-economic factors on the incidence of T2DM among Saudi individuals. This study provides Saudi policy makers with some reliable benchmarks and a better understanding of the diabetes problem to help target planned interventions to reduce the high prevalence of diabetes in Saudi society.

Chapter 6 Diagnosed and Undiagnosed Type Two

Diabetes in Saudi Arabia: Exploring the Socio-economic Factors

6.1 Introduction

Previous studies about the relationship between socio-economic predictors and diagnosed diabetes have been discussed in Chapter 2. These studies indicated that there is a relationship between socio-economic factors and being diagnosed with diabetes. Besides socio-economic predictors, this relationship could also be related to other factors such as health care access. In addition, previous literature has highlighted that some individuals have diabetes without realising it. Undiagnosed diabetes is one of the main causes for the early onset of diabetes complications. Undiagnosed diabetes has become a public health issue and reducing the burden of T2DM is now a high priority in Saudi Arabia. Long-term undiagnosed diabetes is associated with a high risk of serious complications that could lead to early death.

Globally, it has been estimated that 45.8% of the adult population who have diabetes are undiagnosed (Guariguata et al., 2014). The early identification of individuals with undiagnosed diabetes or those who have high-risk of developing diabetes (pre-diabetes) is therefore potentially an effective approach to tackle the burden of this disease and its complications (Wilder et al., 2005, Clark and Utz, 2014). Together, these considerations point to the potential of estimating undiagnosed diabetes in driving policies that could reduce the high prevalence of diabetes. T2DM is usually asymptomatic initially, however insulin resistance has important long-term health complications. For example, chronic high blood sugar damages blood vessels which can lead in turn to damage to a range of organs including eyes, kidneys, nerves, liver and major blood vessels to heart and brain. These complications are often already present in patients with late diagnosis of T2DM. We therefore need to understand better the scale of and reasons underlying undiagnosed T2DM.

This means that public health policy should have the dual objective of preventing T2DM where possible, together with the earlier identification of individuals with diabetes. The early identification of undiagnosed cases in Saudi Arabia will improve health outcomes for those with T2DM. In this context, several studies have already developed a risk score tool based on demographic and anthropometric measures. These are aimed at predicting the likelihood of an individual having undetected diabetes using different diagnosis criteria. However, these risk score models are not directly applicable to the Saudi population due to socio-economic and lifestyle differences.

This chapter is divided into four main sections: introduction, methods, results, discussion, and conclusion. In **Section 6.2**, I describe my approach to the statistical analyses of the association between education level and diagnosed diabetes described in Chapter 4. A review of the literature on the association between education level and diagnosed diabetes and the risk factors discussed in Chapter 3. This review allowed me to develop an analytical framework for the relationship between education level and diagnosed T2DM with a list of possible risk factors. The results of the analysis are divided into three subsections and will be presented in **Section 6.3**. The section presents the characteristics of study participants, and a descriptive and analytical analysis of the relationship between education level and diagnosed diabetes. This chapter will end with a discussion and summary of the conclusions in **Section 6.4**.

The aim of this chapter is to investigate the association between education level and diagnosed diabetes, addressing the following research questions:

1. What is the association between education level and the diagnosed diabetes? Is this association affected by an interaction between education level and gender?
2. Does this relationship persist after controlling for demographic, regional, socio-economic, and healthcare access factors?

My hypothesis is that individuals with lower education levels are more likely to have undiagnosed T2DM. Furthermore, I propose that demographic and health care access factors influence the relationship between socio-economic factors and diagnosis of T2DM, and that this is affected by an interaction between employment status and health insurance.

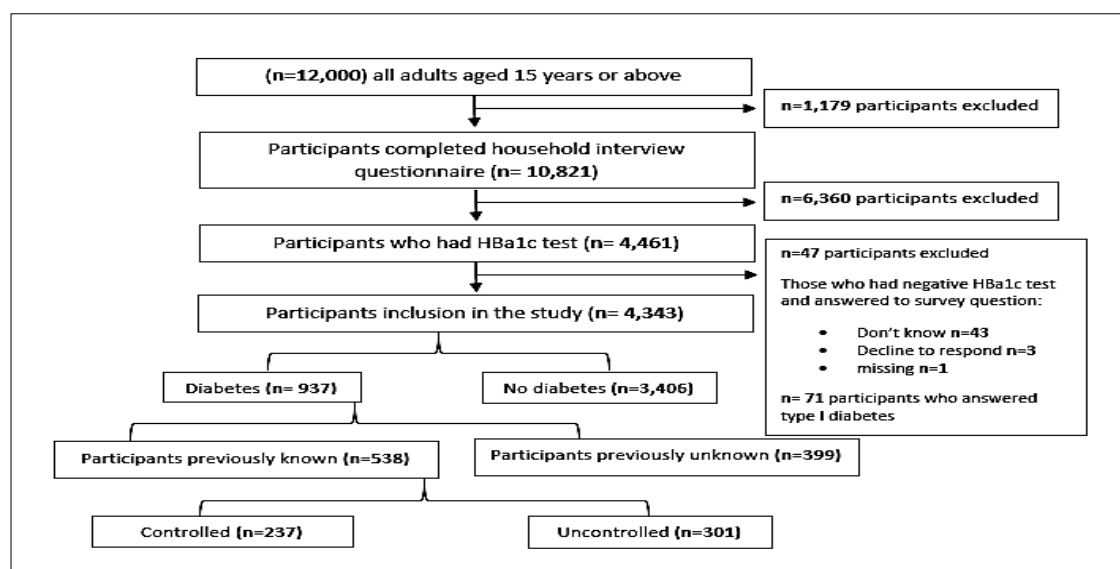


Figure 6.1 Flow chart for participants' inclusion in the study

6.2 Methods

6.2.1 Study design and participants

The study uses data collected in the second phase of the Saudi Health Interview Survey (SHIS), which included only Saudi citizens aged 15 years and above. This is a cross-sectional survey designed to investigate the prevalence of chronic diseases and associated risk factors among Saudi residents from different regions across the country. A summary of the data and the survey methods were described in Chapter 4. Participants who had T1DM (71 individuals) and who live in the Tabuk region were excluded from the study (see section 4.3 for further information on this). The survey applied the HbA_{1c} test to identify individuals with undiagnosed diabetes. A total of 4,461 participants took the HbA_{1c} test of whom 738 had diabetes. Note that of those with high HbA_{1c} , some were previously diagnosed.

6.2.2 The association between education level and diagnosis diabetes

The outcome is the result of cross tabulation between HbA1c test results: ($\text{HbA1c} < 6.5\%$ or $\text{HbA1c} \geq 6.5\%$) and the survey question: “Have you ever been diagnosed with diabetes by a doctor or health professional?” - the classification table is discussed in detailed in Chapter 4. Binary logistic regression was selected for the first analysis due to the dichotomous nature of the study outcome. The purpose of applying this model is to understand the relationship between the outcome variable and possible explanatory variables.

The analytical framework for the relationship between diagnosed diabetes and educational level based on published studies was presented in Chapter 2. The analytical framework for this chapter follows the outcome measure (individuals who are diagnosed or undiagnosed with T2DM). The analytical framework examines whether an individual’s education level influences the likelihood of diagnosis. Different demographic, community, socio-economic and healthcare access factors could influence this relationship. Then, healthcare access characteristics could also influence the likelihood of an individual being diagnosed with T2DM.

Administrative region is another important predictor of community-level based on previous literature indicating that the variation in socio-economic levels between regions could affect an individual’s likelihood of being diagnosed. For example, high-educated individuals are more likely to be aware of diabetes and therefore less likely to have undiagnosed disease. They are also more likely to visit healthcare facilities routinely. As discussed in Chapter 3, there is a free public health care services for all nationals in Saudi Arabia. However, only governmental employees have private health care access through employee health insurance, and the rest of the Saudi population pays out-of-pocket for private services (Almalki et al., 2011, Mourshed et al., 2006, Walston et al., 2008).

Health care access were explored using the following four variables. **Health care provider number** was classified as: no; yes, (one location); yes, (more than one location). This variable described the additional health care places for participants. While all individuals have free access to public primary health care services, this variable measure access to additional health care provision. I hypothesised that individuals who had access to more than one source of health care were more likely to diagnosed.

Type of health care facility last visited for medical attention (Almalki et al., 2011). This variable was categorised into three groups: public; private; and other. I hypothesised that individuals who attend governmental hospitals are less likely to be diagnosed compared to those who attend private sector hospitals. The third variable was whether the individual has had a **routine check-up** (binary classified yes or no). This was used as low-educated individuals are less likely to have a routine check-up of their health compared with high-educated groups and are therefore more likely to be unaware of their diabetes status. However, this factor could also reflect on ease of access to health care facilities (Costanian et al., 2014, Collier et al., 2015, Grintsova et al., 2014).

Finally, the **time between home and health care facility** was used as it might influence the probability of visiting health care services (Albejaidi, 2010, Mourshed et al., 2006). Thus, individuals who lived near to a healthcare facility might be more likely to undergo a routine check-up and obtain an early diagnosis of diabetes. This variable was categorized based on the SHIS database: less than 60 minutes or more than or equal to 60 minutes.

In addition, the **region of residence** could also affect individual socio-economic characteristics, which in turn might influence the likelihood of receiving a diagnosis. Region may also be associated with health care access driven by regional health care policy and by geographical location (e.g. remoteness). Each risk factor discussed in Chapter 2 was included in the analytical framework, as they are all potentially relevant to the outcome and could help in gaining a better understanding of the association between education level and diagnosed diabetes.

Before applying the logistic regression model, bivariate analysis was applied to assess the significance of the relationship between the outcome measure and explanatory variables using Chi-square tests. This identified the most significant risk factors that influence the likelihood of being diagnosed with T2DM ($p\text{-value} < 0.05$). After identifying the significant relationships, five logistic models were fitted for diagnosed diabetes, controlling for different socio-economic, demographic, healthcare access factors along with administrative region. These factors will be explained more fully in the following section. The following table presents the analytical framework and the five logistic regression models.

Table 6.1 The analytical framework for diagnosis diabetes outcome

Outcome measure	Main predictor	Model adjustment	Variables list based on SHIS database
Diagnosis diabetes	Education level	Model 1: Unadjusted	
		Model 2: Demographic factors	Age and sex, and marital status
		Model 3: Geographical location	Administrative region
		Model 4: Socio-economic factors	Employment status and household income
		Model 5: Healthcare access factors	Health care provider number, routine check-up, distance between home and healthcare facility, and hospital type.
		Model 6: Fully adjusted + gender differences	Model 4 + interaction between education level and sex

As mentioned above, the outcome variable in this chapter is whether an individual's T2DM status was previously known/diagnosed or not (undiagnosed). This outcome encompassed 538 individuals who previously knew about their diabetes status and 399 individuals who did not. The analysis used the explanatory variables listed in Table 6.4. The main outcome is diagnosed diabetes defined as individuals who had clinically confirmed T2DM irrespective of their HbA1c level. The relationship between the outcome measures and the study predictors used the classical Chi-square "goodness-of-fit" test. Binary logistic regression was used to address all the statistical analyses, carried out using STATA (Stata Corp. Inc, TX, USA).

Table 6.2 Study explanatory variables list

Study variables		Categories
Demographic factors	Sex	Dichotomous variable: 1= Male, 2= Female. Reference category: Female
	Age	Categorical variable in questionnaire: (1: 15 - 24, 2: 25 – 34, 3: 35 – 44, 4: 45 – 54, 5: 55 – 64, 6: 65+) Reference category: (35 - 44)
	Marital status	This was recoded into three categories: 1 = Currently Married; 2 = Never married; 3 = Not currently married: Separated/divorced, and Widowed. Reference category: Never married
Socio-economic factors	Education level	Recoded into three categories: 1 = Low education level: cannot read or write, and no formal education. 2 = Middle education level: Primary and intermediate school education 3 = High education level: High school, university degree and above. Reference category: High education level
	Employment status	Coded into seven categories: 1 = Government employee; 2 = Non-government employee; 3 = Unemployed; 4 = Retired; 5 = Homemaker. 6 = Self-employed; 7 = Student. Reference category: Governmental employee
	Monthly household income	Recoded into five categories: 1 = Lowest : 1= Less than 3,000 Riyals; 2=3,000 Riyals to less than 5,000 2 = Low : 3=5,000 Riyals to less than 7,000; 4=7,000 Riyals to less than 10,000 3 = Middle : 5= 10,000 Riyals to less than 15,000; 6=15,000 Riyals to less than 20,000; 4 = High : 7=20,000 Riyals to less than 30,000; 5 = Highest: 8= 30,000 Riyals or more ⁵⁶ Reference category: Lowest
Healthcare access factors	Routine check-up	Recoded into two categories: 1 = Yes; 2 = No Reference category: Yes
	Time between home and healthcare facility	Recoded into two categories: 1 = Less than one hour; 2 = More or equal to one hour. Reference category: More or equal to one hour
	Hospital type	Recoded into three categories: 1 = Public: 1: Governmental (MOH) 2 = Private: 5: Private hospital 3 = Other: National Guard, Armed Forces, Military hospital, Other Reference category: Public
	Health care provider number	Recoded into three categories: 1 = No 2 = Yes, there is one place 3 = Yes, there are more than one place Reference category:
Administrative region	Region location	Recoded into five categories based on geographical location: 1 = Central region; 2 = Easter region; 3 = Western region; 4 = Northern region. 5 = Southern region Reference category: Western region

⁵⁶ 1000 Riyal = 266.5\$ = £ 216.03

6.3 Results

6.3.1 Descriptive results

Table 6.3 shows the characteristics of the sample used to examine factors relating to whether someone with diabetes knew about their status before the survey (the ‘previously known diabetes’ or diagnosed class) or were diagnosed subsequently. I tested whether the following factors were significant associated with being diagnosed diabetes or not using univariate analysis. These statistics show the percentage distribution of diagnosed and undiagnosed T2DM based on different socio-economic, demographic, healthcare access and administrative regions. In addition to the descriptive results in **Table 6.3**, the significance of associations between potentially explanatory variables and the study outcome was measured using bivariate analysis. Overall, 57.4% of the affected population had a T2DM diagnosis that predated the survey. Males were more likely to fall into this category (males 60.9%, females 53.2%), as were older groups (three times more likely to be in this category than younger age-groups).

The results also show that individuals in the low education group are more likely to know their diabetes status, with 69.3% being aware, with 41.5% in the higher educated group. Individuals who lived in the Eastern region were the most likely to be aware of their condition (71.0% diagnosed compared to 63.8% in the Western region that showed the next highest percentage). The results also indicated that retired and unemployed individuals had the highest proportion of previously diagnosed T2DM (76.0% and 70.0% respectively); this association was significant. In addition, individuals who are currently unmarried (divorced, separated, and widowed) were more likely to have been previously diagnosed (74.2%) compared to individuals who never married (16.3%), although this is likely to be related to age. Thus, three socioeconomic factors; education level, employment status and marital status have a significant relationship with the pre-diagnosed category. Unexpectedly, the household income variable did not show a significant relationship with pre-diagnosed status.

Several healthcare access variables examined were found to be significantly associated with pre-diagnosed diabetes. A higher proportion of individuals who had routine health check-ups fell into the diagnosed category (64.7%) compared to individuals who did not have routine check-ups (51.1%). Conversely, individuals who lived near to a healthcare facility were less likely to fall into the diagnosed category than those who lived far from healthcare facility - this observation was significant. However, this distance analysis considered only public healthcare centres, so individuals using private healthcare facilities would be missed. This is important, as individuals who attend private and other facilities - such as specialised public hospitals⁵⁷ - show a higher percentage of previous diagnosis than those attending public healthcare facilities.

However, it should be noted that the distance variable was affected by a high level of missing data and inaccuracy in time recording. The results also indicated that having healthcare insurance did not show any obvious association with pre-diagnosed T2DM. Other healthcare variables based on the conceptual framework that failed to show significant associations with pre-diagnosed status were healthcare provider number and the nature of the hospital attended (public or private). The proportions of individuals with pre-diagnosed diabetes are presented in **Figures 6.2, 6.3 and 6.4**. These results show that the highest proportion of undiagnosed diabetes is seen in those who are high-educated, female and who lived in the Central or Southern regions. Conversely, the proportion of individuals with undiagnosed diabetes was lowest among low-educated individuals, males and those who lived in Eastern or Western regions. However, the level of undiagnosed diabetes was still pronounced among these groups.

⁵⁷ Other group included other specialised public hospitals and other multi-centres. Specialised public hospitals are mainly targeting Saudi and expatriates' employees and their families (see Saudi healthcare system structure in chapter two).

Table 6.3 Characteristics of survey participants of diagnosed and undiagnosed T2DM

Variables		Categories N (%)	Total 937	Undiagnosed 399 (42.6)	Diagnosed 538 (57.4)	p
Sex	Female		423	198 (46.8)	225 (53.2)	0.018
	Male		514	201 (39.1)	313 (60.9)	
Age	15 - 24		59	56 (94.9)	3 (5.1)	<0.001
	25 - 34		86	70 (81.4)	16 (41.1)	
	35 - 44		151	89 (58.9)	62 (41.1)	
	45 - 54		216	84 (38.9)	132 (61.1)	
	55 - 64		193	57 (29.5)	136 (70.5)	
	65 +		232	43 (18.5)	189 (81.5)	
Education level	Low education		413	125 (30.3)	288 (69.7)	<0.001
	Middle education		237	106 (44.7)	131 (55.3)	
	High education		287	168 (58.5)	119 (41.5)	
Employment status	Government employee		238	129 (54.2)	109 (45.8)	<0.001
	Non-government employee		27	16 (59.3)	11 (40.7)	
	Unemployed		150	45 (30.0)	105 (70.0)	
	Retired		196	47 (24.0)	149 (76.0)	
	Homework		238	106 (44.5)	132 (55.5)	
	Self-employed		44	16 (36.4)	28 (63.6)	
	Student		42	39 (92.9)	3 (7.1)	
	Other		2	1 (50.0)	1 (50.0)	
Monthly Household income	Lowest/ poorest		406	157 (38.7)	249 (61.3)	0.150
	Low / poor		130	56 (43.1)	74 (56.9)	
	Middle		107	42 (39.3)	65 (60.8)	
	High / rich		64	30 (46.9)	34 (53.1)	
	Highest / richest		46	23 (50.0)	23 (50.0)	
	Unknown or decline to respond		184	91 (49.5)	93 (50.5)	
Marital status	Currently married		701	287 (40.9)	414 (59.1)	<0.001
	Never married		86	72 (83.7)	14 (16.3)	
	Divorced, Separated, and widowed		147	38 (25.8)	109 (74.2)	
	Unknown / decline to respond		3	2 (66.7)	1 (33.3)	

Table 6.3 Characteristics of survey participants of diagnosed diabetes (cont.)

Variables	Categories N (%)	Total 937	Undiagnosed 399 (42.6)	Diagnosed 538 (57.4)	p
Routine check-up	No	483	236 (48.9)	247 (51.1)	<0.001
	Yes	377	133 (35.3)	244 (64.7)	
	Missing	77	30 (39.0)	47 (61.0)	
Time between home and healthcare facility	Less than one hour	667	270 (40.5)	397 (59.5)	0.007
	More than one hour or equal	68	24 (35.3)	44 (64.7)	
	Missing	202	105 (52.0)	97 (48.0)	
Hospital type	Public	641	260 (40.6)	381 (59.4)	<0.001
	Private	73	29 (39.7)	44 (60.3)	
	Other	84	27 (32.1)	57 (67.9)	
	Missing	139	83 (59.7)	56 (40.3)	
Health care provider number/ insurance	No	251	108 (43.0)	143 (57.0)	0.299
	Yes, there is one place	263	122 (46.4)	141 (53.6)	
	Yes, there are more than one place	378	154 (40.7)	224 (59.3)	
	Missing	45	15 (33.3)	30 (66.7)	
Region location	Central region	307	144 (46.9)	163 (53.1)	0.048
	Eastern region	38	11 (29.0)	27 (71.0)	
	Western region	138	50 (36.2)	88 (63.8)	
	Northern region	150	57 (38.0)	93 (62.0)	
	Southern region	304	137 (45.1)	167 (54.9)	

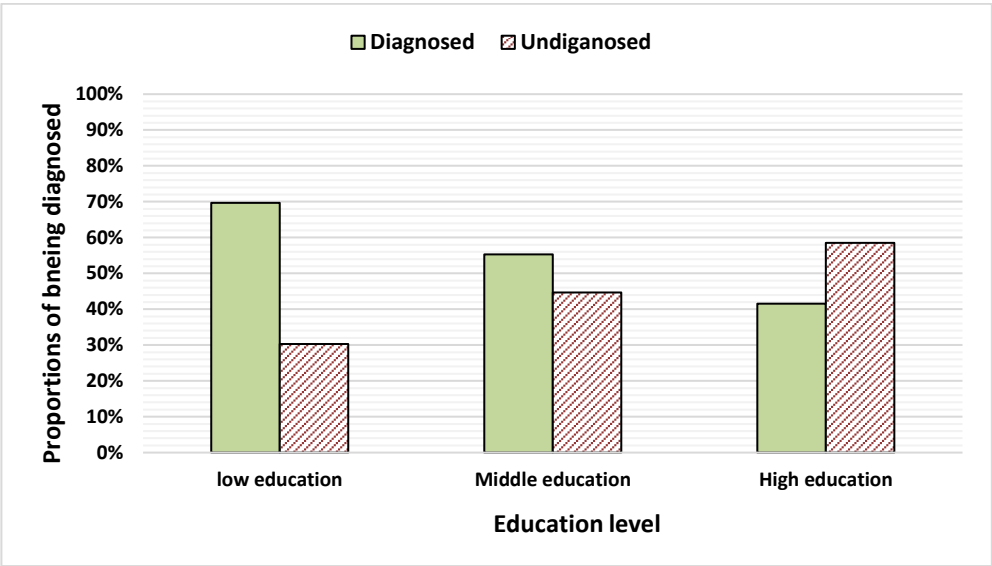


Figure 6.2 Proportions of diagnosed and undiagnosed diabetes by education level

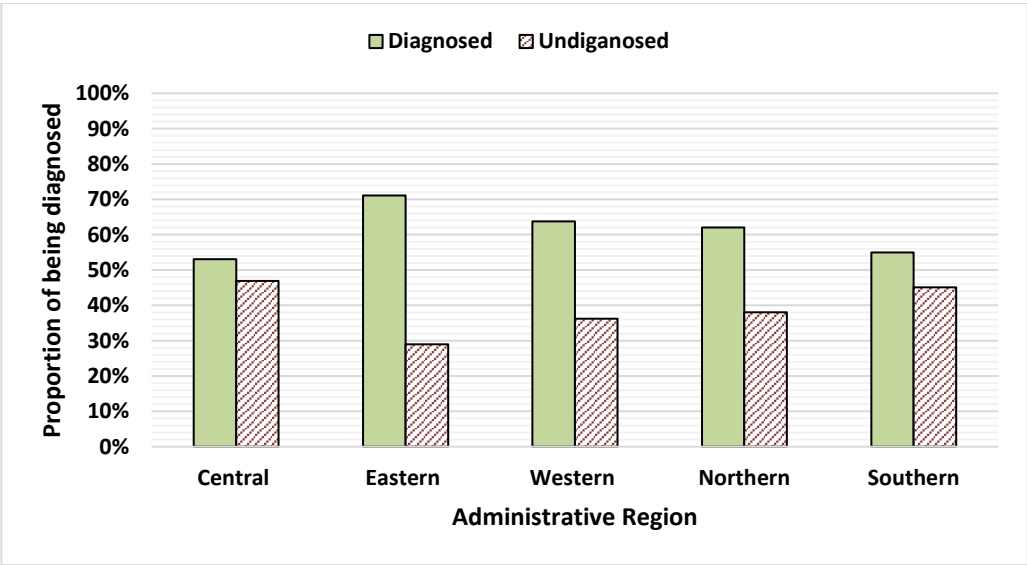


Figure 6.3 Proportions of diagnosed and undiagnosed diabetes by administrative region

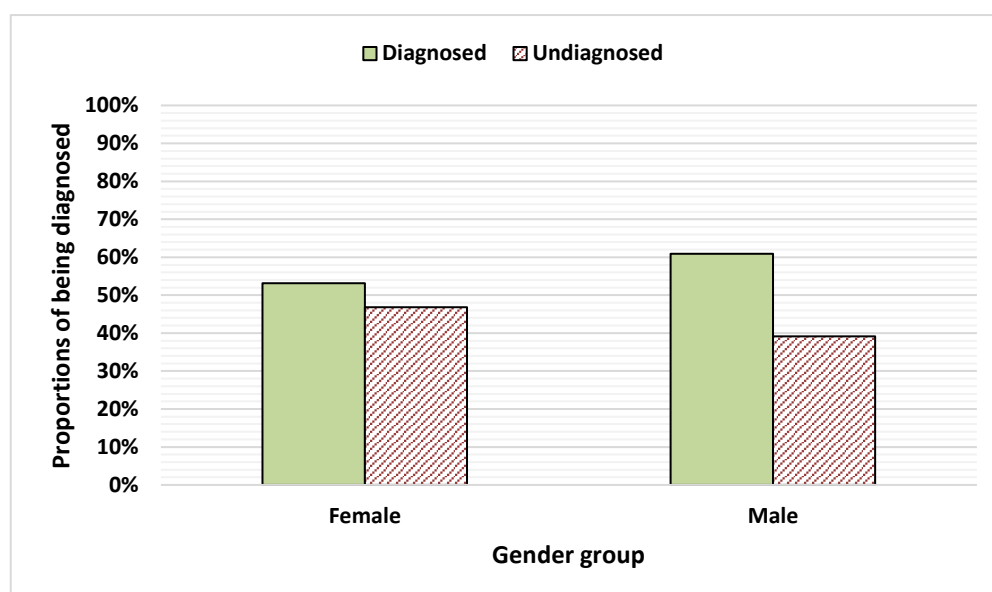


Figure 6.4 Proportions of diagnosed and undiagnosed diabetes by gender group

6.3.2 Analytical results

The following **Table 6.4** presents a series of nested logistic regression models that were constructed based on the analytical framework shown in **Table 6.1**. The models explored the association between previously diagnosed T2DM and education level, with subsequent models adding and controlling for demographic, region, other socioeconomic, and healthcare access factors. The progressive nature of the six models allowed me to test the effect of the various control variables on the intensity of the association between pre-diagnosed T2DM and education level. The first model shows that low-educated individuals were more likely to be previously diagnosed than high-educated individuals (OR: 2.717, p -value<0.05), as were middle-educated individuals (OR: 1.902, p -value<0.05).

In the **second model** (adjusted with demographic variables), however education level did not show a significant relationship with previously diagnosed T2DM, an individual's age appeared to have an important relationship with this outcome; individuals aged 25 - 34 year were less likely to be diagnosed compared with individuals aged 35 - 44 and older (OR: 0.269, p -value<0.05). Similarly, the 45 – 54 age group were more likely to be represented than those aged 35 – 44 and younger (OR: 1.868, p <0.05). Regarding gender, males were more likely to fall into the previously diagnosed category compared to females (OR: 1.570, p -value<0.05). Marital status was not significantly associated with the diagnosed of T2DM in any of the nested logistic models. The administrative region variable was included in the third adjusted model but showed no significant association with previously diagnosed T2DM.

In **model four**, two socioeconomic factors were included in the analysis. Employment status had no significant relationship with diagnosed T2DM outcome in any of following nested models. In contrast, monthly household income was a reliable predictor of previously diagnosed status, with low-income individuals over-represented compared to the lowest income group (OR: 2.045, p-value<0.05). The middle-income group also showed a higher likelihood of previously diagnosed status than lowest-income individuals (OR: 1.996, p-value<0.05). After controlling for health care access predictors in the **fifth model**, all three variables had no significant relationship with individuals who were previously diagnosed with T2DM. Similarly, hospital type, routine check-up and distance between home and health facility had no association with previously diagnosed T2DM in any of the nested models. Adjusting the **final model** with a term for education-sex interaction also had no significant effect.

In conclusion, with regards to demographic variables, sex and marital status were not significantly related to previously diagnosed T2DM. However, an individual's age was a significant demographic indicator. Monthly household income did have an important relationship with diagnosed T2DM in the last three nested models, with low-income individuals having the highest likelihood of being in the previously diagnosed group. Region and other healthcare access factors showed no significant association with previously diagnosed T2DM status. As previously mentioned in Chapter 5, some of these associations were not statistically significant due to the small size of study population that reduced the statistical power of the analysis.

Table 6.4 Generalised binary logistic regression odds ratios and 95% confidence interval of diagnosed T2DM (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Education level												
High education (ref)	1		1		1		1		1		1	
Low education	2.717 (1.764 - 4.185)	0.000	1.178 (0.679 - 2.042)	0.560	1.202 (0.692 - 2.090)	0.514	1.717 (0.925 - 3.185)	0.086	1.771 (0.945 - 3.317)	0.074	1.91 (0.769 - 4.748)	0.163
Middle education	1.902 (1.190 - 3.040)	0.007	1.563 (0.905 - 2.699)	0.109	1.601 (0.918 - 2.793)	0.097	1.934 (1.092 - 3.426)	0.024	1.985 (1.122 - 3.51)	0.018	2.02 (0.773 - 5.296)	0.151
Age group												
35 – 44 (ref)			1		1		1		1		1	
15 - 24			0.058 (0.008 - 0.448)	0.006	0.059 (0.008 - 0.438)	0.006	0.047 (0.003 - 0.713)	0.028	0.036 (0.002 - 0.590)	0.020	0.036 (0.002 - 0.591)	0.020
25 - 34			0.269 (0.119 - 0.606)	0.002	0.277 (0.123 - 0.622)	0.002	0.295 (0.125 - 0.698)	0.005	0.281 (0.123 - 0.639)	0.002	0.282 (0.123 - 0.649)	0.003
45 - 54			1.868 (1.065 - 3.273)	0.029	1.888 (1.073 - 3.322)	0.028	1.716 (0.962 - 3.064)	0.068	1.546 (0.857 - 2.791)	0.148	1.538 (0.851 - 2.780)	0.154
55 - 64			1.858 (1.004 - 3.438)	0.049	1.919 (1.044 - 3.526)	0.036	1.628 (0.832 - 3.185)	0.155	1.632 (0.828 - 3.217)	0.157	1.622 (0.822 - 3.198)	0.163
65 +			4.054 (1.967 - 8.352)	0.000	4.303 (2.085 - 8.879)	0.000	3.748 (1.666 - 8.432)	0.001	3.464 (1.518 - 7.904)	0.003	3.505 (1.517 - 8.094)	0.003
Sex												
Female (ref)			1		1		1		1		1	
Male			1.570 (1.002 - 2.461)	0.049	1.579 (1.011 - 2.467)	0.045	1.289 (0.684 - 2.427)	0.432	1.367 (0.744 - 2.512)	0.314	1.428 (0.635 - 3.210)	0.388
Marital status												
Never married (ref)			1		1		1		1		1	
Currently married			1.438 (0.310 - 6.671)	0.643	1.462 (0.333 - 6.426)	0.615	1.326 (0.276 - 6.375)	0.725	1.369 (0.268 - 6.987)	0.706	1.375 (0.273 - 6.917)	0.699
Divorced, Separated, and widowed			2.159 (0.417 - 11.199)	0.359	2.199 (0.446 - 10.835)	0.333	2.153 (0.402 - 11.512)	0.370	2.166 (0.378 - 12.416)	0.386	2.153 (0.380 - 12.176)	0.385
Decline to respond or unknown			0.579 (0.057 - 5.915)	0.645	0.622 (0.064 - 6.063)	0.683	0.574 (0.061 - 5.356)	0.626	0.853 (0.078 - 9.355)	0.897	0.876 (0.081 - 9.438)	0.913

Model 1: unadjusted; **Model 2:** adjusted by demographic factors (age, sex and marital status); **Model 3:** includes also administrative regions; **Model 4:** includes also socio-economic factors (employment status and monthly household income); **Model 5:** includes also Healthcare access factors (Distance between home and health facility, routine check-up, and hospital type); **Model 6:** includes also education level and sex interaction term, **P:** p-value (significant in **BOLD**); **OR:** Odds Ratios; **CI:** Confidence interval.

Table 6.4 Generalised binary logistic regression odds ratios and 95% confidence interval of diagnosed T2DM (Household weighted models) (cont.)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Administrative Regions												
Western region (ref)					1		1		1		1	
Central region					1.319 (0.769 - 2.265)	0.314	1.251 (0.726 - 2.154)	0.420	1.296 (0.735 - 2.283)	0.370	1.298 (0.731 - 2.306)	0.374
Eastern region					1.302 (0.453 - 3.740)	0.624	1.603 (0.533 - 4.814)	0.400	1.695 (0.563 - 5.106)	0.348	1.697 (0.563 - 5.110)	0.347
Northern region					1.254 (0.576 - 2.733)	0.569	1.196 (0.537 - 2.668)	0.661	1.193 (0.543 - 2.618)	0.661	1.185 (0.537 - 2.616)	0.674
Southern region					1.069 (0.615 - 1.858)	0.813	1.077 (0.610 - 1.899)	0.798	1.117 (0.623 - 2.101)	0.664	1.139 (0.618 - 2.098)	0.677
Employment status												
Government employee (ref)							1		1		1	
Non-government employee							0.674 (0.133 - 3.409)	0.633	0.770 (0.149 - 3.983)	0.756	0.776 (0.150 - 3.995)	0.761
Unemployed							1.268 (0.562 - 2.862)	0.568	1.285 (0.578 - 2.861)	0.539	1.275 (0.565 - 2.873)	0.558
Retired							1.696 (0.787 - 3.655)	0.177	1.658 (0.760 - 3.614)	0.204	1.683 (0.770 - 3.676)	0.192
Homeworker							0.954 (0.459 - 1.979)	0.899	0.996 (0.482 - 2.059)	0.991	0.976 (0.457 - 2.085)	0.950
Self-employed							1.074 (0.447 - 2.583)	0.873	1.031 (0.416 - 2.552)	0.947	1.051 (0.422 - 2.614)	0.916
Student							1.539 (0.180 - 13.126)	0.694	2.281 (0.262 - 19.83)	0.455	2.309 (0.264 - 20.202)	0.449
Other							1.091 (0.064 - 18.552)	0.952	0.812 (0.048 - 13.84)	0.886	0.806 (0.047 - 13.718)	0.882
Monthly household income												
Lowest (ref)							1		1		1	
Low							2.045 (1.059 - 3.948)	0.03	2.088 (1.086 - 4.017)	0.027	2.086 (1.084 - 4.017)	0.028
Middle							1.996 (1.016 - 3.919)	0.05	2.012 (1.018 - 3.977)	0.044	2.009 (1.018 - 3.965)	0.044
High							1.536 (0.664 - 3.551)	0.32	1.406 (0.590 - 3.345)	0.441	1.403 (0.590 - 3.334)	0.443
Highest							1.561 (0.637 - 3.821)	0.33	1.582 (0.617 - 4.047)	0.339	1.592 (0.619 - 4.087)	0.334
Missing							0.573 (0.334 - 0.983)	0.04	0.557 (0.323 - 0.961)	0.035	0.553 (0.321 - 0.953)	0.033

Table 6.4 Generalised binary logistic regression odds ratios and 95% confidence interval of previously unknown diabetes (Household weighted models) (cont.)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Hospital type												
Governmental (MOH) (ref)									1		1	
Private									0.764 (0.352 – 1.652)	0.493	0.762 (0.351 - 1.654)	0.491
Other									1.083 (0.589 - 1.993)	0.797	1.079 (0.584 - 1.990)	0.808
Missing									0.433 (0.235 - 0.799)	0.007	0.431 (0.233 - 0.798)	0.007
Routine check-up												
Yes (ref)									1		1	
No									0.716 (0.461 - 1.110)	0.136	0.714 (0.459 - 1.108)	0.133
Missing									0.398 (0.187 – 0.846)	0.017	0.398 (0.187 - 0.848)	0.017
Distance between home and health facility												
More than 60 min									1		1	
Less than 60 min									0.424 (0.171 - 1.044)	0.062	0.422 (0.170 - 1.046)	0.062
Missing									0.534 (0.195 - 1.459)	0.221	0.532 (0.193 - 1.464)	0.222
sex × Education												
											1	
Male × low education											0.869 (0.291 – 2.596)	0.802
Male × Middle education											0.979 (0.311 - 3.073)	0.971

6.4 Discussion and conclusion

This analysis was carried out to explore the association between education level and pre-diagnosis of T2DM. This led to the development of a risk score model for identifying individuals who are at high risk of having undiagnosed diabetes. The risk score model will be discussed in more depth in the following chapter. A descriptive analysis of the results revealed that individuals from lower socioeconomic groups were less likely to have undiagnosed T2DM. As expected, most individuals with abnormally high Hb1Ac had diagnosed T2DM (57.4%), with the remaining individuals (43.6%) classified as having undiagnosed T2DM. However, the prevalence of undiagnosed T2DM is particularly high compared to that reported in other high-income countries (Beagley et al., 2014). The proportion of undiagnosed diabetes is estimated at 28.4% in high-income countries, only 40.7% of that seen in the MENA region. This is probably resulting from a combination of limited access to healthcare and poor awareness of diabetes among the population.

Diagnosed diabetes was most prevalent in males, individuals aged over 65, people who attended private hospitals and those who lived in the Eastern administrative region. In contrast, undiagnosed diabetes was more prevalent in younger age groups and individuals with higher education levels and who attended public hospitals. These findings are in accordance with previous literature on the relationship between education levels and diagnosed T2DM. However, health care access has emerged as an additional important risk factor for diagnosed T2DM in this study. These observations are significant, as previous epidemiological studies in Saudi Arabia on diagnosed diabetes have concentrated on demographic and lifestyle risk predictors, with little consideration of socioeconomic and health care access factors.

The gender difference was significant and was expected as males had a higher prevalence of T2DM than females, consistent with previous studies. The effect of age was also significant. As discussed in the chapter five, older age groups are more likely to have T2DM. This result appears consistent with previous literature as older individuals are more likely to use health care services. Note that a higher prevalence does not necessarily translate into a higher proportion of individuals who are aware of their condition. This applies to both the apparent gender and age effects.

The unadjusted model indicates that there is a significant positive relationship between undiagnosed T2DM and educational level. This finding is inconsistent with a previous US study, which provided evidence that better educated individuals were more likely to be diagnosed with T2DM (Robbins et al., 2005). It might be expected that better educated individuals had more knowledge about the disease and its symptoms. The results of the present study confirm that education level is an important variable, but not after controlling with other factors. The discrepancy with the US study in terms of the direction of the association might be explained by the fact that, in the Saudi population studied here, there is a strong relationship between age and education level. This would lead to an excess of diagnosed diabetes among less educated individuals, as this group would tend to be older (Bener et al., 2009), even when controlling for age. Thus, the different age-distribution in the current study means that those in older age-groups are more likely to be contained in the low education category. This interpretation is likely to be correct, as the effect of education level disappeared after correcting the model for sex, age, and marital status. This is an unexpected result, because less well-educated individuals were more likely to have diagnosed T2DM if they were male, and/or fell into the middle and older age groups. Age and gender were both independently and significantly associated with diagnosed T2DM.

Studies reviewed in chapter three show that individuals in the older and middle aged groups were more likely to be diagnosed with T2DM than those in the younger age group, with no gender differences (Bener et al., 2009). Gender was significantly associated in first two models; thus this observation is at odds with previous studies showing that females were more likely to be diagnosed with T2DM than males (Al-Nuaim, 1997). However, this effect was not significant in the final model. One possible explanation could be the strong inverse relationship between education level and age. These results therefore indicate that education level is not a key predictor of T2DM diagnosis after controlling for other demographic variables.

Marital status was not significantly associated in this model, however previous studies indicated that marital status was associated with diagnosed diabetes (more prevalent in the married and previously married groups). The results of the current study do, however, show a different trend. It has been reported that married and previously married individuals have a higher likelihood of being diabetic (Murad et al., 2014), and so are more likely to be diagnosed with T2DM than the never married group. Even after controlling the model for administrative region, an individual's age and sex were significantly associated with diagnosed T2DM. This result indicates that there is no regional variation in pre-diagnosis of T2DM.

Previous studies were more focused on the influence of regional variation in healthcare access and quality on the prevalence of T2DM, reflecting regional differences in urbanisation level (Brown et al., 2004, Sommer et al., 2015). Individuals who lived in rural areas were less likely to be diagnosed than those who lived in urban areas (Sicree and Shaw, 2007). This result might be expected, as the Eastern region is classified as a socioeconomically advantaged region with correspondingly high-quality education and health-care provision. This result reached significance, and has been identified in previous studies, indicating that geographic region is a proxy of community socioeconomic status and urbanisation - see the previous discussion in chapters two and three.

In terms of controlling the model for other socioeconomic factors, the results demonstrated that middle-educated individuals were more likely to be diagnosed with T2DM compared to high-educated individuals. The result of this model observe that education level is now significantly associated with diagnosed T2DM, and that only certain age groups show association with diagnosed T2DM. Gender, region and employment status appeared did not show significant association with diagnosed T2DM. However, monthly household income was significantly associated, however this was only observed in low and middle-income groups. Some of these findings are not in line with the findings of past studies, which provided evidence that high socioeconomic status individuals in the United States were more likely to be diagnosed with T2DM compared to low SES individuals (Clark and Utz, 2014, Wilder et al., 2005, Robbins et al., 2005). A possible explanation of this finding might be that individuals in low and middle-income groups are less likely to be well-educated and/or in younger age groups, then they are more likely to be diagnosed with T2DM diabetes based on previous models. Employment status is another variable that might influence this relationship; however, it was not significantly associated with diagnosed T2DM as any effect was removed by the strong relationship between age and diagnosed diabetes.

With regard to the effect of health care access, middle-educated individuals were more likely to be diagnosed with T2DM. Interestingly, the results highlighted that hospital type, routine check-up, and distance between home and health care facility were not significant predictors of diagnosed T2DM. The association with health-care access is probably explained by the fact that middle-educated individuals fall mainly into middle or older age groups who have increased prevalence of T2DM. This result is inconsistent with the previous research on health care utilisation and diagnosis of T2DM discussed in Chapter two, which highlighted that health care access factors were important predictors of diagnosed T2DM. Individuals who were pre-diagnosed with T2DM were more likely to have better health care access than those who were unaware of their status.

A previous study identified that individuals who had insurance and routinely visited health care in Mexico were more likely to have diagnosed T2DM (Zhang et al., 2010), however it did not consider the importance of health care access influence as a proxy of socioeconomic status for diagnosed T2DM. This is an area of active research in the U.S., where the diversity in health care utilisation may significantly influence T2DM outcomes. Some studies have found an association between undiagnosed T2DM and lack of health care access (Zhang et al., 2008), indicating that access to health care services provides early detection through screening in health care centres and prevention by adherence to healthier lifestyle (Clark and Utz, 2014). This point was discussed in Chapter 2 with reference to a concern in Saudi Arabia concerning a lack of community health care studies that could impact on the incidence of T2DM and its early diagnosis. Individuals from lower socio-economic groups are less likely to participate in private health care facilities in Saudi Arabia. Access to these facilities relies on the ability to self-fund, or having work place health insurance schemes that cover private hospitals and additional health care services (Walston et al., 2008).

Individuals who are unemployed, students, housewives and employees without health care insurance were more likely to be unaware of their T2DM status for a long time. In light of these considerations, I added an employment/healthcare access factor interaction term to the fully adjusted model. This showed that there is no significant interaction between employment and health care insurance in relation to diagnosed T2DM. Some public healthcare facilities in the country provide poor health care services that do not carry out routine blood tests (Almalki et al., 2011). This might explain the low likelihood of early T2DM detection in some rural regions compared with other more urban regions in Saudi Arabia. In the present study, the influence of the gender and education level interaction terms on diagnosed T2DM was not significant. Only age and monthly household income have a major influence on the relationship between education level and diagnosed T2DM in Saudi Arabia.

To conclude, the purpose of the current study was to investigate the relationship between education level and T2DM diagnosis. It has shown that age and monthly income are the most important risk factors for diagnosed T2DM and bring out several questions in need of further investigation. The following chapter will follow on from this study by developing a prediction model for undiagnosed diabetes.

Chapter 7 Development of Diabetes and Pre-Diabetes

Prediction Risk Models: Identifying Undiagnosed Cases among the Saudi population

7.1 Introduction

This chapter extends the previous analysis in an attempt to help develop a T2DM risk model for predicting undiagnosed cases, as we need to better understand the scale of and reasons underlying undiagnosed T2DM. This should have the dual objective of preventing T2DM where possible, and the earlier identification of individuals with undiagnosed diabetes. The early identification of undiagnosed cases in Saudi Arabia will improve health outcomes for those with T2DM. In this context, several studies to date have already developed risk score tools based on demographic and anthropometric measures. These tools predict the likelihood of an individual having undetected diabetes using a range of different diagnosis criteria. However, these risk score models are not directly applicable to the Saudi population due to socio-economic and lifestyle differences.

This chapter is divided into five main sections: introduction, background and rationale, methods, results, discussion, and conclusion. **Section 7.2** reviews previous studies that aimed to develop a risk score model for undiagnosed T2DM. In **Section 7.3**, the underlying statistical analysis, development, and validation of the prediction model for undiagnosed T2DM and prediabetes will be described. The results are divided into three subsections and will be presented in **Section 7.4**. The first subsection presents the characteristics of study participants, the following two subsections will address the development and validation of undiagnosed T2DM and prediabetes prediction risk score models. This chapter will end with a discussion and summary of the main conclusions in **Section 7.5**.

7.2 Background and rationale

Screening of individuals for the early detection of undiagnosed diabetes is potentially an effective approach to tackle the complications of the disease and ensure adequate control in the future (Wilder et al., 2005, Clark and Utz, 2014). A review of T2DM diagnosis criteria concluded that there are differences in diabetes diagnosis guidelines in the GCC countries (Alhyas et al., 2011).

This, along with differences in socio-economic and lifestyle factors, means that diabetes risk score models based on the wider GCC region cannot be applied to the Saudi population. A screening tool to identify individuals at high-risk of undiagnosed diabetes in Saudi Arabia would therefore be of great value, particularly if it is based on Saudi data. The tool would enable the identification of individuals at high-risk of progressing to undiagnosed T2DM, allowing interventions such as close monitoring, life-style advice, and potentially pre-emptive medication. This in turn would help to reduce the burden of T2DM complications that set in early after the onset of diabetes. An accurate prediction model is an essential pre-requisite for such a screening tool, and it would also help with the development of new public health policies to reduce the prevalence of undiagnosed diabetes (Collins et al., 2011).

A considerable amount of research has been published on undiagnosed T2DM risk score modelling. These studies suggest that using the risk score model results in clinical practice is particularly valuable in those countries with the highest risk of diabetes. Once individuals at high risk of developing diabetes are identified using the risk score assessment tool, they can then undergo one or more screening tests that include: a fasting blood glucose test; oral glucose tolerance test (OGTT); or measurement of blood *HbA1c* levels. Despite the large number of studies in different populations aimed at developing a risk score model for T2DM, there have been few attempts to deliver such a model in GCC countries. Significantly, previous prediction models have not given much attention to socio-economic factors. This observation was one of the main motivations behind the present study. The following sub-section provides an overview of risk prediction models for T2DM, including a consideration of methods for their development and validation.

The following tables provide a list of relevant prediction models for undiagnosed diabetes and prediabetes (in **Table 7.1 and 7.2**). Different terms have been used to identify relevant articles that presented risk prediction model for undiagnosed diabetes. To find relevant studies, a literature search of Medline, Science Direct, Google Scholar and Web of Science for data sources reporting undiagnosed diabetes conducted from 1990 to 2018 using the search terms: 'diabetes', 'undiagnosed', 'prediction model', 'prediabetes'.

Only those studies with the primary aim of developing risk score models for undiagnosed diabetes and prediabetes are included in this chapter. I excluded any articles aimed at developing prediction models for T1DM, involving gestational diabetes or based on children. In addition, relevant articles referred to in these primary sources were also reviewed. This approach identified a spectrum of studies based on a range of source data sets and predictors used to build the prediction model.

Most of the models identified included demographic, anthropometric and lifestyle factors. The most commonly observed predictors were age, waist circumference, body mass index (BMI) or obesity, family history of diabetes and hypertension. A single Chinese study was unique in the empirical literature in that it developed separate prediction models based on gender (Xie et al., 2010). While a few studies included socio-economic factors, these were not found to be significant. Furthermore, only four studies have developed diabetes risk score models specifically for the GCC region: these were set in Kuwait (Al Khalaf et al., 2010b), Saudi Arabia (Memish et al., 2015), Oman (Al-Lawati and Tuomilehto, 2007) and the UAE (Sulaiman et al., 2018).

However, none of the four GCC studies used socio-economic factors in developing the model. The Sulaiman study set in the UAE found gender, education level and income status were not significantly associated with undiagnosed diabetes in the UAE. However, this model only covered Sharjah city and the Northern Emirates. The Kuwaiti study by Al-Khalaf observed that exercise, diet, and smoking were not associated with undiagnosed diabetes. Similarly, the Omani study found that gender, family history of hypertension and smoking were also not significantly associated. The most common risk predictors used in the GCC region prediction model were age, waist-circumference, family history of diabetes and hypertension. However, the Saudi model developed by Memish observed that smoking status and a history of gestational diabetes were additional predictors for the Saudi population. Family history and a sibling with diabetes were also identified as valuable predictors.

Another study by Handlos et al. (2013) developed two MENA regional prediction models for diabetes and diabetes or prediabetes based on a pooled sample of individuals from Saudi Arabia (Riyadh, and Jeddah), UAE (Dubai, Sharjah and Abu Dhabi), Algeria (Blida) collected between November 2010 and December 2011. The final model included age, BMI, family history, number of siblings with diabetes, gestational diabetes for females and ethnicity. The current study therefore addresses two significant gaps in our knowledge by assessing the importance of socio-economic factors in the identification of undiagnosed diabetes, and by applying these specifically to a Saudi population that has to date been poorly represented.

Regarding the *diagnostic criteria*, most studies in the literature used the Oral Glucose Tolerance test (OGTT) and/or the Fasting Plasma Glucose (FPG) diagnostic criteria. However, one recent study in the UAE used the HbA1c test in order to identify diabetes status (Sulaiman et al., 2018). Most previous studies have based their models on individuals aged 30 years or above (middle-aged groups).

However, other models addressing populations in India (Mohan et al., 2005, Ramachandran et al., 2005), the USA (Bang et al., 2009) and China (Gao et al., 2010) as well the GCC countries were based on individuals aged 20 years and above. The use of a lower threshold of 20 years is probably due to the higher rates of T2DM in these countries, reflecting an earlier age of onset for disease necessitating the inclusion of younger age groups in the risk models.

With regard to *model building*, five studies used univariate screening to include risk predictors to the model (Ruige et al., 1997, Gao et al., 2010, Chaturvedi et al., 2008, Glümer et al., 2004, De Sousa et al., 2009); in two other studies the approach used was unclear (Keesukphan et al., 2007, Ko et al., 2010). Two further studies included all predictors in the final model (Mohan et al., 2005, Ramachandran et al., 2005). The remaining studies used automatic variable selection techniques such as forward selection (Al Khalaf et al., 2010b, Ko et al., 2010), backward elimination (Ruige et al., 1997, Al-Lawati and Tuomilehto, 2007, Borrell et al., 2007, Gao et al., 2010, Glümer et al., 2004, De Sousa et al., 2009), and stepwise regression (Baan et al., 1999, Keesukphan et al., 2007) to identify the most appropriate risk predictors for the final multivariate model.

Two studies used internal validation by randomly splitting the data into development and validation datasets (Glümer et al., 2004, Ramachandran et al., 2005). However, for the majority of the studies, the model was validated using another dataset (Baan et al., 1999, Al-Lawati et al., 2002, Bang et al., 2009, Chaturvedi et al., 2008, Gao et al., 2010, Glümer et al., 2004, Keesukphan et al., 2007, Ko et al., 2010, Mohan et al., 2005, De Sousa et al., 2009). In most cases, the performance of the prediction model was evaluated using C-statistic discrimination measures (the area under ROC, sensitivity, specificity). Some studies applied further validation strategies behind the discrimination measures.

However, only two studies (the Chinese and Egyptian risk models) assessed model performance using a calibration measure (Hosmer-Lemeshow goodness of fit test) to compare predicted risks to observed risks (Gao et al., 2010, Tabaei and Herman, 2002). Regarding model presentation, most studies derived a simple scoring system from the relevant risk model (Ruige et al., 1997, Al-Lawati and Tuomilehto, 2007, Baan et al., 1999, Bang et al., 2009, Chaturvedi et al., 2008, Gao et al., 2010, Glümer et al., 2004, Ko et al., 2010, Mohan et al., 2005, De Sousa et al., 2009, Ramachandran et al., 2005). These prediction model studies resulted in a simple point system obtained by multiplying or dividing the regression coefficients by a constant (typically 10) and then rounding the values to the nearest integer.

Based on the above summary of published prediction models and findings from previous chapters, I generated an analytical framework for better understanding the association between different risk factors and undiagnosed diabetes. This is shown in diagrammatic form in chapters 2 and 6. The diagram includes a set of risk predictors based on the previous literature and some additional individual and community socio-economic and healthcare access factors. As mentioned in the previous section, the most common classes of risk predictor for undiagnosed diabetes were demographic and lifestyle factors. I have not found any systematic attempt to consider socio-economic and healthcare access factors in a prediction model. This research study will attempt to add predictors from these unrepresented classes to the prediction model.

For example, socio-economic factors such as education level and household income could influence an individual's lifestyle and access to healthcare. Based on the literature review in Chapter 2, highly educated individuals were more likely to have been diagnosed with T2DM compared to less well-educated individuals. This could be explained by variation in lifestyle affecting susceptibility to disease, but also by probability of diagnosis due to improved access to health care, workplace testing and awareness of the disease and its symptoms. However, individuals with lower socio-economic status may have greater morbidity prompting them to attend health care for health problems which might also result in increased testing.

Table 7.1 Summary of previously published prevalence of T2DM prediction models (undiagnosed) and their risk factors

Author	Population	Significant variables included in the final prediction model	Definition of undiagnosed diabetes	Sample size	Age group	Measure performance
Ruige <i>et al.</i> (1997)	The Netherlands	Frequent thirst, pain when walking with a need to slow down, shortness of breath when walking, age, gender, obesity (BMI), obesity (men), family history of diabetes, use of antihypertensive drugs, reluctance to use a bicycle for transportation	Participants classified based on WHO criteria 75-g OGTT	N = 2,364 (1993) N = 786 (1995)	(50 - 74) (45 - 74)	AUC = 0.69 AUC = 0.62
Griffin <i>et al.</i> (2000)	United Kingdom	Age, gender, prescribed antihypertensive medication, prescribed steroids, BMI, family history of diabetes, smoking status	WHO criteria	N = 1,077	40 - 64	AUC = 0.80 Sensitivity = 0.77 Specificity = 0.72
Glümer <i>et al.</i> (2004)	Denmark	Age, BMI, gender, known hypertension, physical activity, family history of diabetes	Individuals without known diabetes and with FPG ≥ 7.0 mmol/L or 2-h PG ≥ 11.1 mmol/L defined as SDM	N = 6,784	30 - 60	First half: AUC = 0.804 Second half: AUC = 0.761 Pilot study AUC = 0.803
Mohan <i>et al.</i> (2005)	India	Age, abdominal obesity (waist circumference), physical activity, family history of diabetes	Diabetes diagnosis based on WHO criteria, 2-hr PG ≥ 200 mg/dL	N = 2,350	≥ 20	AUC = 0.698 Sensitivity = 0.725 Specificity = 0.601
Ramachandran <i>et al.</i> (2005)	India	Age, family history of diabetes, BMI, waist circumference, physical activity	Diabetes diagnosis based on 2-h PG ≥ 11.1 mmol/L	(2000) Cohort 1 n = 4,993; Cohort 2 n = 5,010. (1995) Cohort 3 n = 2,002 (1999) Cohort 4 n = 676	≥ 20	Sensitivity = 0.766, Specificity = 0.599
Wang <i>et al.</i> (2015)	Southern Chinese	Age, family history of diabetes, BMI, waist circumference.	Individuals without known diabetes and with FPG ≥ 7.0 mmol/L or 2-h PG ≥ 11.1 mmol/L	Guangzhou N = 3,162 Tibet N = 1,289	≥ 20	Female: AUC = 0.75 Men: AUC = 0.73

ADA, American Diabetes Association; **BMI**, body mass index; **WHO**, World Health Organisation; **FPG**, fasting plasma glucose; **OGTT**, oral glucose tolerance test; **CVD**, cardiovascular disease; 2-h PG, two-hour 75-g post-load plasma glucose level; **IGT**, impaired glucose tolerance; **IFG**, impaired fasting glucose; **SDM**, screen-detected diabetes; **BMI**, body mass index; **WHO/IDF**, World Health Organisation/International Diabetes Federation; **2-h PG**, two-hour 75-g post-load plasma glucose level; **ROC**, Receiver Operating Characteristic; **AUC**, Area Under the Curve.**HL**: Hosmer-Lemeshow

Table 7.1 Summary of previously published prevalence of T2DM prediction models (undiagnosed) and their risk factors (cont.)

Author	Population	Significant variables included in the final prediction model	Definition of undiagnosed diabetes	Sample size	Age group	Measure performance
Keesukphan <i>et al.</i> (2007)	Thailand	Age, BMI, history of hypertension	75-g OGTT carried out based on WHO criteria	N = 429	38 - 58	Sensitivity = 0.968 and Specificity = 0.24
Chaturvedi <i>et al.</i> (2008)	India	Age, blood pressure, waist circumference, family history of diabetes	Undiagnosed diabetes defined FPG $\geq$ 126 mg/dL ($\geq$ 7.0 mmol/L) but were not aware of their glycaemic status	N = 4,044 (1991-1994) N = 10,566 (2001 - 2003)	36 - 64	Sensitivity = 0.73 Specificity = 0.56
Heikes <i>et al.</i> (2008)	United States	Age, waist circumference, history of gestational diabetes, family history of diabetes, ethnicity, high blood pressure, weight, height, exercise, parental diabetes	Undiagnosed diabetes based on FPG and/or 2-h OGTT and absence of having been told that individuals have diabetes	N = 6,620	40 - 74	AUC = 0.85 Sensitivity = 0.88 Specificity = 0.75
Bang <i>et al.</i> (2009)	United States	Age, sex, family history of diabetes, history of hypertension, obesity (BMI or waist circumference), physical activity	Undiagnosed diabetes defined as FPG $\geq$ 7.0 mmol/L ($\geq$ 126 mg/dL)	N = 5,258 (1999-2004) N = 19,728 (2005-2006)	$\geq$ 20	AUC = 0.79 AUC = 0.74
Borrell <i>et al.</i> (2007)	United States	Age, sex, ethnicity, family history of diabetes, self-reported hypertension, hyper-cholesterolaemia, periodontal disease	FPG $\geq$ 126 mg/dL	4,830	$\geq$ 20	
Gao <i>et al.</i> (2010)	China	Age, waist circumference, family history of diabetes	Based on WHO/IDF 2006 criteria for individuals without known diabetes, FPG $\geq$ 7.0 mmol/L and/or PG $\geq$ 11.1 mmol/L	N = 1,986 N = 4,336	20 - 74	AUC = 0.673 Men: HL $\chi^2 = 6.23, p = 0.62$ Female: HL $\chi^2 = 10.36, p = 0.24$
Gray <i>et al.</i> (2010)	United Kingdom	Age, gender, ethnicity, first-degree family history of diabetes, antihypertensive therapy or history of hypertension, waist circumference, BMI	Diagnosed based on WHO criteria with FPG $\geq$ 7 mmol/L and/or 2-h PG $\geq$ 11.1 mmol/L. IFG defined as FPG between 6.1 and 6.9 mmol/L	N = 6,186 N = 3,171	40 - 75	AUC = 0.72
Bernabe-Ortiz <i>et al.</i> (2016)	Peru	Age, diabetes in first-degree relatives, and waist circumference.	Undiagnosed diabetes defined as FPG $\geq$ 7.0 mmol/L ($\geq$ 126 mg/dL)	N = 2,472 N = 2,945	$\geq$ 35	AUC = 0.68 Sensitivity = 0.70 Specificity = 0.59
Heianza <i>et al.</i> (2013)	Japan	Age, sex, family history of diabetes, smoking, BMI, and hypertension	Undiagnosed diabetes defined as FPG $\geq$ 7.0 mmol/L (or HbA1c $\geq$ 6.5%)	N = 33,335 N = 7,477	18 - 88	AUC = 0.806 Sensitivity = 0.727 Specificity = 0.681

Table 7.1 Summary of previously published prevalence of T2DM prediction models (undiagnosed) and their risk factors (cont.)

Author	Population	Significant variables included in the final prediction model	Definition of undiagnosed diabetes	Sample size	Age group	Measure performance
Xie <i>et al.</i> (2010)	China (2000 - 2001)	Men: waist circumference, age Women: waist/ hip ratio, age	Participants without previous diagnosis of diabetes were categorised according to the ADA diagnostic criteria: undiagnosed diabetes (FPG $\geq$ 7.0 mmol/L)	15,540	35 - 74	Men: sensitivity = 0.59; specificity = 0.63; women: sensitivity = 0.61; specificity = 0.71.
De Sousa <i>et al.</i> (2009)	Brazil	Age, BMI, known hypertension	FPG > 126 mg/dL (7.0 mmol/L), that is, provisional diagnosis of diabetes according to ADA criteria, classified as T2DM patients	1,224	$\geq$ 35	AUC = 0.772
				467		AUC = 0.720
Ko <i>et al.</i> (2010)	Hong Kong	Age, sex, BMI, hypertension, dyslipidaemia, family history of diabetes, gestational diabetes	75-g OGTT based on WHO 1998 criteria (FPG $\geq$ 7.0 mmol/L and/or 2-h PG $\geq$ 11.1 mmol/L)	N = 2,448, n = 3,734, n = 1,513	27 - 47	AUC = 0.735
Tabaei and Herman (2002)	Egypt (1992 - 1993)	Age, random plasma glucose, postprandial time, sex, BMI	Undiagnosed diabetes defined based on FPG $\geq$ 126 mg/dL and/or 2-h PG $\geq$ 200 mg/dL.	N = 1,023 N = 1,065 US subjects	(31 - 59) (31 - 61)	Sensitivity = 0.65 Specificity = 0.96 HL $\chi^2 = 5.27, p = 0.73$
Barengo <i>et al.</i> (2017)	North Colombia	Age, waist circumference, using hypertension medication, family history of diabetes	Undiagnosed diabetes defined based on FPG $\geq$ 126 mg/dL and/or 2-h PG $\geq$ 200 mg/dL.	N = 553	18 - 74	AUC = 0.74 Sensitivity = 0.73 Specificity = 0.67
Chien <i>et al.</i> (2010)	Taiwan Chinese ethnic	Age, sex, waist circumference, BMI, family history, hypertension	HbA_{1c} test based on WHO criteria, if $HbA_{1c} \geq 7.0\%$ indicates diabetes	N = 17,773	$\geq$ 30	AUC = 0.770 Sensitivity = 0.76 to 0.96 Specificity = 0.39 to 0.63
Heltberg <i>et al.</i> (2018)	Denmark	BMI, age, antihypertensive medication, cardiovascular disease, fitness, waist circumference.	HbA_{1c} test based on WHO criteria, if $HbA_{1c} \geq 6.5\%$ indicates diabetes	N = 21,205	$\geq$ 20	AUC = 0.771
Pramono <i>et al.</i> (2010)	Indonesia	Age, obesity, central obesity, hypertension, and smoking habit	Undiagnosed diabetes defined based on 2-h PG $\geq$ 200 mg/dL.	N = 18,956	$\geq$ 18	AUC = 0.64

Table 7.1 Summary of previously published prevalence of diabetes prediction models (undiagnosed) and their risk factors (cont.)

Author	Population	Significant variables included in the final prediction model	Definition of undiagnosed diabetes	Sample size	Age group	Measure performance
Al-Lawati and Tuomilehto (2007)	Oman (1991) (2001)	Age, waist circumference, BMI, family history of diabetes, hypertension	Diagnosed diabetes based on WHO 1998 criteria OGTT (FPG 11.1 mmol/l 2-h post 75-g glucose load)	N = 4,881 N = 1,432	≥ 20	AUC = 0.83 AUC = 0.76
Al Khalaf <i>et al.</i> (2010)	Kuwait (2007)	Age, waist circumference, blood pressure medication, diabetes in sibling	ADA 2003 criteria. If FPG was ≥ 7.0 mmol/L or random glucose was ≥ 11.1 mmol/L, participants were classified as having newly diagnosed diabetes	562	≥ 20	AUC = 0.82 Sensitivity = 0.87 Specificity = 0.64
Memish <i>et al.</i> (2015)	Saudi Arabia (2009)	Age, history of gestational diabetes, smoking, family history of diabetes, and central obesity	OGTT (fasting plasma glucose ≥ 7.0 mmol/L or 2-h post-load glucose ≥ 11.1 mmol/L)	1,485	≥ 20	Sensitivity = 0.766 Specificity = 0.521
Sulaiman <i>et al.</i> (2018)	UAE (2013) Sharjah and North Emirates	Age, family history of diabetes, hypertension, BMI, and waist-to-hip ratio	HbA _{1c} test based on WHO criteria, if HbA _{1c} ≥ 6.5% indicates diabetes	872	≥ 18	AUC = 0.82 Sensitivity = 0.75 Specificity = 0.70

Table 7.2 Summary of previously published prevalence of diabetes and prediabetes prediction models (undiagnosed) and their risk factors

Author	Population	Significant variables included in the final prediction model	Definition of undiagnosed diabetes	Sample size	Age group	Measure performance
Xin <i>et al.</i> (2010)	China	Age, body mass index, waist/hip ratio (WHR), hypertension, family history of diabetes.	Fasting plasma glucose (FPG) and FPG ≥ 5.6 mmol/L had a 75-g oral glucose tolerance test.	N = 2,826	≥ 35	Sensitivity = 0.653 Specificity = 0.725 AUC = 0.689
Heikes <i>et al.</i> (2008b)	United States	Age, waist circumference, weight, height, family history, exercise	Fasting plasma glucose (FPG) and oral glucose tolerance test.	N = 7,092	≥ 20	Sensitivity = 0.75 Specificity = 0.65 AUC = 0.75
Gray <i>et al.</i> (2010)	United Kingdom Multi-ethnic	Age, ethnicity, sex, first-degree family history, hypertension or antihypertensive, waist circumference	Individuals without known diabetes and with FPG ≥ 7.0 mmol/L or 2-h PG ≥ 11.1 mmol/L IFG: FPG = 6.1 to 6.9 mmol/L or 2-h PG = 7.8 to 11 mmol/L	N = 6,186 N = 3,171	40 - 75	AUC = 0.72 Sensitivity = 0.81 Specificity = 0.45
Handlos <i>et al.</i> (2013)	MENA Saudi Arabia UAE Algeria	Age, BMI, family history, gestational diabetes, ethnicity	HbA _{1c} ≥ 42 mmol/mol (≥ 6.0%)	N = 6588	30 - 75	AUC = 0.67 – 0.70 Sensitivity = 0.75 Specificity = 0.50 – 0.57

The aim of this chapter is to explore whether education level, socio-economic and healthcare access factors are predictive of undiagnosed diabetes and to develop a prediction tool to enable screening of individuals at high risk of the condition. The chapter addresses the following research questions:

- Do socio-economic and health care factors contribute to prediction models of undiagnosed diabetes and high-risk of diabetes in the whole population excluding those with diagnosed diabetes?
- Which socio-economic factors should be included in a prediction risk score to identify individuals who are at risk of having undiagnosed T2DM or Pre-T2DM?
- Can I use a prediction risk score model including socio-economic and health care access factors to identify high risk individuals with undiagnosed T2DM or pre-T2DM?

7.3 Methods

The analysis addresses the research questions listed above in that it aims to develop a prediction model for undiagnosed T2DM and pre-T2DM among the whole Saudi population (excluding those individuals with diagnosed diabetes). It develops separate prediction models based on two related outcome measures. The first considers two categories: undiagnosed T2DM (individuals who had no previous diagnosis of T2DM, including status unknown, and had $\text{HbA1c} \geq 6.5\%$) and normal (individuals with no previous diagnosis of T2DM, including status unknown, and had $\text{HbA1c} < 6.5\%$). The second outcome measure will use a broader definition for affected individuals that includes those with pre-T2DM. This considers two categories: individuals with no previous diagnosis of T2DM or pre-T2DM, including status unknown, and $\text{HbA1c} \geq 5.7\%$, again compared to normal ($\text{HbA1c} < 5.7\%$). Due to the nature of outcome measures, the prediction models were developed using binary logistic regression based on a similar dataset to that used in previous chapters.

7.3.1 Development of prediction model

The study used categorical variables without interaction terms to develop straightforward and relevant prediction models that can be used to generate a risk score appropriate for screening. To develop a risk score model, all predictors that are significantly associated with the main outcome variables in bivariate analysis must be included. I developed the final prediction model by selecting from different prototype models using e.g. AIC and BIC, and applying performance measures such as AUC, sensitivity, and specificity to construct an effective and fitted prediction model.

Two common strategies have been used to choose the most important predictors for the outcome (Pavlou et al., 2015). In the first approach, univariable screening, the association between each predictor and main outcome is examined individually and only statistically significant predictors are included in the final model. The second class of approach uses a stepwise model for selection of the most significant predictors, such as forward selection and backward elimination. The model selection procedure is a systematic process for selecting and building the model by adding or deleting predictor variables. Backward elimination starts with the most complex model and then deletes the least significant terms sequentially; in contrast, a forward procedure adds terms sequentially to a simple model.

In this way, all predictors that do not meet the pre-defined threshold for statistical significance are deleted from the prediction model. However, both approaches have the disadvantage that minor changes in the predictor selection procedure or in the sample could influence the final choice of predictors. Hence, in this study I incorporated other measures to develop the models and applied two alternative measures to quantify and estimate their quality.

The two measures were the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) – also known as the Schwarz information criterion (SIC) (Kuha, 2004). The AIC is an estimate based on extension of the Maximum Likelihood function including a number of parameters. The BIC/SIC is an estimate function of the posterior probability of the model. Using such measures, the best model is the one with the lowest AIC or BIC value. The disadvantage of using either AIC or BIC is that there is a risk of overfitting the model as both measures consider the parameter number. The AIC and BIC can be defined based on the following equations:

$$AIC = -2 \times \ln(L) + 2 \times k \quad (7.1)$$

$$BIC = -2 \times \ln(L) + \ln(N) \times k \quad (7.2)$$

Where k is the number of estimated parameters, N is the data sample size, and L is the maximised value of the likelihood function of the model. However, due to the importance of developing a model with high sensitivity, I also compared models based on discrimination and calibration approaches such as sensitivity, specificity, the area under the receiver-operating characteristic curve (AUC) (Lalkhen and McCluskey, 2008), the Youden index and the Hosmer-Lemeshow “goodness of fit” statistic. The following table describes the confusion matrix based on sensitivity, specificity, and other measures:

Table 7.3 Confusion Matrix/ Classification Algorithm for clinical prediction models

TEST SCORE		The Truth / Gold Standard		
		Has the disease	Does not have the disease	
	POSITIVE (+ve)	True Positive (TP)	False Positive (FP)	Positive Predictive Value (PPV)
	NEGATIVE (-ve)	False Negative (FN)	True Negative (TN)	Negative Predictive Value (NPV)
		Sensitivity	Specificity	

The area under the receiver-operating characteristic curve (AUC) is a graph of sensitivity versus specificity for various cut-off definitions of a positive diagnostic test result. The AUC value ranges from 0 to 1, where a perfect test has an AUC value of 1. A comparison of the different models based on these measures allows selection of the best prediction model. I applied the binary logistic regression approach to build prediction models for predicting those individuals at high risk of undiagnosed T2DM and Pre-T2DM, and to identify the most important predictors for the risk score model (Steyerberg, 2008).

The probability of having undiagnosed diabetes was modelled by log-odds or logistic:

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni} \quad (7.1)$$

Where x_{ni} 's: model predictors, the β_n 's: the regression coefficients. The outcome variable followed binomial distribution and the probability can be re-written as:

$$p(\text{diabetes}) = \frac{e^{(\beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni})}}{1 + e^{(\beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni})}} \quad (7.2)$$

The reference values for each categorical explanatory variable were determined in the final model, and covariates were checked so as not to omit any incorrectly. In addition to above steps, the results were used to create a weighted scoring system by rounding up the regression coefficients/odds ratios in the final model to the nearest integer (Bang et al., 2009). Adding these scores then gives a total score for each individual. Such a point-based model may have a useful role in classifying individuals according to their risk of disease so that those who are at highest risk can then be offered further testing in a clinical environment. Sensitivity analysis will be carried out to test the robustness of the final model by using the available data and ignoring any missing information for participants. I will run the final prediction model in two situations: 1) an analysis that consider only individuals with complete case records; and 2) an analysis including an extra category for each variable to include missing values, allowing inclusion of all participants in the model.

7.3.2 Validation of prediction model

After developing a prediction model, it is important to validate it. The most common strategies used to evaluate model performance are internal validation, external validation and use of performance measures, all of which I will discuss in the following sections. Internal validation of a prediction model can be achieved by applying one of four strategies: apparent validation; the split sample approach (if the sample is large enough); cross-validation; or bootstrapping (Steyerberg et al., 2001, Collins et al., 2011). The most appropriate strategy for internal validation generally involves resampling of the data (bootstrapping) unless the sample size is large, in which case the split-sample strategy will be most appropriate (Efron and Tibshirani, 1994). One drawback of the *split-sample* strategy is that it underestimates model performance and provides unstable estimates because only half of the sample is used to build the model.

In contrast, *bootstrap re-sampling* is more stable and provides unbiased estimates of model performance. Compared to bootstrap resampling, cross-validation may not accurately estimate model performance, especially if the number of repetitions is low. This approach is applied with 50% or 10% of the sample as a test part. Then one part of the data is used to estimate the model and assess the independent part. After that the performance average is the result repeated over 2 or 10 repetitions (Steyerberg, 2008).

Irrespective of the internal validation method adopted, external validation is essential to generalise the applicability of prediction model. External validation uses historical samples (temporally differentiated), another cohort, or samples from a different geographical location (spatially differentiated) in the prediction model (Steyerberg, 2008, Altman et al., 2009). However, although I requested additional datasets from Saudi Arabia and Oman, I did not receive any of them, and I was unable to apply external validation to the models in this study.

I therefore decide to use the bootstrapping approach to assess performance. The purpose of internal validation is to estimate the accuracy of the model and to determine the quality of the prediction risk scores based on the ability to classify individuals correctly in terms of their diabetes risk. In other words, to determine whether its predictive power is as good on a random sample of new patients as it was on the training set. Most previous studies of this type report that the bootstrapping approach is the most efficient validation technique (Efron and Tibshirani, 1994), and that it provides accurate estimates of model performance compared to other internal validation methods (Steyerberg et al., 2001). Importantly, the bootstrapping approach can be used when no external cohort dataset of patients is accessible or available.

Although I also used the split-sample procedure, I mainly used the bootstrap method to perform internal validation of the prediction model, and to estimate the model variability. The bootstrap sample is a simulation of the original population sample constructed built by random sampling with replacement, followed by estimation of one parameter from each sample such as the area under the receiver operating characteristic (ROC) curve. The standard errors and confidence intervals were extracted based on 100 bootstrap samples (Steyerberg et al., 2001). The approach involved first generating the bootstrap sample from the original sample using replacements (Efron and Tibshirani, 1994) and subjecting it to the prediction model. Logistic regression was then used to predict the risk of diabetes among individuals and construct a model for the bootstrap sample. Finally, the results of multiple logistic models from the original and bootstrap data were compared and analysed using standard validation measures (Efron and Tibshirani, 1994, Steyerberg, 2008, Steyerberg et al., 2010).

7.3.3 Evaluation of prediction model performance

After developing and validating the prediction model, it is next important to evaluate its performance. Suitable measures include the proportion of high-risk undiagnosed individuals; other performance metrics such as discrimination; and calibration approaches such as sensitivity, specificity, and the area under the receiver-operating characteristic curve (AUC) as a discrimination statistic. The area under the receiver operating characteristic (ROC) curve (AUC) was used to compare the discrimination ability of the two new prediction models between the original sample and the bootstrap sample.

It was also used to test how well the model parameter can discriminate between the group at high risk of undiagnosed diabetes and the normal group, and to evaluate model performance by estimating the AUC for both samples. In addition, the Hosmer-Lemeshow “goodness of fit” test was used as a calibration performance measure to show the closeness of the match between predicted and observed frequencies. Finally, I compared model performance for both original sample and the bootstrap sample, where the original sample shows model development performance and the bootstrap sample gives an estimate of the model validation performance (Steyerberg et al., 2001). SHIS data was used for model development and validation throughout.

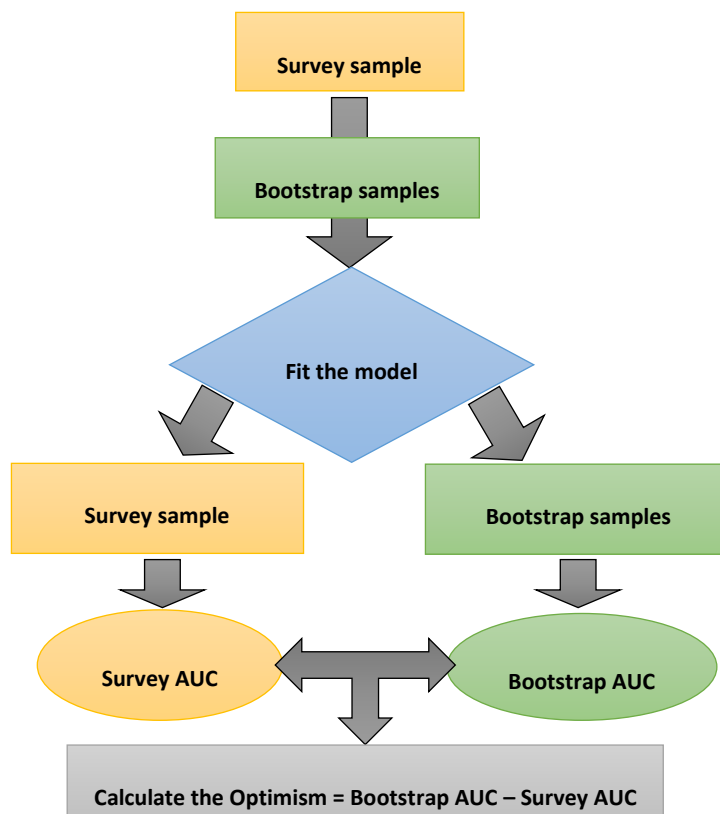


Figure 7.1 Flow chart of prediction model validation Using Bootstrap Method

Two main outcomes will be investigated in this chapter. Diagnosed diabetes is defined as individuals who had clinically confirmed T2DM irrespective of their actual HbA1c level. The first outcome of interest here is undiagnosed diabetes defined as individuals with no previous diagnosis of diabetes and HbA1c level $\geq 6.5\%$. The second outcome is a combined category that includes individuals with undiagnosed diabetes and pre-diabetes. This is defined as individuals with no previous diagnosis of diabetes and a lower threshold for HbA1c level of ($\geq 5.7\% - <6.5\%$). The following table describes the predictor and outcome variables used for developing the prediction models. The relationship between the outcome measures and study predictors used the classical Chi-square “Goodness-of-fit” test. Binary logistic regression was used to address all the statistical analyses, and carried out using STATA (Stata Corp. Inc, TX, USA).

Table 7.4 Outcome and explanatory variables codes

Variables	Categories
Undiagnosed T2DM	0 = Normal, 1 = Undiagnosed
Undiagnosed T2DM and Pre-T2DM	0 = Normal, 1 = Undiagnosed of T2DM and Pre-T2DM
Sex	1= Male, 2= Female
Age	1 = 15 - 24, 2 = 25 – 34, 3 = 35 – 44, 4 = 45 – 54, 5 = 55 – 64, 6 = 65+
Marital status	1 = Currently Married 2 = Not currently married: includes never married. 3 = Separated/divorced, and Widowed.
Education level	1 = Low education level 2 = Middle education level 3 = High education level
Employment status	1 = Government employee; 2 = Non-government employee; 3 = Unemployed; 4 = Retired; 5 = Homeworker; 6 = Self-employed; 7 = Student.
Monthly household income	1 = Lowest: Less than 3,000 Riyals; 3,000 Riyals to less than 5,000 2 = Low: 5,000 Riyals to less than 7,000; 7,000 Riyals to less than 10,000 3 = Middle: 10,000 Riyals to less than 15,000; 15,000 Riyals to less than 20,000; 4 = High: 20,000 Riyals to less than 30,000. 5 = Highest: 30,000 Riyals or more
Body Mass Index (BMI)	1 = Underweight or normal; 2 = Overweight; 3 = Obese
Hypertension (HTN)	1 = Yes; 0 = No
Routine check-up	1 = Yes; 0 = No
Time between home and healthcare facility	1 = Less than one hour; 2 = More than one hour or equal
Hospital type	1 = Public; 2 = Private; 3 = Other
Health care provider number	0 = No; 1 = Yes, there is one place; 2 = Yes, there are more than one place
Region location	1 = Central region; 2 = Eastern region; 3 = Western region; 4 = Northern region; 5 = Southern region

7.4 Results

7.4.1 Prediction model of undiagnosed diabetes

The previous chapter compared the association between education level and diagnosed or undiagnosed diabetes to investigate whether there were differences between the two groups. In this section, I focus on the association between socio-economic factors and undiagnosed diabetes. Hence, I compare the group with undiagnosed diabetes and the normal group considering the whole population in general rather than diagnosed diabetes cases (individuals who know that they have diabetes). The prediction model I develop here is therefore based on undiagnosed diabetes in comparison to the normal group. **Table 7.5** shows the basic demographic, socio-economic, lifestyle, health care access and administrative regional factors for the study participants. The percentage of distribution between undiagnosed diabetes and normal for each group of study variables is presented. Of the study participants, 399 cases (10.5%) had undiagnosed T2DM and 3,404 cases (89.5%) were unaffected. Participants with undiagnosed T2DM were more likely to be older, male, live in the Southern region, and have higher BMI and blood pressure.

Regarding the distribution of socio-economic factors, participants with undiagnosed T2DM were more likely to be low educated, retired, and unmarried. The proportion of undiagnosed T2DM for females (9.6%) is lower than in males (11.5%). However, this variable was not significant associated with undiagnosed T2DM. Twenty five percent of individuals aged 55 to 64 or older had undiagnosed T2DM significantly higher than for younger age groups. Twice the proportion of individuals with low educational status had undiagnosed diabetes compared to the higher education group (16.1% vs 8.4%) respectively. This factor showed a significant association with undiagnosed T2DM, consistent with previous studies and reflecting factors such as greater awareness of the disease and healthier lifestyles in better-educated individuals. In keeping with earlier descriptive results, a higher proportion of retired individuals had undiagnosed T2DM (23.0%) in contrast to the student group with the lowest proportion (5.5%), and employment status was significantly associated with undiagnosed T2DM. This latter observation is as expected, since retired individuals are more likely to be in older age groups, and students' group would be in the youngest age category.

There was a higher proportion of undiagnosed T2DM among high-income individuals, however the income variable did not show significant association. This finding is inconsistent with past studies showing that low-income individuals have an increased incidence of undiagnosed T2DM and might reflect differences in the socio-economic context of the Saudi population. Marital status showed a significant relationship with undiagnosed diabetes, with a higher proportion of affected individuals in the unmarried category. A high proportion of obese and overweight individuals have undiagnosed T2DM and this relationship is significant. This result is expected given that obesity is one of the main predictors of T2DM. Individuals living in the Southern region have the highest proportion of undiagnosed T2DM; individuals who lived in Western region have the lowest. This factor showed a highly significant association. Hypertensive individuals were more likely to have undiagnosed diabetes. Unexpectedly, healthcare access factors were not significantly associated with undiagnosed T2DM.

Regarding the prediction model, further tests were performed by adding other socio-economic factors to the model individually. The results indicated that age strongly influenced education level and employment status. However, household income was significantly associated in some groups irrespective of age. Education level was significantly associated with undiagnosed diabetes; however, adding education level to the prediction model gave higher AIC and BIC values (See Appendix E, Table E.2). With regard to region, all three individual socio-economic factors were significant in the model if region was included, or if both region and age were excluded from the analysis. I chose to carry out the analysis in four ways: excluding region; excluding age, excluding age and region; and including age and region as eight groups (five regions and three age groups) with four groups of household income. Excluding age gave better sensitivity but lower specificity; however, excluding the region brings better specificity. The results are summarised in Appendix E (summary of all models).

Table 7.5 Characteristics of survey participants with undiagnosed T2DM

Variables	Categories N (%)	Total 3,803	Undiagnosed 399 (10.5)	Normal 3,404 (89.5)	p
Sex	Female	2,060	198 (9.6)	1,862 (90.4)	0.054
	Male	1,743	201 (11.5)	1,542 (88.5)	
Age	15 - 24	885	56 (6.3)	829 (93.7)	<0.001
	25 - 34	997	70 (7.0)	927 (93.0)	
	35 - 44	892	89 (10.0)	803 (90.0)	
	45 - 54	540	84 (15.6)	456 (84.4)	
	55 - 64	229	57 (24.9)	172 (75.1)	
	65 +	260	43 (16.5)	217 (83.5)	
Education level	Low education	777	125 (16.1)	652 (83.9)	<0.001
	Middle education	1,028	106 (10.3)	922 (89.7)	
	High education	1,995	168 (8.4)	1,827 (91.6)	
Employment status	Government employee	1,204	129 (10.7)	1,075 (89.3)	<0.001
	Non-government employee	148	16 (10.8)	132 (89.2)	
	Unemployed	434	45 (10.4)	389 (89.6)	
	Retired	204	47 (23.0)	157 (77.0)	
	Homework	981	106 (10.8)	875 (89.2)	
	Self-employed	113	16 (14.2)	97 (85.8)	
	Student	704	39 (5.5)	665 (94.5)	
Monthly Household income	Lowest/ poorest	1,545	157 (10.2)	1,388 (89.8)	0.298
	Low / poor	612	56 (9.2)	556 (90.8)	
	Middle	510	42 (8.2)	468 (91.8)	
	High / rich	242	30 (12.4)	212 (87.6)	
	Highest / richest	188	23 (12.2)	165 (87.8)	
Marital status	Currently married	2,449	287 (11.7)	2,162 (88.3)	<0.001
	Never married	1,076	72 (6.7)	1,004 (93.3)	
	Divorced, Separated, and widowed	265	38 (14.3)	227 (85.7)	
Routine check-up	No	2,573	236 (9.2)	2,337 (90.8)	<0.001
	Yes	984	133 (13.5)	851 (86.5)	

Table 7.5 Characteristics of survey participants with undiagnosed T2DM (cont.)

Variables	Categories N (%)	Total 3,803	Undiagnosed 399 (10.5)	Normal 3,404 (89.5)	p
Time between home and healthcare facility	Less than one hour	2,416	270 (11.2)	2,146 (88.8)	0.525
	More than one hour or equal	244	24 (9.8)	220 (90.2)	
Hospital type	Public	2,250	260 (11.6)	1,990 (88.4)	0.117
	Private	367	29 (7.9)	338 (92.1)	
	Other	244	27 (11.1)	217 (88.9)	
Health care provider number	No	1,118	108 (9.7)	1,010 (90.3)	0.311
	Yes, there is one place	1,190	122 (10.3)	1,068 (89.8)	
	Yes, there are more than one place	1,339	154 (11.5)	1,185 (88.5)	
Region location	Central region	1,356	144 (10.6)	1,212 (89.4)	<0.001
	Eastern region	126	11 (8.7)	115 (91.3)	
	Western region	738	50 (6.8)	688 (93.2)	
	Northern region	559	57 (10.2)	502 (89.8)	
	Southern region	1,024	137 (13.4)	887 (86.6)	
Body mass index	Underweight or normal	1,212	90 (7.4)	1,122 (92.6)	<0.001
	Overweight	1,204	131 (10.9)	1,073 (89.1)	
	Obese	1,387	178 (12.8)	1,209 (87.2)	
Hypertension	Yes	815	114 (14.0)	701 (86.0)	<0.001
	No	2,988	285 (9.5)	2,703 (89.5)	

Development of the prediction model

- **Comparison between various prediction models**

Before further refinement of the prediction model, I compared the four initial models described above on their ability to predict undiagnosed T2DM cases in the general population using four main predictors. I assessed these using a combination of approaches including BIC and AIC selection measures, AUC, sensitivity, specificity, Hosmer-Lemeshow p-value and chi-square tests. These results are summarised in **Table 7.6**. The following four prediction models have been assessed based on different performance and model selection measures in order to identify the best prediction model with undiagnosed T2DM.

Overall, four main predictors had significant associations with undiagnosed T2DM. **Model 1** used three main predictors for undiagnosed T2DM: age, routine check-up, and region. This model had the lowest BIC value compared to the other models and a very good area under the curve value. However, the sensitivity was low in comparison to the other models and was not further developed. **Model 2** used four predictors: age, routine check-up, region, and body mass index. This model was excluded based on its low sensitivity, greater BIC value and low Hosmer-Lemeshow p-value. **Models 3 and 4** used different combinations of predictors; both had similar or higher sensitivity values to the first two. However, model 3 had higher AUC and specificity values than the fourth model and was used for further refinement.

Table 7.6 Summary statistics for different prediction models based on model selection and performance measures

Model	predictors	Observations	BIC	AIC	AUC [95%CI]	Sensitivity	Specificity	C	Youden	HL-p	HL-c
1	Age + Check-up + Region	3,557	2349.23	2281.29	0.6606 [0.6305, 0.6907]	0.55	0.71	-2.00	0.260	0.4939	7.40
2	Age + Check-up + Region +BMI	3,557	2360.17	2279.87	0.6674 [0.6378, 0.6971]	0.54	0.72	-2.04	0.263	0.2484	10.24
3	Age + Region +BMI	3,557	2357.99	2283.87	0.6638 [0.6339, 0.6937]	0.64	0.62	-2.18	0.263	0.3034	9.48
4	Age + Check-up +BMI	3,557	2349.68	2294.09	0.6524 [0.6230, 0.6818]	0.64	0.59	-2.31	0.229	0.6004	6.42

BIC: Bayesian Information Criterion; **AIC:** Akaike Information Criterion; **AUC:** Area Under the Curve; **C:** Optimal Cut-off value; **Y:** Youden-index: Sensitivity + Specificity - 1 ; **HL-p:** Hosmer-Lemeshow P value; **HL-c:** Hosmer-Lemeshow Chi-square “high Hosmer-Lemeshow p-value and low chi-square values indicated a goodness-of-fit”

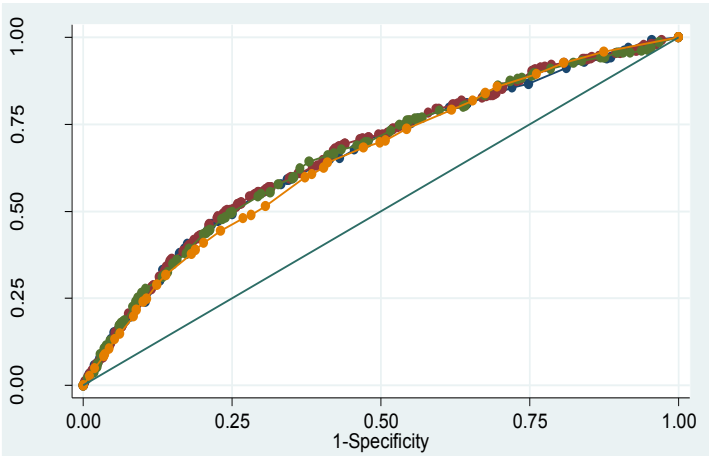


Figure 7.2 Receiver-operating characteristic (ROC) curves for various prediction models in the presented study.⁵⁸

⁵⁸ **BLUE, Model one:** AGE+CHECK-UP+REGION (AUC: 0.6606); **RED, Model two:** AGE+CHECK-UP+REGION+BMI (AUC: 0.6674); **GREEN, Model three:** AGE+REGION+BMI (AUC: 0.6638); **ORANGE, Model four:** AGE+CHECK-UP+BMI (AUC: 0.6524); **BLACK, REFERENCE** (AUC: 0.500)

- **Constructing the prediction model**

The results of the multivariate logistic regression model are listed in **Table 7.7**. Effects due to gender, socio-economic factors and health care access factors were not statistically significant and were excluded in subsequent analyses, leaving age, region, and BMI as the most appropriate predictors of undiagnosed T2DM. The following presents the final regression model derived from the dataset. Age, region, and body mass index are split into multiple categories to capture the individual risk score, giving an optimal cut-off point of 10.16. In the model, the main outcome is presence of undiagnosed T2DM: (Yes = undiagnosed T2DM, No = Normal). The predictor variables are age (six categories), BMI (three categories) and Region (five categories). The risk of an individual experiencing an event depends on combining the risk factors attributed to the three predictors. Then, the logistic regression model used is:

$$\text{Patient's risk of being undiagnosed of type two diabetes} = \frac{e^{(\text{patient's risk score})}}{(1 + e^{(\text{patient's risk score})})} \quad (7.3)$$

$$\text{where, patient's risk score} = \text{intercept} + (\beta_1 \times \text{age}) + (\beta_2 \times \text{region}) + (\beta_3 \times \text{BMI}) \quad (7.4)$$

In addition, $(\beta_1 \times \text{age}) + (\beta_2 \times \text{region}) + (\beta_3 \times \text{BMI})$ are estimated in order to calculate the risk score for each patient. For example, the risk score for a patient aged 45 years old with BMI ≥ 30 , who lived in Southern region would be calculated as:

$$y_{45, \text{bmi} \geq 30, \text{south}} = -2.797 + 0.847 + 0.372 + 0.266 = -1.312$$

$$\text{Therefore, the predicted risk for this patient would be: } \frac{e^{(-1.312)}}{(1 + e^{(-1.312)})} = 0.2122 \times 100 = 21.22\%$$

Age was significantly associated with undiagnosed T2DM, with individuals aged between 45 and 54 more likely to have undiagnosed T2DM than aged between 15 -24 years old (OR: 2.33). This effect has been observed previously as discussed above. Similarly, high body mass index showed significant association with undiagnosed T2DM, with obese individuals having a high likelihood of undiagnosed T2DM compared to those whose BMI was within the healthy range or below (OR: 1.45). Administrative region was also a significant predictor of undiagnosed T2DM, with individuals living in the Southern region more likely to be affected than those in the Central region (OR: 1.30). Unlike the previous two factors, this unexpected finding offers a new insight into the epidemiology of undiagnosed T2DM. In light of this, I extended the analysis using bivariate analysis in order to assess the association between an individual's socio-economic status and administrative region as a proxy of community/region socio-economic level.

Table 7.7 Risk factors and final development prediction model of undiagnosed T2DM

Risk factor	Categories	Odds Ratio [95% CI]	β	P-value	Score assigned
	Intercept		-2.797		
Age	15 - 24	1	Reference	--	0
	25 - 34	1.01 [0.69, 1.46]	0.007	0.969	1
	35 - 44	1.43 [1.00, 2.07]	0.363	0.049	4
	45 - 54	2.33 [1.61, 3.39]	0.847	<0.001	8
	55 - 64	4.34 [2.85, 6.61]	1.468	<0.001	15
	65+	2.69 [1.75, 4.16]	0.991	<0.001	10
BMI	Underweight or normal	1	Reference	--	0
	Overweight	1.28 [0.95, 1.71]	0.244	0.104	2
	Obese	1.45 [1.09, 1.93]	0.372	0.011	4
Region	Central region	1	Reference	--	0
	Eastern region	0.70 [0.37, 1.36]	-0.351	0.294	- 4
	Western region	0.59 [0.43, 0.84]	-0.512	0.003	- 5
	Northern region	0.92 [0.66, 1.28]	-0.087	0.606	-1
	Southern region	1.30 [1.01, 1.68]	0.266	0.041	3

Two possible predicted results are analysed in the confusion matrix presented in **Table 7.8**. A positive result implies an individual has diabetes; a negative result means they do not based on the cut-off point discussed above (≥ 10.16). The classification table was constructed using 3,803 individuals who were screened using the prediction model. Out of the 3,803 cases, the test predicted 1,549 cases of undiagnosed T2DM and 2,254 cases without diabetes. The clinical data shows that in fact 399 individuals in the sample have T2DM; the remaining 3,404 individuals do not have diabetes.

Table 7.8 Confusion matrix for undiagnosed T2DM

		True Disease Status (Gold standard = HbA1c)		
Test results (Predicted test)		Abnormal	Normal	Total
	Positive (+ve)	TP = 256	FP = 1,293	1,549
	Negative (-ve)	FN = 143	TN = 2,111	2,254
	Total	399	3,404	3,803
Abnormal: Undiagnosed (HbA1c $\geq$ 6.5%); Normal: (HbA1c < 6.5%); Positive: cut-point $\geq$ - 2.18; Negative: cut-point < -2.18				

The percentage of individuals with a positive screening test that truly have T2DM ranged from 14.7 % - 18.5%. However, the percentage of individuals with a negative screening test that truly do not have T2DM ranged from 92.6% - 94.6%. The positive and negative likelihood ratios measured the prediction model performance. **The accuracy** assesses how often is the classifier correct overall. This is given by: $(TP+TN)/total = (256+2,111)/3,803 = 0.6224 = 62.2\%$. The **misclassification rate**, in contrast, assesses how often it is wrong overall. This is given by: $(FP+FN)/total = (1,293+143)/3,803 = 0.3776 = 37.8\%$. A high-risk of undiagnosed diabetes can be defined as a predicted probability of greater than or equal to 0.1011; conversely, individuals are at low risk if their predicted probability is lower than 0.1011. In general, the predicted probabilities of undiagnosed T2DM increased among middle-aged and obese individuals.

Individuals who lived in the Eastern and Western regions were less likely to have undiagnosed T2DM, followed by individuals who lived in the Central and Northern regions. Individuals who lived in the Southern region have a high-risk of undiagnosed T2DM. The predicted probability of undiagnosed T2DM for individuals who lived in the Central region ranged from 5.8% (aged 15 -24 and BMI < 25) - 27.7% (aged 55 – 64 and BMI $\geq$ 30). There is a trend towards higher predicted probabilities of undiagnosed of T2DM among obese individuals and those aged 35 years old and over (11.3%, 17.1%, and 27.7%); and among middle-aged individuals who lived in the Eastern or Western regions and had high BMI ranging from 11% - 21.3%. I conducted internal validation of the prediction model using split-sample validation (further information in Appendix E) together with bootstrap methods and an assessment of the variability in model estimates. The performance of the validation model is very similar to that of the development model with a minimal increase in the AUC value to 0.665. Overall, the performance of the bootstrap sample is very similar to the original sample using a range of performance measures.

- **Performance measures of the prediction model**

Table 7.9 summarise the performance statistics of the final prediction model for undiagnosed T2DM. The Hosmer-Lemeshow test showed that the final model fitted well (p-value = 0.6401). The Saudi diabetes risk score model was constructed based on the coefficients used in the final prediction model. The risk model gave an AUC value of 0.67 (95%CI: 0.64, 0.69). The optimal cut-off point for undiagnosed diabetes was ≥ 2.62 using the Youden index; the optimal-empirical cut-off point was ≥ -2.18 (≥ 10.16). With this cut-off point, participants were categorised as at high risk of T2DM with sensitivity 0.64, specificity 0.62; a PPV of 0.6416, and an NPV of 0.6202. The accuracy of the prediction model is **0.622** (given by $(TP+TN)/total = (256+2111)/3,803$), while the misclassification rate is **0.378** (given by $(FP+FN)/total = (1293+143)/3,803$).

Table 7.9 Summary statistics for different prediction models based on model selection and performance measures

Models	observation	AUC	Sensitivity	Specificity	Optimal cut-point	Youden index	HL-c	HL-p
Original sample	3,803	0.664	0.64	0.62	-2.18	0.262	6.06	0.6401
Bootstrapping sample	3,803	0.665	0.64	0.62	-2.18	0.262	6.06	0.6401
HL- C: Hosmer-Lemeshow Chi-square; HL-p: Hosmer-Lemeshow p-value								

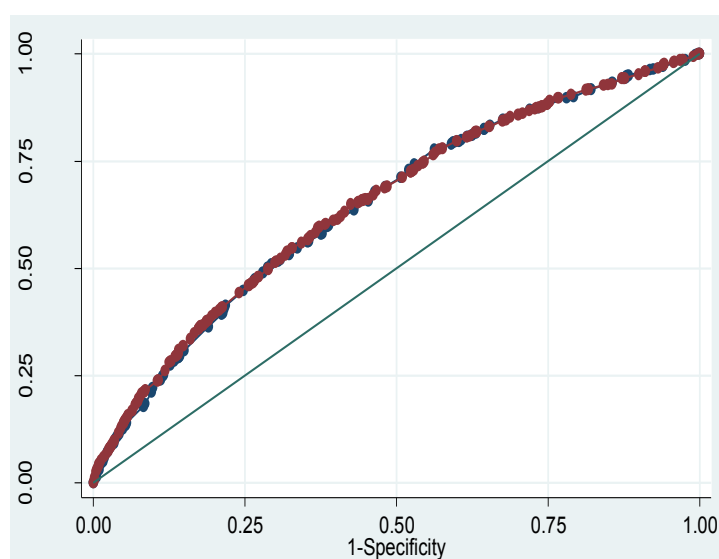


Figure 7.3 Receiver operating characteristics (ROC) curve showing performance of original sample and bootstrap sample for undiagnosed diabetes

Finally, the performance of the model was also assessed using discrimination measures as presented in **Table 7.10** and **Figure 7.4**. A calibration measure - the Hosmer-Lemeshow (HL) test - was used to compare between observed and predicted probabilities, and to assess the model as fitting or calibrating well. For this analysis, the data are usually divided into 10 equal groups. This allows formulation of a null hypothesis that there are no significant differences between observed and predicted values. Thus the, **null hypothesis (H0)** is that the observed values are closely matched with the expected values. The complementary **alternative hypothesis (H1)** is that the observed values are not matched with the expected values. **Table 7.10** below presents the results of the HL goodness-of-fit test which reveals that the observed values match closely to the expected values, and that the model fits the data well. The p-value is greater than 0.05 threshold, meaning that I can reject the alternative hypothesis. The Hosmer-Lemeshow statistic is calculated using the following formula:

$$\text{Hosmer} - \text{Lemeshow} = \sum_{q=1}^{10} \frac{(\text{Observed}(1) - \text{Expected}(1))^2}{\text{Expected}(1)} + \frac{(\text{Observed}(0) - \text{Expected}(0))^2}{\text{Expected}(0)} \quad (7.5)$$

Table 7.10 Observed and predicted T2DM status using the test score as a predictive variable in the groups defined by Hosmer-Lemeshow Goodness-of-fit test

Group/ quantile	Likelihood/ probability of being undiagnosed diabetes	Present (undiagnosed)		Absent (normal)		Total
		Observed (1)	Expected (1)	Observed (0)	Expected (0)	
1	< 0.0507	21	18.5	412	414.5	433
2	0.0507 - 0.0579	18	23.3	400	394.7	418
3	0.0579 - 0.0722	26	21.0	282	287.0	308
4	0.0722 - 0.0742	27	28.7	363	361.3	390
5	0.0742 - 0.0822	27	29.1	337	334.9	364
6	0.0822 - 0.1027	31	38.5	367	359.5	398
7	0.1027 - 0.1128	49	47.2	386	387.8	435
8	0.1128 - 0.1426	47	40.6	258	264.4	305
9	0.1426 - 0.1869	62	62.4	312	311.6	374
10	> 0.3336	91	89.6	287	288.4	378
Results	H-L Statistic = 6.06	Degree of freedom = 8		Significance (P-value) = 0.6401		

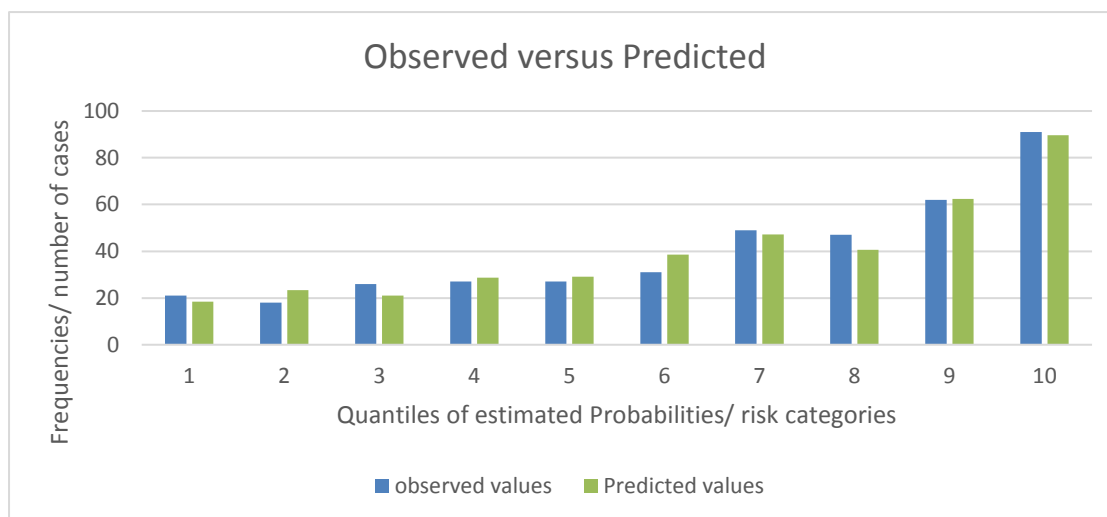


Figure 7.4 Calibration of the test score: observed and expected of undiagnosed T2DM

7.4.2 Prediction model of undiagnosed diabetes and pre-diabetes

The composite pre-diabetes and undiagnosed diabetes outcome showed a similar relationship with socio-economic, demographic factors and region to that discussed in the previous section for undiagnosed diabetes alone - **Table 7.11**. A significant exception was the observation that gender was now a significant risk factor, with a higher proportion of Saudi males displaying the composite phenotype (36.7%) compared to females (31.7%).

Table 7.11 Characteristics of survey participants with undiagnosed diabetes and prediabetes

Variables	Categories N (%)	Total 3,803	Undiagnosed ⁵⁹ 1,291 (33.9)	Normal 2,512 (66.1)	p
Sex	Female	2,060	652 (31.7)	1,408 (68.3)	0.001
	Male	1,743	639 (36.7)	1,104 (63.3)	
Age	15 – 24	885	213 (24.1)	672 (75.9)	<0.001
	25 – 34	997	257 (25.8)	740 (74.2)	
	35 – 44	892	326 (36.6)	566 (63.4)	
	45 – 54	540	237 (43.8)	303 (56.1)	
	55 – 64	229	127 (55.5)	102 (44.5)	
	65 +	260	131 (50.4)	129 (49.6)	
Education level	Low education	777	338 (43.5)	439 (56.5)	<0.001
	Middle education	1,028	340 (33.1)	688 (66.9)	
	High education	1,995	613 (30.7)	1,382 (69.3)	
Employment status	Government employee	1,204	422 (35.1)	782 (64.9)	<0.001
	Non-government employee	148	53 (35.8)	95 (64.2)	
	Unemployed	434	141 (32.5)	293 (67.5)	
	Retired	204	117 (57.4)	87 (42.6)	
	Homework	981	336 (34.3)	645 (65.7)	
	Self-employed	113	49 (43.4)	64 (56.6)	
	Student	704	167 (23.7)	537 (76.3)	
Monthly Household income	Lowest/ poorest	1,545	520 (33.7)	1,025 (66.3)	0.118
	Low / poor	612	191 (31.2)	421 (68.8)	
	Middle	510	167 (32.8)	343 (67.2)	
	High / rich	242	97 (40.1)	145 (59.9)	
	Highest / richest	188	70 (37.2)	118 (62.8)	
Marital status	Currently married	2,449	899 (36.7)	1,550 (63.3)	<0.001
	Never married	1,076	276 (25.7)	800 (74.3)	
	Divorced, Separated, and widowed	265	112 (42.3)	153 (57.7)	
Routine check-up	No	2,573	833 (32.4)	1,740 (67.6)	0.020
	Yes	984	359 (36.5)	625 (63.5)	

⁵⁹ Undiagnosed diabetes and prediabetes

Development of the prediction model

• Comparison between various prediction models

Before constructing the final prediction model, I compared a range of prediction models based on BIC, AIC selection measures, the area under the curve (AUC), sensitivity, specificity, and Hosmer-Lemeshow p-values to choose the best prediction model for identifying undiagnosed T2DM or pre-diabetic cases. **Table 7.12** and **Figure 7.5** shows the results of this analysis for six prediction models. Although all six models are well-fitted based on the Hosmer-Lemeshow p-value, three of the models displayed lower sensitivity values (**Model 3**: 0.56, **Model 4**: 0.59, and **Model 5**: 0.51). The remaining three models had higher sensitivity values and lower AIC values than other prediction models. The BIC value was lowest in models 2, 3 and 4. Overall, Model 2 performed best in this comparison as it displayed the highest Hosmer-Lemeshow p-value and lowest BIC value. As a result, I selected Model 2 as the basis for development of the second prediction model.

• Constructing prediction model

The following table presents the final prediction model for the composite undiagnosed T2DM/pre-diabetes outcome (**Table 7.13**). The main outcome is presence of undiagnosed T2DM or pre-diabetes: (Yes = undiagnosed T2DM or prediabetes, No = Normal). The predictor variables are age (six categories), BMI (three categories), region (five categories), gender (two categories) and hypertension (two categories). The risk of an individual experiencing an event depends on combining the risk factors attributed to the five predictors. The logistic regression model I used is:

$$\text{Patient's risk of being undiagnosed of pre diabetes} = \frac{e^{(\text{patient's risk score})}}{(1 + e^{(\text{patient's risk score})})} \quad (7.6)$$

where, $\text{patient's risk score} = \text{intercept} + (\beta_1 \times \text{age}) + (\beta_2 \times \text{region}) + (\beta_3 \times \text{BMI}) + (\beta_4 \times \text{Sex}) + (\beta_5 \times \text{HTN})$

In addition, $(\beta_1 \times \text{age}) + (\beta_2 \times \text{region}) + (\beta_3 \times \text{BMI}) + (\beta_4 \times \text{Sex}) + (\beta_5 \times \text{HTN})$ are estimated in order to calculate the risk score for each patient as follows: For example, the risk score for female and male patients aged 55 with BMI ≥ 30 , who had high blood pressure and lived in Western region would then be calculated as:

Female patient:

$$y_{\text{female}} = -1.359 + 1.444 + 0 + 0.508 + (-0.436) + (0.198) = 1.85 - 1.795 = 0.055$$

The predicted risk of female patient would be: $\frac{e^{(0.055)}}{(1 + e^{(0.055)})} = 0.5137 \times 100 = 51.37\%$

Male patient:

$$y_{male} = -1.359 + 1.444 + 0.247 + 0.508 + (-0.436) + (0.198) = 2.097 - 1.795 = 0.302$$

The predicted risk of male patient would be: $\frac{e^{(0.302)}}{1+e^{(0.302)}} = 0.5749 \times 100 = 57.49\%$

Age was significantly associated with undiagnosed T2DM/pre-diabetes. Individuals aged 35 – 44 had a higher likelihood of being affected than individuals aged 15 – 24 (OR: 1.54; 95%CI: 1.24, 1.91). Interestingly, gender was a good demographic predictor of undiagnosed T2DM/pre-diabetes as patients were more likely to be affected if they were male (OR: 1.68, 95%CI: 1.13, 1.49). Body mass index was also an important predictor of affected status, with high BMI individuals at significantly higher risk (OR: 1.68, 95%CI: 1.39, 1.49). Region was another significant predictor of undiagnosed T2DM/pre-diabetes. Individuals living in the Western region were less likely to be affected than those living in the Central region. Finally, hypertension significantly increased the risk of affected status.

Table 7.12 Summary statistics for different prediction models based on model selection and performance measures

Model	predictors	Observations	BIC	AIC	AUC	Sensitivity	Specificity	Cut-point	Youden index	HL-p
1	Age + Sex +Region +BMI+ HTN	3,803	4739.63	4652.22	0.6519	0.64	0.59	-0.701	0.230	0.2487
2	Age + Sex +Region +BMI	3,803	4736.34	4655.17	0.6511	0.65	0.58	-0.708	0.225	0.7907
3	Age + Sex +Region	3,803	4752.64	4683.96	0.6378	0.56	0.65	-0.640	0.216	0.4530
4	Age + Sex +BMI	3,803	4733.60	4677.41	0.6412	0.59	0.62	-0.672	0.204	0.8923
5	Age + Region +BMI	3,803	4741.59	4666.68	0.6474	0.51	0.71	-0.526	0.216	0.7301
6	Age + Sex +BMI+ HTN	3,803	4736.39	4673.96	0.6426	0.64	0.57	-0.772	0.211	0.5451

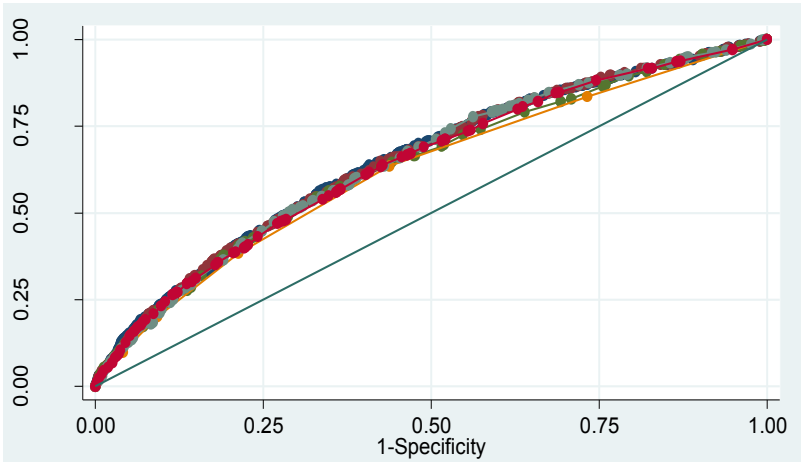


Figure 7.5 Receiver-operating characteristic curves for various prediction models in the presented study

Table 7.13 Risk factors and final development prediction model of undiagnosed T2DM and pre-diabetes

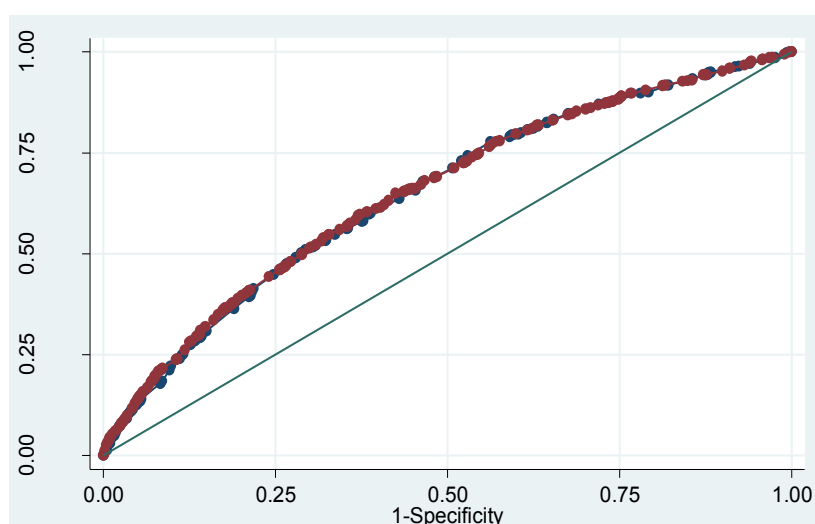
Risk factor	Categories	Odds Ratio [95% CI]	β	P-value	Score
	Intercept		-1.359		
Age	15 - 24	1	reference	--	0
	25 - 34	0.96 [0.78, 1.20]	-0.038	0.731	-1
	35 - 44	1.51 [1.22, 1.88]	0.413	0.000	4
	45 - 54	1.94 [1.52, 2.47]	0.662	0.000	7
	55 - 64	3.14 [2.28, 4.32]	1.144	0.000	12
	65+	2.64 [1.94, 3.58]	0.970	0.000	11
Sex	Female	1	reference	--	0
	Male	1.28 [1.11, 1.48]	0.247	0.001	3
BMI	Underweight or normal	1	reference	--	0
	Overweight	1.51 [1.25, 1.81]	0.409	0.000	4
	Obese	1.66 [1.38, 2.00]	0.508	0.000	5
Region	Central region	1	reference	--	0
	Eastern region	0.54 [0.35, 0.82]	-0.621	0.004	- 6
	Western region	0.65 [0.53, 0.79]	-0.436	0.000	- 4
	Northern region	0.79 [0.63, 0.98]	-0.239	0.032	- 3
	Southern region	1.01 [0.84, 1.20]	-0.007	0.938	-1
Hypertension	No	1	reference	--	0
	Yes	1.22 [1.02, 1.45]	0.198	0.026	

Two possible predicted results are analysed in the confusion matrix presented in **Table 7.14**, similar to that presented for the first prediction model in the previous section. The classification table was constructed using 3,803 individuals who were screened using the second precision model. Out of the 3,803 individuals, the test predicted 1,850 cases of undiagnosed T2DM/pre-diabetes and 1,953 unaffected individuals. The clinical data shows that in fact 1,291 individuals in the sample have T2DM; the remaining 2,512 individuals do not have diabetes. The percentage of individuals with a positive screening test that truly have undiagnosed T2DM/pre-diabetes ranged from 42.3% - 46.8%. The percentage of individuals with a negative screening test that are truly unaffected ranged from 83.0% - 85.6%. The positive and negative likelihood ratios measured the prediction model performance. The **accuracy** of the prediction model is **0.607** (given by $(TP+TN)/total = (824+1,486)/3,803$), while the **misclassification rate** is **0.393** (given by $(FP+FN)/total = (467+1,026)/3,803$).

Table 7.14 Confusion matrix of undiagnosed pre-T2DM

		True Disease Status (Gold standard = HbA1c)		Total
		Abnormal	Normal	
Test results (Predicted test)	Positive (+ve)	TP = 824	FP = 1,026	1,850
	Negative (-ve)	FN = 467	TN = 1,486	1,953
	Total	1,291	2,512	3,803

Abnormal: Undiagnosed (HbA1c $\geq$ 5.7%); Normal: (HbA1c < 5.7%); Positive: cut-point $\geq$ - 0.701; Negative: cut-point < -0.701

**Figure 7.6** Receiver operating characteristics (ROC) curve showing performance of original sample and bootstrap sample for undiagnosed diabetes and prediabetesPerformance measures of the prediction model

The summary statistics for the undiagnosed T2DM/pre-diabetes prediction model are presented in **Table 7.15**. The Saudi diabetes risk score model was constructed based on the coefficients in the final model. This risk model gave an AUC value of **0.65 (95%CI: 0.63, 0.67)**, and the optimal cut-point for undiagnosed diabetes using Youden index was ≥ 2.25 , and optimal-empirical cut-point ≥ -0.701 (≥ 0.3316). With this cut-off point, participants were categorised as at high risk of undiagnosed T2DM/pre-diabetes with sensitivity: 0.64, specificity: 0.59, a PPV of **0.6416** and an NPV of **0.6202**. **Table 7.15** shows the performance of the risk score tool for detecting and identifying undiagnosed T2DM/pre-diabetes cases at various values for the cut-off point.

- **Validation of the prediction model**

I internally validated the development model using a bootstrapping approach and assessed the variability in model estimates as described for the previous model. The performance of the validation model is very similar to the development model with a minimal increase in the AUC value to 0.67. Overall, the performance of bootstrap sample is very similar to the original sample using a range of performance measures.

Table 7.15 Summary statistics for different prediction models based on model selection and performance measures

Models	observation	AUC	Sensitivity	Specificity	Optimal cut-point	Youden index	HL-c	HL-p
Original sample	3,803	0.6519	0.64	0.59	-0.701	0.225	0.2487	10.24
Bootstrapping sample	3,803	0.6519	0.64	0.59	-0.701	0.225	0.2487	10.24

The Hosmer-Lemeshow (HL) test was used as a calibration measure to assess the model by comparing between observed and predicted probabilities. The Hosmer-Lemeshow test showed that the model fitted well (**p-value = 0.2487**). **Table 7.16** and **Figure 7.6** shows the results of the HL goodness-of-fit test. The observed values are a close match to the expected values, which means that the model is fitting well. The p-value is greater than the 0.05 threshold, meaning that I can reject the alternative hypothesis that the observed values do not match the expected values.

Table 7.16 Observed and predicted occurrence of prediabetes using the test score as a predictive variable in the groups defined by the Hosmer-Lemeshow Goodness-of-fit test

Group/ quantile	Likelihood/ probability of being undiagnosed diabetes	Present (undiagnosed)		Absent (normal)		Total
		Observed (1)	Expected (1)	Observed (0)	Expected (0)	
1	< 0.1993	85	71.5	328	341.5	413
2	0.1993 - 0.2266	68	72.1	280	275.9	348
3	0.2266 – 0.2609	83	96	311	297.9	394
4	0.2609 - 0.2911	125	120.1	305	309.9	430
5	0.2911 - 0.3263	96	101.3	233	227.7	329
6	0.3263 - 0.3688	148	145	272	275	420
7	0.3688 - 0.3938	129	131.6	213	210.4	342
8	0.3938 - 0.4522	175	166.4	216	224.6	391
9	0.4522 - 0.5072	159	172.6	203	189.4	362
10	> 0.6779	223	214.3	151	159.7	374
Results	H-L Statistic = 10.24	Degree of freedom = 8		Significance (P-value) = 0.2487		

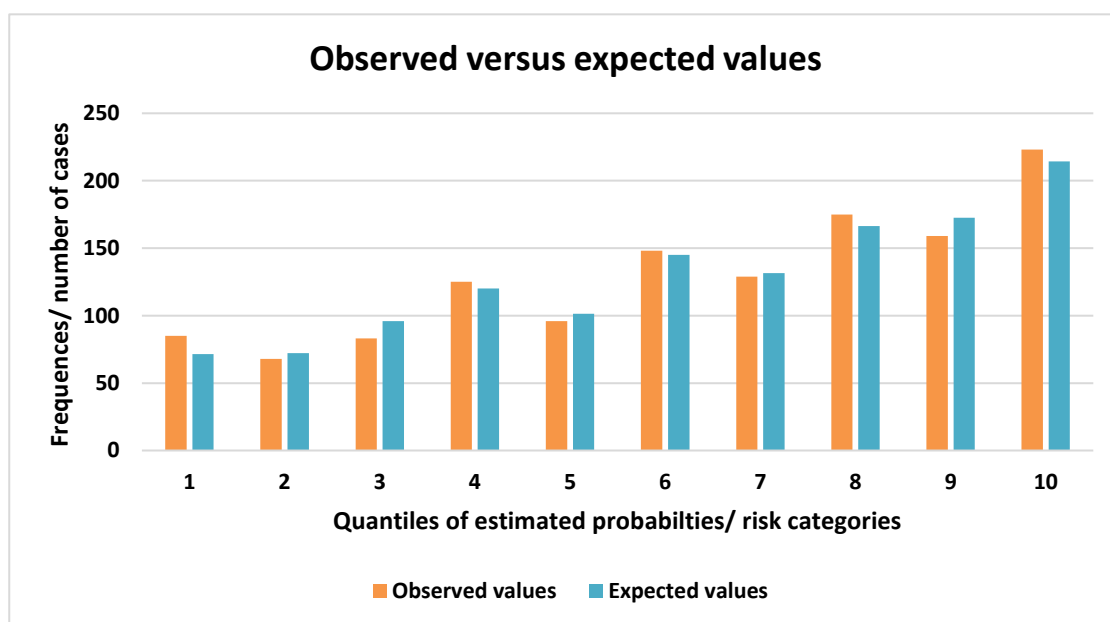


Figure 7.7 Calibration of the test score: observed and expected of undiagnosed prediabetes

7.5 Discussion and conclusion

In this chapter, I carried out an analysis to support development of two risk score models to identify individuals who are likely to have undiagnosed diabetes (model one), or a composite of undiagnosed diabetes with prediabetes (model two). The purpose of the study was to develop a specifically Saudi prediction model that included consideration of education level and socioeconomic status along with demographic, lifestyle, and health care access factors. Following up the previous analysis, two past binary logistic regression models (for undiagnosed T2DM and undiagnosed diabetes and pre-diabetes) were used to develop the prediction models. In each case, a variety of selection procedures were applied to identify the best predictors and to optimise the model. The ability of the prediction models to classify and identify high risk individuals with undiagnosed T2DM and prediabetes were formally assessed using a range of performance measures. The prediction models were also internally validated using a bootstrapping approach. The analysis allowed the production of a risk score model for T2DM that adds to the diagnostic criteria currently available in Saudi Arabia.

Overall, based on descriptive results, the evidence indicated that the likelihood of undiagnosed T2DM is higher for low-educated individuals. Males aged 55 - 64 who were obese and had no formal education had the highest risk of T2DM. However, an important consideration is that males in the middle income group who undergo routine health checks are more likely to be diagnosed with T2DM. Undiagnosed T2DM was positively associated with low-educated individuals (those who cannot read or write and have no formal education), consistent with past studies (Abdella et al., 1998, Costanian et al., 2014, Robbins et al., 2005, Wilder et al., 2005, Bener et al., 2009). These observations are explained by the fact that better-educated individuals are more likely to be aware of diabetes symptoms and its complications. The disease is more likely to be detected early in these individuals, as they tend to adopt a healthier lifestyle, be aware of diabetes, its symptoms and its complications, and to attend at health-care centres regularly (Clark and Utz, 2014). As a result, this group is more likely to have diagnosed T2DM, irrespective of the underlying prevalence of the disease.

Regarding employment status, unemployed, retired and homemaker groups were more likely to have undiagnosed T2DM than employed individuals. This association was particularly marked in the retired group. The excess of undiagnosed T2DM in these groups probably reflects their more limited access to healthcare services compared to employed individuals (Rabi et al., 2006). However, after applying selection procedures and comparing different prediction models, this study indicates that none of the socioeconomic factors showed significant association with undiagnosed diabetes and pre-diabetes. This finding is consistent with a previous study in the UAE that indicated education level and income were not associated (Sulaiman et al., 2018).

In terms of risk predictors, the findings indicate that the significant risk factors for undiagnosed T2DM are age (older age groups), obesity (BMI greater than or equal 30), and administrative region. These results are in line with most previous studies, which have provided evidence that age, and body mass index were important risk predictors for T2DM (Ruige et al., 1997, Baan et al., 1999, Tabaei and Herman, 2002, Glümer et al., 2004, Mohan et al., 2005, Ramachandran et al., 2005, Bang et al., 2009, Ko et al., 2010, Gray et al., 2010). The model developed in this study demonstrates that administrative region is the primary community socioeconomic risk predictor for both undiagnosed T2DM and pre-diabetes. This reflects the greater prevalence of these outcomes among individuals who lived in the Southern and Northern regions, particularly the higher frequency of undiagnosed T2DM in obese individuals living in the Southern region. In addition, the findings also indicate that the combination of risk factors giving the highest risk of undiagnosed diabetes and prediabetes are seen in older males who are obese and hypertensive. These results are consistent with previous studies indicating that age, hypertension, gender, body mass index, and region were important factors (Xin et al., 2010, Heikes et al., 2008b, Gray et al., 2010, Handlos et al., 2013).

As the case was in the previous model, administrative region emerged as a significant risk predictor in the model for undiagnosed diabetes and pre-diabetes. This is a surprising finding, that is to my knowledge the first study in the region to highlight the importance of an individual's place of residence as a predictor for T2DM. Using this approach to allocate resources to the high-risk regions could contribute to reducing the high prevalence of diabetes. At an individual level, the predictor could be included in a compound measure to identify individuals with pre-diabetes or undiagnosed T2DM. This would contribute to the early diagnosis and the better management of the disease and reduction of complications. In addition to better outcomes for patients, the measure could have significant economic benefit as a previous Saudi health economic study has revealed significant differences in health care costs between diagnosed diabetics (\$3,686) and non-diabetic (\$380) individuals (Alhowaish, 2013).

The apparent association of undiagnosed T2DM with administrative region might reflect previous studies in Saudi Arabia indicating regional variation in obesity and T2DM trends (Al-Nuaim, 1997). However, variation in healthcare access and individual socioeconomic status might also contribute to the effect as a recent analysis showed that socioeconomic level varied between provinces in Saudi Arabia (AlOmar et al., 2018). As discussed in the previous chapter, the increased risk of undiagnosed T2DM in low-educated individuals in Saudi Arabia can partly be explained by the fact that they tend to be older and they are less likely to be aware of the disease irrespective of gender. Note that the underlying risk of T2DM is higher in males, and that BMI is an important risk predictor in Saudi Arabia and across the GCC region.

Together with age, monthly household income and access to regular health care check-ups, these represent the key predictors in the diagnosed T2DM risk score model. Diabetes in other high-income countries is associated with individual/micro socioeconomic status; however, diabetes in Saudi Arabia appears more associated with macro/community socioeconomic level. High body mass index is an important risk predictor in most past studies and in the GCC region including Saudi Arabia. However, the most significant risk predictors were age, monthly household income, and health care check-up for diagnosed risk score model. Therefore, it is clear that age is an important predictor.

Regarding model performance, the scores displayed moderate sensitivity and specificity that are sufficient to identify individuals at high-risk of presenting with undiagnosed T2DM. In comparison to other countries in the GCC region (Oman, Kuwait, UAE), the findings of my model are quite consistent with regard to model performance. However, there are some differences in the risk predictors for undiagnosed T2DM. Unlike studies in other GCC regions, this study did not find sex and hypertension as useful risk predictors for undiagnosed T2DM. However, these two factors were significant predictors of the composite undiagnosed T2DM/pre-diabetes outcome. Furthermore, my risk model did not include waist circumference and family history of diabetes, as there was insufficient information in the dataset. However, my model provides the insight that administrative region is an important risk predictor. In terms of the sensitivity and specificity of the model's ability to identify individuals with undiagnosed diabetes/pre-diabetes, the scores are only moderate. Overall, the model performance is lower than that seen in the MENA prediction model. While it should be noted that age and BMI were the only predictors that my model and the MENA model have in common.

In terms of performance, the Omani study (Al-Lawati and Tuomilehto, 2007) had a higher AUC value (0.76 - 0.83) than my model, with higher sensitivity (0.63, 0.79) and lower specificity (0.73, 0.78). The Kuwaiti model (Al Khalaf et al., 2010b) gave a higher AUC value (0.82), higher sensitivity (0.87) and lower specificity (0.64), however this study was not validated by using another dataset. The UAE risk score model (Sulaiman et al., 2018) used similar diagnostic criteria to my model, but included different risk factors. The modest performance of this model can likely be attributed to heterogeneity in factors such as the diagnosis criteria, and differences in the underlying prevalence of T2DM between the geographical location where the prediction model was developed and subsequently tested.

Key differences between the published studies and my model are that the current study did not find gender and hypertension as predictive for diagnosed T2DM, although they were useful predictors of undiagnosed T2DM and pre-diabetes. Furthermore, I excluded consideration of waist circumference and family history of diabetes, as there was not enough information in the current dataset. The current model uses gender as a risk predictor for undiagnosed T2DM. This is similar to other studies in the UK (Gray et al., 2010, Griffin et al., 2000), Europe (Glümer et al., 2004, Baan et al., 1999) and the USA (Bang et al., 2009, Heikes et al., 2008a). The major new insight from the current study has been the identification of administrative region as an important risk predictor. The low performance of my model compared with other GCC region countries as Oman, UAE and Kuwait can be attributed to heterogeneity in different factors (the diagnosis criteria, and the prevalence of T2DM), between the geographical location where the prediction model was developed and tested.

Although age and body mass index were identified as significant risk predictors in previous studies, adding them to my model did not improve the prediction. Similarly, I did not include socioeconomic factors such as education level, employment status and monthly household income because they were not significant predictors of diabetes after model comparisons based on BIC, AIC selection measures, AUC, sensitivity, specificity and the Hosmer-Lemeshow p-value.

To conclude, the purpose of the current study was to develop a holistic prediction model for both undiagnosed T2DM and pre-diabetes. The study was successful in achieving this aim, and identified age, body mass index, and region as the most significant predictors for detecting undiagnosed T2DM. The most important predictors of the composite outcome of undiagnosed T2DM/pre-diabetes are age, body mass index, region, gender, and hypertension. The most obvious finding to emerge from this analysis is therefore that body mass index and administrative region are particularly meaningful factors in Saudi population.

Chapter 8 Socio-economic Factors Related to Controlled Diabetes in Saudi Arabia: Studying the Medication and Lifestyle Behavioural Patterns for Successful Diabetes Management

8.1 Introduction

Diabetes mellitus is one of the most serious public health issues in the 21st century (Shaw et al., 2010, Whiting et al., 2011). The WHO reported that diabetes mellitus is the third highest risk for early mortality after tobacco use and hypertension (WHO, 2009), and a contributor to 14.5% of all-cause mortality in those aged less than 60 years old. It has been estimated that in 2015, 5 million people aged 20 – 74 years died from diabetes-related causes, accounting for 12.8% of global mortality in that year (Ogurtsova et al., 2017). The same study revealed that the proportion of diabetes-related death in adults under 60 years was 29.6% and 72.7% in high and low-income countries, respectively. In addition to early mortality, diabetes complications reduce life quality and contribute to disability.

Controlled diabetes is characterised by a normalisation of blood glucose levels. A recent recommendation by the American Diabetes Association (ADA) defined poorly controlled diabetes as either having an HbA1c level more than or equal to 7% or fasting blood glucose (FPG) level more than 7.2 mmol/l for patients, excluding pregnant women (Association, 2016). Successful control of diabetes is strongly associated with a reduced risk of complications and mortality among patients (Group, 1998). Better diabetes management can therefore delay the onset of diabetes complications. Greenfield et al. (2009) found that patients with an HbA1c value less than or equal 6.5% had a lower 5-year lower likelihood of developing cardiovascular disease. They also found evidence that achieving lower HbA1c level in older patients reduced the incidence of comorbidities such as eye and foot problems, heart attack, and stroke.

In a study carried out in Oman, Al Balushi et al. (2014) showed that patients with controlled diabetes were less likely to have high diastolic blood pressure coupled with elevated levels of total cholesterol and low-density lipoprotein. Given this overwhelming evidence for the importance of diabetes control, in this chapter I examine the factors associated with blood sugar control as measured by HbA1c levels. I focus on patients with a pre-existing diagnosis of T2DM diabetes as shown in **Figure 8.1**.

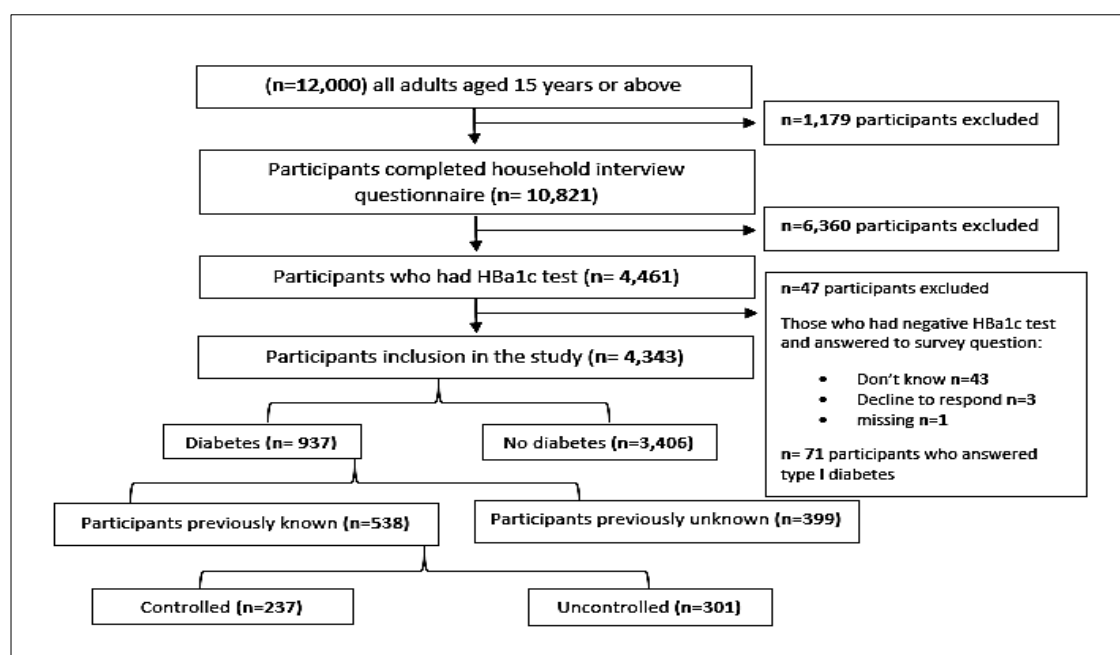


Figure 8.1 Flow chart for participants' inclusion in the study

Several studies have considered socio-economic risk factors for uncontrolled diabetes. As reviewed in Chapter 2, the most common factors influencing control of diabetes are age, gender, living in a deprivation area, socio-economic status, diabetes duration, family history, weight, diabetes-related knowledge, medication use, diet and use of insulin. To improve our understanding of why some diagnosed patients have poor control, we need to study the risk factors at both the individual and regional level (Stringhini et al., 2012, Grintsova et al., 2014, Clark and Utz, 2014). Self-management of disease is a good example of an individual level factor with significant impact on control of diabetes. In terms of regional factors, administrative region and urbanisation level are strongly associated with diversity in healthcare access, quality of healthcare facilities including access to specialist care, accessibility to healthy food, and the availability of public places and sport facilities for physical activity.

These environmental factors probably contribute to variation in the management and consequent control of diabetes in Saudi Arabia. It follows that gaining a better understanding of the relationship between education level and control of diabetes by considering individual and regional factors could make an important contribution to reducing the unacceptably high prevalence of uncontrolled diabetes.

Additional risk factors for uncontrolled diabetes I will consider in this section are an individual's knowledge of diabetes, how long they have had the condition, the type of healthcare facility providing care, diet, physical activity and medication with particular reference to insulin usage. For example, it has been reported that poor control of diabetes is linked to a lack of knowledge among patients about T2DM and its management (Fritschi et al., 2017), coupled with a reluctance to change lifestyle and behaviours (Matthews et al., 2009). A Jordanian study by Al-Khawaldeh et al. (2012) provided evidence that diabetes patients with poorly controlled disease were more likely to lack knowledge regarding the condition and the importance of insulin.

Thus, diabetes knowledge is one of the main individual level factors associated with controlled diabetes. In this context knowledge refers to an understanding of not only the disease and its complications, but also how best to manage diabetes through dieting, physical activity, adherence to medication and regular blood sugar level checks. Lack of such knowledge leads to uncontrolled diabetes through lack of medication compliance and a failure to monitor blood glucose adequately (Al-Qazaz et al., 2011, Barbara and Krass, 2013, Ong et al., 2014, Alramadan et al., 2018, Fallatah, 2018).

A Chinese study observed that self-efficacy - the confidence that an individual has in their ability to carry out a task - indirectly influenced self-care behaviours such as taking medication (Gao et al., 2013). In addition to taking medication and regular monitoring of blood glucose, following a healthy diet is another important aspect related to control of diabetes. An experimental study by Redmond et al. (2007) observed that nutrition intervention coupled with education about the importance of HbA1c level was effective and reduced HbA1c values. Tayyem et al. (2013) reported that self-efficacy was significantly associated with better glycaemic control among T2DM patients following an intervention based on diet and calorie counting to control carbohydrate consumption.

Duration of diabetes is another potential factor related to control of diabetes. Previous studies have claimed that a shorter duration of T2DM was associated with better glycaemic control (Bi et al., 2010, Lee et al., 2016, Khattab et al., 2010, Costanian et al., 2014, Alramadan et al., 2018, D'Souza et al., 2015, Alzaheb and Altemani, 2018). However, Al-Lawati et al. (2012) found the opposite in that patients with long duration of T2DM (≥ 5 years) had better glycaemic control compared to patients with short duration. Other studies failed to find a significant association between diabetes duration and glycaemic control (Abdelmoneim and Al Homrany, 2002, Al Balushi et al., 2014, Al-Kaabi et al., 2009). Thus, the evidence is inconclusive, and any such association is probably heavily context dependent.

Quality of health care provision is an important factor that could account for the association between socio-economic factors and diabetes control and mortality (Espelt et al., 2011). Some studies have identified an association between healthcare facility type and glycaemic control. Al-Baghli et al. (2010) found that Saudi patients who visited private health facilities had better glycaemic control compared to those who visited MOH or multiple health care facilities. In a UAE study, patients who were followed up in primary healthcare centres had a higher likelihood of poor glycaemic control compared to those who were followed up in hospitals (Malik et al., 1999). These findings supported my contention in Chapter 3 that provision for T2DM care in public healthcare services is generally poor across the GCC region.

To gain a more detailed understanding of how these factors combine to influence management of T2DM, we need to understand how management patterns for T2DM differ between regions, and how they influence outcomes. By diabetes management pattern, I mean the combination of medication (oral drug treatment and insulin) and behavioural modification (diet and physical activity) approaches. Accordingly, in this chapter I chose to focus on diabetes management patterns based on diet, medication, and insulin. I also investigated prescribed diet and physical activity to identify predictors of whether changes in lifestyle behaviour are likely to be successful, and how such predictors are related to an individual's socio-economic characteristics.

We already know based on numerous studies that the majority of the Saudi population are physically inactive and follow a poor diet (Majeed et al., 2014). In this chapter, I test the hypothesis that patients with diagnosed diabetes are more likely to modify their lifestyle and adopt beneficial behaviours. Rather than simply examine whether behavioural factors vary between patients with controlled and uncontrolled diabetes, I have adopted an approach based on identifying the main risk factors influencing lifestyle choices and behavioural characteristics among diabetics. I hypothesised that there is relationship between the behaviours that are key to T2DM management and the socio-economic factors that are the main focus of this thesis.

Some studies have studied the interaction between lifestyle modifications (i.e. diet and physical activity) and medication (i.e. both oral and insulin use). Patients who are prescribed insulin and anti-diabetes drugs may have poor glycaemic control (Khattab et al., 2010), however this is confounded by the fact that the combination of insulin and oral medication is an aggressive approach more likely to be applied to refractory disease (Amarasekara et al., 2015). Nonetheless, over two-thirds of the diabetics in this study failed to follow dietary regimens, did not participate in physical activity, did not check up on their blood glucose regularly at home and exhibited low compliance with diabetes treatment (Khattab et al., 2010).

In relation to physical activity, two studies failed to find an association between physical activity and glycaemic control (Al-Kaabi et al., 2008, Al-Kaabi et al., 2009). Conversely, other studies have highlighted the importance of both diet and physical activity in combination with oral medication and insulin treatment for the effective control of diabetes. Both Al-Nuaim et al. (1998) and Al-Lawati et al. (2012) reported that patients who followed a dietary regimen had better glycaemic control than those who only used oral medication and insulin. Al-Hayek et al. (2012) and Alramadan et al. (2018) found that patients in Saudi Arabia with HbA1c <7.0% had lower medication use than those with HbA1c $\geq$ 7.0%. This suggests that on its own, medication is insufficient for effective diabetes management and that a combination of medication and behavioural modification approaches is required. Alramadan et al. (2018) highlighted the higher prevalence of poor glycaemic control among physically inactive patients with limited intake of fruit and vegetables. Conversely, a Bahraini study found that patients who followed a balanced diet high in fruit and vegetables had a lower mean of HbA1c level than those with a poor diet (Shamsi et al., 2013).

The prevalence of poor control among patients with T2DM in the GCC region ranged from 54% - 65% in Oman (Al Balushi et al., 2014, D'Souza et al., 2015); 86.5% to 88.8% in Bahrain (Fikree et al., 2006, Nasser, 2007); 68% to 69% in the UAE (Al-Kaabi et al., 2009, Al-Kaabi et al., 2008); 55% in Kuwait (Badawi et al., 2015) (HbA1c > 9.0%); and 59% to 70.7% in Saudi Arabia (HbA1c $\geq$ 7.0%) (Al-Hayek et al., 2012, Alsulaiman et al., 2016). A review by Alhyas et al. (2011) indicated that the generally poor control of T2DM in the GCC region reflected missing diabetes data, poor standardisation of T2DM diagnostic criteria and a lack of diabetes management guidelines within primary healthcare centres. In this study I argue that socio-economic factors and place of residence also have a major impact on an individual's diabetes control status. As discussed in Chapter 2, low SES individuals are less likely to self-manage and adhere to medication/treatment resulting in poor glycaemic control.

Several previous studies in Saudi Arabia and the wider GCC region have investigated the risk factors for poor glycaemic control. These were based on examining lifestyle factors such as physical activity, dietary habits, and weight loss among those treated for T2DM. Few studies, however, have explored the impact of individual and regional socio-economic level on the relationship between control of diabetes, lifestyle and self-care factors. Where they do exist, they were insufficient to explore the influence of individual and regional socio-economic status on diet and physical activity. To achieve control and management goals, all potential risk factors for inadequate control of T2DM need to be considered in establishing diabetes management guidelines. Gaining a better understanding of the relationship between education level, lifestyle factors and regional variation in the control of diabetes is a necessary first step to improving management strategies to delay the onset of diabetes.

The Saudi Diabetes National guidelines recognise that diet and physical activity are the two main lifestyle and behavioural factors influencing control of high blood sugar levels. However, my review in Chapter 2 of diabetes and management guidelines in different high-income countries such as UK, Finland, Sweden, US, and GCC neighbouring countries revealed that region/place of residence is another important factor that should be considered in the design of diabetes management strategies. In this chapter, I will extend this observation by analysing the prevalence of uncontrolled diabetes in relation to educational level and other risk factors. I will then examine the most important lifestyle factors influencing behaviour among diabetic patients in Saudi Arabia.

In summary, the analysis addresses the following questions:

1. What are the factors independently related to the control of diabetes in Saudi Arabia, with specific focus on education level? How do demography, administrative region, health care access, self-care and other socio-economic factors relate to the control of diabetes?
2. Is there an association between controlled diabetes and patterns of diabetes management? If an association is present, how do these patterns vary between administrative regions?
3. What are the key factors in current behavioural management approaches for the diabetic Saudi population? Do these take account of patients' socio-economic factors and place of residence?

The overall aim of this analysis is to explore factors associated with poor glycaemic control in patients with diagnosed T2DM in Saudi Arabia and investigate the statistical association between a patient's education level and the prevalence of uncontrolled T2DM. The study considers a set of healthcare access and self-care factors (e.g. prescribed diet, use of medication). By contributing to a better understanding of the factors influencing the control of diabetes, this analysis will help healthcare providers develop strategies to improve the control of diabetes in the Saudi population.

This chapter is divided into four main sections: Background and rationale; methods; results; discussion and conclusion. The risk factors for uncontrolled diabetes, aims and research questions are discussed in **Section 8.2**. **Section 8.3** defines the study variables and explains the statistical methods that I use to address the chapter's aims. **Section 8.4** describes the results of the analysis divided into three main themes; the association between education level and controlled diabetes; the most important factors behind the diversity of lifestyle factors seen in diabetic patients; and the impact of regional variation and healthcare access. The chapter concludes with a discussion and summary of conclusions in **Section 8.5**.

8.2 Methods

8.2.1 The association between education level and inadequate control of T2DM

The total sample size available for analysis in this chapter was 538 individuals; however, I observed that some individuals had some missing responses to relevant survey questions. I therefore restricted my sample to subjects with complete data to ensure that the outcome of controlled diabetes was accurate. This carried the risk that by excluding subjects with incomplete data I might bias the results. To evaluate the impact this had on the validity of this study, we need to consider the role of bias and how it influences a study's statistical precision. Firstly, **selection bias** results from any choices made relating to inclusion of participants or factors affecting study participation, where these give rise to a relationship between study outcome and exposure that differs between study participants and those who were excluded or who were non-respondents (Tripepi et al., 2010).

Moreover, selection bias increases when study participants are not randomly selected from the target population, as there is a significant chance that the study sample might not be representative (Hammer et al., 2009). The most important consideration in representative cross-sectional population studies is high participation rate. The reasons for non-response or refusal to answer a question need to be considered when interpreting the results to avoid biased results.

This bias was mentioned in Chapter 5. This study used household weights to ensure that the sample was representative of the population and could be used to present valid results. I therefore applied household weights in the analysis to mitigate selection bias.

To assess whether there is a difference between those who were included and excluded from this analysis due to missing data on specific variables, an analysis comparing the two groups on a range of factors was undertaken. This is shown in Appendix G and indicate that participants who were included in the study differ in from those who did not answer with respect to factors which might influence the control diabetes outcome such socioeconomic factors. There was no significant relationship between those who completed and did not complete survey questions on the dimensions of age, sex, or marital status factors based on chi-square test. Nonetheless, there was a trend for excluded individuals with missing data to be older than included subjects. Females were also more likely to have missing information and there were disparities socioeconomic factors, including work status, education, income, and region.

In particular, the sample did not cover all socioeconomic groups equally, however the controlled diabetes sample is not truly representative of the diabetic population. I used the outcome "controlled diabetes" as the dependent variable in the analysis. This binary outcome variable is defined as: Controlled diabetes: individuals with an existing diagnosis of diabetes and HbA1c level less than 6.5%. Uncontrolled diabetes: individuals with an existing diagnosis of diabetes and HbA1c level more than 6.5%.

Furthermore, the issue of **confounding** can cause under-estimation of the observed association between exposure and health outcome. For example, age and sex variables are confounding variables that could influence the strength/power of the relationship in adjusted models. Confounding refers to mixing effects, where a study aims to assess the relationship between outcome and exposure by adding an additional factor or set of factors to this relationship (Skelly et al., 2012). Confounding variables could hide the actual association, or incorrectly support the relationship between the outcome and medication when no actual relationship between them is present (Skelly et al., 2012). The influence of confounding variables could make it difficult to observe a true association between medication and outcome unless methods are used to mitigate the effect. This issue could also be a primary reason for failing to achieve statistically significant results, and/or lead to inaccurate estimates (underestimate or overestimate) of the actual relationship between the study's outcome and exposure (Skelly et al., 2012).

Therefore, statistical analysis needs to take account of the possibility of bias introduced by confounder effects. One approach is to calculate adjusted estimates for the association between the exposure and outcome whether can be found and remain after taking out the effects of confounder. Overall, if the crude estimate differs from the adjusted estimate by 10% or more, the factor should be considered as confounder and the adjusted estimates used to reflect the effect of the exposure (Skelly et al., 2012). One possible method to check the possibility of confounder effect is to split the statistical test into sub-groups comprising different levels of the confounder (Hammer et al., 2009). Most observational studies are evaluated by regression models. Possible confounders can be added to the models as explanatory risk factors. The confounder effect is checked by comparing different models with and without the confounder (Hammer et al., 2009). However, in this chapter, the usual confounders are age and gender. As I found that these variables are not significantly associated with the main outcome of the study it is unlikely to be influenced by confounding.

In addition to above, another potential issue in this study is the occurrence of **survival bias**. This occurs when assessing the influence of treatment on individuals as exposed subjects need to survive up to a fixed time point before they receive treatment or are diagnosed (Zhou et al., 2005). This bias increases in observational studies that use one-point time from the beginning at which to follow up patients. Hence, the analyses fail to consider or recognise survival time for subjects that die or are lost before treatment resulting biased estimates (Zhou et al., 2005). The occurrence of this bias can be recognised when an individual's survival affects classification into two comparison groups. For example, individuals who were previously diagnosed and are taking medication are more likely to survive. We assume that the majority of those who are controlling their diabetes were taking medication. The results may not fully reflect the true relationship between outcome and taking medication due cross-sectional nature of the study design.

I used the outcome "controlled diabetes" as the dependent variable in the analysis. This binary outcome variable is defined as:

- *Controlled diabetes*: individuals with an existing diagnosis of diabetes and HbA1c level less than 6.5%
- *Uncontrolled diabetes*: individuals with an existing diagnosis of diabetes and HbA1c level less than 6.5%

To address the first and second research questions outlined in section 8.1, a bivariate analysis was initially applied to test for association between four main factors and explanatory variables using the Chi-square test with threshold for significance of 0.05. I then built models to

explore the association between controlled diabetes and education level, before and after adjusting for demographic, socio-economic, health care access and self-care behaviour characteristics. The models were weighted by household, and odds ratios estimated for education level and other risk factors. The following figure and table present the analytical framework and four logistic regression models.

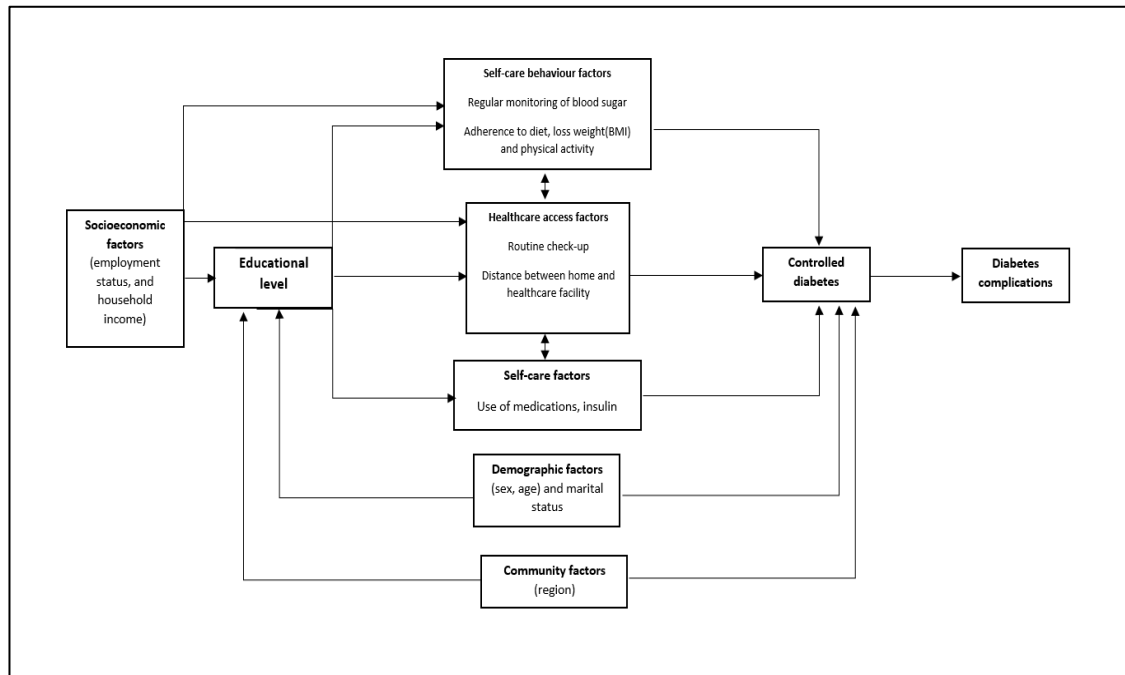


Figure 8.2 Analytical framework for the relationship between education level and inadequate control of T2DM

Table 8.1 The analytical framework of controlled diabetes

Main predictor	Model adjustment	Secondary predictors list based on SHIS database
Education level	Model 1: Unadjusted	Educational level only
	Model 2: Demographic factors, Administrative region, and Socio-economic factors	Age and sex, marital status, administrative region, working status and monthly household income.
	Model 3: Self-care behaviour factors	Test blood sugar at home, loss weight (BMI), prescribed diet, and physical activity.
	Model 4: Health care access and self-care factors	Routine check-up, visit for manage diabetes, use of medications, and insulin.

8.2.2 Medication and lifestyle patterns among diabetic patients

In the second analysis, I addressed the second research question by analysing the relationship between controlled diabetes and the combination of medication, diet, and insulin. I also examined the difference between administrative regions in relation to demographic and socio-economic factors using bivariate analysis, assessing significance using the Chi-square test. The final analysis examined how socio-economic risk factors related to the prescription of diet and physical activity. Prescribed diet was treated as a binary dependent variable, reflecting whether a diabetic patient had a prescribed diet (yes or no). Consequently, I used the binary logistic modelling approach to explore the association between prescribed diet and socio-economic, demographic, or healthcare access factors. A stepwise selection procedure was used to identify the most significant factors associated with diet and physical activity, before carrying out regression modelling focused on the most important indicators. To measure **physical activity**, I used the International Physical Activity questionnaire (Craig et al., 2003) to categorise participants into four groups: (1) meet vigorous physical activity criteria; (2) meet moderate physical activity criteria; (3) insufficient physical activity to meet vigorous or moderate levels; and (4) no physical activity (see section 5.3.2 in Chapter 5).

Physical activity is an ordered variable with more than two categories; thus, it requires the use of the ordered logit model (OLM - also known as the proportional odds model). The main assumptions of this approach are 1) that the dependent or outcome variable is measured at an ordinal level and that 2) there is no multi-collinearity (i.e. no two independent variables or predictors are associated with each other).

This latter assumption is frequently violated (Williams, 2016). The parallel lines or proportional odds assumption requires that any paired grouping of outcome category gives similar results for various categories of covariate variables/predictors. If this assumption is invalid for one of the model predictors, the results from ordered logit model may be biased and inaccurate. In such cases, some researchers switch the modelling approach to a multinomial logit model (MLM) or use the generalised ordered logit (GOL)/partial proportional odds model. These approaches do not consider the order of categories and produce more parameters as a consequence of not including the proportional odds constraint. The generalised ordered logit model equation contains two main elements: the proportional odds model and the partial proportional odds model that allows the coefficients of the variable to vary the thresholds of the response variable.

There are, however, a number of issues with the GOL model (Williams, 2016). First, it works best when there are few independent variables. A large number of predictors can violate the model; in such cases the MLM will be preferred. Secondly, the GOL model can result in negative predicted probabilities. This can occur when N is small in some categories of response or outcome variable. Finally, the proportional odds assumption can be significant, even in large sample sizes with small violations. These factors often lead to confusion over the best model to use, even after assessing the models using information criteria such as AIC or BIC measures when adding criteria. In the current study, the outcome did not satisfy the OLM assumptions, so I chose to use the multinomial logistic regression model.

Multinomial logistic modelling uses an extension of the binary logistic model to assess the association between an unordered or nominal categorical outcome variable and independent predictors. MLR allows for regression to transform probabilities into log-odds by the logit function on the outcome variable as discussed in Chapter 4. It can be used to calculate the probability of each category for the outcome variable. The multinomial logistic model can be written as (Agresti and Kateri, 2011):

$$\log \left(\frac{p_{ik}}{p_{i3}} \right) = \alpha_k + \beta_k x_i \quad 1 \leq k \leq 2 \quad (8.1)$$

Where: $\log \left(\frac{p_{ik}}{p_{i3}} \right)$ = Log odds of the outcome variable; p_{i3} = the reference category.

α_k = intercept (constant); β_k = control variable coefficient.

Linear probability model:

$$p_i = E(y_i) = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} \quad (8.2)$$

The following transformation model is used to find the link function:

$$g(p_i) = g(E(y_i)) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} \quad (8.3)$$

Where: p_i is the category of outcome; $g(p_i)$ is the logit function of the outcome variable category.

$$p_i = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki})}$$

After fitting the final model, the predicted probabilities can be calculated by using the following equation:

$$p_{ik} = \frac{\exp(\alpha_k + \beta_k x_i)}{1 + \sum_{k=1}^2 \exp(\alpha_k + \beta_k x_i)} \quad (8.4)$$

$$p_{i3} = \frac{1}{1 + \sum_{k=1}^3 \exp(\alpha_k + \beta_k x_i)} \quad (8.5)$$

Where: $k = 1, 2, 3$; p_{i3} = the predicted probability of baseline category.

MLR does not assume normality or linearity, but any relationship between categories of outcome variable are ignored. As described above, the physical activity variable used has four categories. However, the first and second categories (high and moderate) were combined due to the low numbers of individuals in the high category, giving three categories of outcome: (high or moderate, low, and none). The following table shows the explanatory variables I used to assess the relationship between controlled diabetes and education level as the main predictor. In the event, I had to reclassify most of the independent variables in a similar way due to the reduction in sample size. The analysis presented here used the explanatory variables shown in **Table 8.2**. All analyses were carried out using STATA software version 14 (Stata Corp. Inc, TX, USA) and Microsoft Office Excel (2013) to construct all figures.

Table 8.2 Study explanatory variables list

Study variables		Categories
Demographic factors	Sex	Dichotomous variable: 1 = Male, 2 = Female; <i>Reference category</i> : Female
	Age	Recoded into three categories: (1 = 20 – 39; 2 = 40 – 59; 3 = 60+) <i>Reference category</i> : (60+)
	Marital status	This was recoded into two categories: 1 = Currently Married; 2 = Not currently married: Never married; Separated/divorced, and Widowed. <i>Reference category</i> : Currently Married
Administrative region	Region location	Recoded into five categories based on geographical location: 1 = Central region; 2 = Eastern region; 3 = Western region; 4 = Northern region; 5 = Southern region; <i>Reference category</i> : Central region
Socio-economic factors	Education level	Recoded into three categories: 1 = Low education level: cannot read or write, and no formal education. 2 = Middle education level: Primary and intermediate school education 3 = High education level: High school, university degree and above. <i>Reference category</i> : High education level
	Working status	Coded into three categories: 1 = Currently working : Employee: Government and Non-government employee; 2 = Not currently working : Unemployed; Retired; 3 = Never worked : Homemaker; Self-employed; Student. <i>Reference category</i> : currently working
	Monthly household income	Recoded into five categories: 1 = Lowest : Less than 3,000 Riyals; 3,000 Riyals to less than 5,000 5,000 Riyals to less than 7,000; 2 = Low : 4=7,000 Riyals to less than 10,000 3 = Middle : 5= 10,000 Riyals to less than 15,000; 4 = High : 6=15,000 Riyals to less than 20,000; 5 = Highest : 7=20,000 Riyals to less than 30,000; 8= 30,000 Riyals or more; 6 = Don't know/ decline to respond or missing <i>Reference category</i> : Highest
Self-care behavior factors	loss weight (BMI) Body Mass Index	Recoded into two categories: 1 = Underweight or Normal; 2 = Overweight. 3 = Obese <i>Reference category</i> : Underweight or Normal
	Physical activity levels	Recoded into three categories: 1 = Moderate or High; 2 = Low; 3 = Not active <i>Reference category</i> : High or moderate
	Prescribed diet	Coded into two categories: 1 = Yes; 2 = No. <i>Reference category</i> : Yes
Health care access and Self-care factors	Routine check-up	Recoded into two categories: 1 = Yes; 2 = No <i>Reference category</i> : No
	visit for manage diabetes	Coded into two categories: 1 = Yes; 2 = No. <i>Reference category</i> : Yes
	Test blood sugar at home	Recoded into two categories: 1 = No; 2 = Yes. <i>Reference category</i> : Yes
	Use of medications	Coded into two categories: 1 = Yes; 2 = No. <i>Reference category</i> : Yes
	insulin	Coded into two categories: 1 = Yes; 2 = No. <i>Reference category</i> : Yes

8.3 Results

8.3.1 The association between education level and controlled diabetes

The percentage of controlled diabetes in various groups defined by socio-economic, health care access and self-care factors are presented in **Table 8.3**. In the overall sample, the percentage of individuals with 'uncontrolled diabetes' was higher than those with 'controlled diabetes' (56.1% vs 43.9% respectively). More than half of individuals currently taking medication had uncontrolled diabetes, indicating that they were still not managing their blood sugar sufficiently. This might indicate that their medication was inadequate or reflect poor compliance with therapy. Males appeared more likely to have controlled diabetes (46.0% vs 40.9% respectively). In addition, low-educated groups that cannot read or write group were less likely to have controlled diabetes than highly educated groups (university degree or above) (40.3% vs 50.9% respectively).

However, education level did not have a significant relationship with controlled diabetes. Similarly, individuals defined as 'currently working' have a higher proportion of controlled diabetes than other working status groups ('never worked' or 'not currently working'). Based on self-care behaviour factors, individuals who were following a prescribed diet and were physically active were more likely to have controlled diabetes compared with other groups ('no prescribed diet' or 'physical activity low', physical activity none'). In terms of administrative region, Individuals who lived in the Southern and Northern regions were most likely to have 'uncontrolled diabetes'.

Table 8.3 Characteristics of diabetes patients with controlled and uncontrolled diabetes

Variables		Categories N (%)	Total 462	Uncontrolled 259 (56.1)	Controlled 203 (43.9)	p-value
Demographic factors						
Sex	Female		259	110 (59.1)	76 (40.9)	0.274
	Male		203	149 (54.0)	127 (46.0)	
Age	20 - 39		38	19 (50.0)	19 (50.0)	0.541
	40 - 59		212	116 (54.7)	96 (45.3)	
	60 +		212	124 (58.4)	88 (41.5)	
Marital status	Currently married		355	192 (54.1)	187 (45.9)	0.119
	Not currently married		107	67 (62.6)	40 (37.4)	
Socio-economic factors						
Education level	Low education		236	141 (59.8)	95 (40.3)	0.175
	Middle education		118	65 (55.1)	53 (44.9)	
	High education		108	53 (49.1)	55 (50.9)	
Working status	Never worked		112	64 (57.1)	48 (42.9)	0.147
	Currently working		112	54 (48.2)	58 (51.8)	
	Not currently working		238	141 (59.2)	97 (40.8)	
Monthly Household income	Lowest		214	119 (55.6)	95 (44.4)	0.751
	Low		65	39 (60.0)	26 (40.0)	
	Middle		57	29 (50.9)	28 (49.1)	
	High		33	19 (57.6)	14 (42.4)	
	Highest		22	10 (45.5)	12 (54.5)	
	Missing		71	43 (60.6)	28 (39.4)	
Region location						
Administrative Region	Central		140	73 (52.1)	67 (47.9)	0.000
	Eastern		24	13 (54.2)	11 (45.8)	
	Western		77	27 (35.1)	50 (64.9)	
	Northern		89	55 (61.8)	34 (38.2)	
	Southern		132	91 (69.0)	41 (31.0)	

Table 8.3 Characteristics of diabetes patients with controlled and uncontrolled diabetes (cont.)

Variables	Categories N (%)	Total 462	Uncontrolled 259 (56.1)	Controlled 203 (43.9)	p-value
Health care access factors					
Routine check-up	Never	235	132 (56.2)	103 (43.8)	0.961
	Yes	227	127 (56.0)	100 (44.0)	
Health care visit for manage diabetes in last month	Yes	403	227 (56.3)	176 (43.7)	0.763
	No	59	32 (54.2)	27 (45.8)	
Self-care factors					
Have you ever taken medications (two weeks)	Yes	396	230 (58.1)	166 (41.9)	0.032
	No	66	29 (43.9)	37 (56.1)	
Insulin	Yes	184	107 (58.2)	77 (41.8)	0.461
	No	278	152 (54.7)	126 (45.3)	
Self-care behavior factors					
Test blood sugar at home	Yes	259	139 (56.0)	109 (44.0)	0.995
	No	203	120 (56.0)	94 (44.0)	
Prescribed diet	Yes	361	195 (54.0)	166 (46.0)	0.094
	No	101	64 (63.4)	37 (36.6)	
Physical activity	High or moderate	40	20 (50.0)	20 (50.0)	0.696
	Low	15	9 (60.0)	6 (40.0)	
	None	407	230 (56.5)	177 (43.5)	
Body mass index (BMI)	Under or normal weight	60	28 (46.7)	32 (53.3)	0.277
	Overweight	161	91 (56.5)	70 (43.5)	
	Obese	241	140 (58.0)	101 (42.0)	

The odds ratios (ORs) and 95% confidence interval (CI) from the weighted binary logit regression models of unadjusted and adjusted associations between controlled diabetes and education level are shown in **Table 8.4**. This analysis shows that there is no association between education level and controlled diabetes. However, administrative region, prescribed diet and taking medication appeared as significant predictors of controlled diabetes across all the models. Diabetes patients who lived in Western region were significantly more likely to control their diabetes than those who lived in the Central region (OR: 2.91, CI: 1.45, 5.85). However, patients who lived in Southern region were less likely to have controlled diabetes than those who lived in the Central region (OR: 0.51 CI: 0.27, 0.96). It appears that diabetes patients who lived in eastern or northern regions do not significantly differ from those who lived in central region.

Regarding other socio-economic factors, neither working status nor monthly household income were predictors of controlled diabetes. For self-care lifestyle factors, prescribed diet is associated with control of diabetes, as diabetes patients who are not following a prescribed diet are less likely to have controlled diabetes (OR: 0.32 CI: 0.16, 0.61). However, neither physical activity nor BMI are significant predictors of controlled diabetes. In addition, after controlling the model for self-care and healthcare access factors, patients who had never taken diabetes medication are significantly more likely to have controlled diabetes than those who are taking medication (OR: 2.35, CI: 1.13, 4.86). However, 'insulin use', a 'recent healthcare visit' and 'routine check-up' were not associated with control of diabetes.

As mentioned in methods section, the results of this study may not nationally representative due to differences in socioeconomic factors between those who were included in the study and those who were excluded. However, this study used household weights to ensure a representative sample of the population that could be used to present valid results. Hence, this weight was also applied in the analysis to mitigate the effect of selection bias.

In the results from the adjusted models, although the odds appeared significant, the effect of confounders may reduce the significance levels. I think that this possibility is unlikely for two reasons. First, the effect persisted even if the analysis was divided by confounder subgroup levels. Secondly, both sex and age factors did not show significant association with the main outcome of the study (control diabetes). Hence, this study is unlikely to be influenced by confounding bias, as this would require the confounder to be associated or related to both exposure and outcome. In the event, most of the sample were taking medication and this factor did achieve significance; however, as mentioned above, this observation is unreliable due to the cross-sectional nature of the study.

Table 8.4 Binary logistic regression odds ratios and 95% confidence interval for controlled diabetes (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Education level										
High education (reference)	1		1		1		1			
Low education	0.802 (0.439 - 1.464)	0.473	0.954 (0.406 - 2.240)	0.913	1.134 (0.447 - 2.879)	0.791	1.188 (0.467 - 3.023)	0.718		
Middle education	1.062 (0.534 - 2.111)	0.864	1.056 (0.450 - 2.479)	0.900	1.237 (0.521 - 2.939)	0.630	1.318 (0.556 - 3.124)	0.530		
Age group										
60+ (reference)			1		1		1			
20 - 39			0.855 (0.289 - 2.521)	0.776	0.847 (0.284 - 2.525)	0.766	0.738 (0.235 - 2.319)	0.603		
40 - 59			0.791 (0.427 - 1.466)	0.456	0.739 (0.396 - 1.380)	0.344	0.726 (0.384 - 1.374)	0.325		
Sex										
Female (reference)			1		1		1			
Male			1.239 (0.565 - 2.716)	0.593	1.082 (0.504 - 2.323)	0.839	1.108 (0.501 - 2.449)	0.801		
Marital status										
Currently married (reference)			1		1		1			
Not currently married			0.703 (0.339 - 1.454)	0.342	0.609 (0.304 - 1.224)	0.164	0.623 (0.304 - 1.277)	0.196		
Administrative Region										
Central (reference)			1		1		1		1	
Eastern			0.587 (0.195 - 1.767)	0.343	0.984 (0.299 - 3.239)	0.979	0.949 (0.283 - 3.182)	0.956	1.090 (0.359 - 3.314)	0.879
Western			1.913 (0.912 - 4.008)	0.086	2.537 (1.216 - 5.296)	0.013	2.373 (1.135 - 4.959)	0.022	2.908 (1.445 - 5.853)	0.003
Northern			0.677 (0.314 - 1.459)	0.320	0.739 (0.336 - 1.624)	0.452	0.674 (0.309 - 1.465)	0.319	0.814 (0.412 - 1.609)	0.554
Southern			0.441 (0.232 - 0.865)	0.017	0.479 (0.238 - 0.966)	0.040	0.419 (0.207 - 0.848)	0.016	0.513 (0.273 - 0.963)	0.038
Model 1: unadjusted; Model 2: adjusted by demographic factors (sex and age) and marital status, administrative regions, socio-economic factors (working status and monthly household income). Model 3: includes also self-care behavior factors (prescribed diet, physical activity, test blood sugar at home and BMI). Model 4: includes also Health care access and self-care factors (routine check-up, and visit for manage diabetes, (Use of medications, and insulin). Model 5: final fitted model, P: p-value (significant in BOLD); OR: Odds Ratios; CI: Confidence interval.										

Table 8.4 Binary logistic regression odds ratios and 95% confidence interval for controlled diabetes (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Working status										
Never worked (reference)			1		1					
Currently working			1.097 (0.404 - 2.983)	0.856	1.304 (0.472 - 3.602)	0.609	1.243 (0.442 - 3.498)	0.680		
Not currently working			0.522 (0.223 - 1.219)	0.133	0.570 (0.253 - 1.282)	0.174	0.576 (0.249 - 1.335)	0.199		
Monthly household income										
Highest (reference)			1		1		1			
Lowest			0.632 (0.130 - 3.065)	0.569	0.467 (0.118 - 1.855)	0.288	0.385 (0.096 - 1.539)	0.177		
Low			0.351 (0.064 - 1.911)	0.226	0.296 (0.065 - 1.349)	0.113	0.254 (0.055 - 1.170)	0.079		
Middle			0.388 (0.069 - 2.163)	0.280	0.320 (0.067 - 1.520)	0.152	0.244 (0.051 - 1.162)	0.077		
High			0.314 (0.054 - 1.821)	0.197	0.238 (0.046 - 1.225)	0.086	0.192 (0.036 - 1.034)	0.055		
Missing			0.461 (0.091 - 2.344)	0.350	0.482 (0.113 - 2.049)	0.323	0.389 (0.089 - 1.695)	0.209		
Test blood sugar at home										
Yes (reference)					1		1			
No					1.127 (0.661 - 1.923)	0.661	1.019 (0.589 - 1.760)	0.947		
Body Mass Index (BMI)										
Under or normal weight (reference)					1		1			
Overweight					0.722 (0.308 - 1.694)	0.455	0.704 (0.311 - 1.596)	0.401		
Obese					0.571 (0.248 - 1.315)	0.188	0.536 (0.242 - 1.186)	0.124		
Physical activity										
High or moderate (reference)					1		1			
Low					1.278 (0.229 - 7.105)	0.779	1.155 (0.183 - 7.284)	0.878		
None					0.661 (0.259 - 1.691)	0.388	0.581 (0.224 - 1.507)	0.265		

Model 1: unadjusted; **Model 2:** adjusted by demographic factors (sex and age) and marital status, administrative regions, socio-economic factors (working status and monthly household income). **Model 3:** includes also self-care behavior factors (prescribed diet, physical activity, test blood sugar at home and BMI). **Model 4:** includes also Health care access and self-care factors (routine check-up, and visit for manage diabetes (Use of medications, and insulin), **Model 5:** final fitted model, **P:** p-value (significant in **BOLD**); **OR:** Odds Ratios; **CI:** Confidence interval.

Table 8.4 Binary logistic regression odds ratios and 95% confidence interval for controlled diabetes (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)	
	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
Prescribed diet										
Yes (reference)					1		1		1	
No					0.273 (0.134 - 0.558)	0.000	0.232 (0.112 - 0.479)	0.000	0.315 (0.164 - 0.607)	0.001
Visit for manage diabetes in last month										
Yes (reference)							1			
No							1.435 (0.605 - 3.403)	0.412		
Routine check-up										
Never (reference)							1			
Yes							0.766 (0.453 - 1.295)	0.319		
Have you ever taken medications										
Yes (reference)							1		1	
No							2.234 (1.007 - 4.959)	0.048	2.348 (1.134 - 4.858)	0.021
Insulin										
Yes (reference)							1			
No							1.358 (0.781 - 2.358)	0.278		
Model 1: unadjusted; Model 2: adjusted by demographic factors (sex and age) and marital status, administrative regions, socio-economic factors (working status and monthly household income). Model 3: includes also self-care behavior factors (prescribed diet, physical activity, test blood sugar at home and BMI). Model 4: includes also Health care access and self-care factors (routine check-up, and visit for manage diabetes (Use of medications, and insulin), Model 5: final fitted model. P: p-value (significant in BOLD); OR: Odds Ratios; CI: Confidence interval.										

8.3.2 Medication and behavioural forms for management diabetes

Table 8.5 lists the different patterns of diabetes management seen among Saudi patients. The proportion of the outcome 'controlled diabetes' by management pattern indicates that most patients are using medication and diet (40.0%), diet, medication, and insulin (29.4%), or medication only (10.4%) for control of their diabetes. However, I did not observe any significant differences in management patterns between 'controlled' and 'uncontrolled' diabetic patients. In **Table 8.6** I analyse the diabetes management patterns by administrative region. Patients who lived in the Central region were more likely to follow a combination of diet, medication and insulin for managing their diabetes in contrast to other regions that were more likely to follow a simpler combination of prescribed diet and medication only.

Table 8.5 Diabetes management patterns and control of diabetes in Saudi patients

Management forms	Controlled diabetes outcome (p-value=0.097)		
	Uncontrolled diabetes 259 (56.1)	Controlled diabetes 203 (43.9)	All diagnosed diabetes 462
No diet, medication or insulin	8 (40.0)	12 (60.0)	20
Diet only	12 (48.0)	13 (52.0)	25
Diet and medication	100 (54.0)	85 (45.0)	185
Diet and insulin	6 (40.0)	9 (60.0)	15
Medication only	32 (66.7)	16 (33.3)	48
Medication and insulin	21 (77.8)	6 (22.2)	27
Insulin only	3 (50.0)	3 (50.0)	6
Diet, medication and insulin	77 (56.6)	59 (43.4)	136

Table 8.6 Diabetes management patterns among Saudi patients by administrative region

Management forms	Administrative regions				
	Central	Eastern	Western	Northern	Southern
No diet, medication or insulin	6 (4.3)		6 (7.8)	2 (2.3)	6 (4.6)
Diet only	10 (7.1)	1 (4.2)	5 (6.5)	2 (2.3)	7 (5.3)
Diet and medication	48 (34.3)	9 (37.5)	28 (36.4)	49 (55.0)	51 (38.6)
Diet and insulin	7 (5.0)			3 (3.4)	5 (3.8)
Medication only	12 (8.6)	7 (29.2)	12 (15.6)	4 (4.5)	13 (9.8)
Medication and insulin	3 (2.1)	2 (8.3)	3 (3.9)	7 (7.8)	12 (9.1)
Insulin only		3 (12.5)		1 (1.1)	2 (1.5)
Diet, medication and insulin	54 (38.6)	2 (8.3)	23 (29.8)	21 (23.6)	36 (27.3)

Green colour: most form used

Based on previous results, administrative region, prescribed diet and taking medication are the most important factors influencing control of diabetes. In addition, diabetes management guidelines in Saudi Arabia emphasise the importance of changing behaviour to follow a better diet and become more physically active. A key objective in this chapter is therefore the identification of lifestyle patterns which influence the control of diabetes. The results in **Table 8.7** show the odds ratios (OR) and 95% confidence interval (CI) from weighted binary logit regression models of the likelihood of following a prescribed diet. This analysis revealed no significant effect of age, sex, and marital status.

In terms of patient education level, no education group was a significant predictor of following a prescribed diet. However, I did detect an association between administrative region and prescribed diet in diabetes patients. Patients living in the Eastern, Western and Southern regions were significantly less likely to follow a prescribed diet (OR: 0.109, 0.307, 0.368) than those who lived in Central region. Patients who lived in the Northern region did not significantly differ from those who lived in the Central region in terms of prescribed diet. Working status also showed significant association with prescribed diet.

Chapter 8

Patients who have 'never worked' are significantly more likely to follow a prescribed diet than those who are 'currently working' (OR: 3.297), however there was no significant difference between patients identified as 'not currently working' and those who were 'currently working'. Furthermore, patients who do not regularly test blood sugar at home are less likely to follow a prescribed diet than those who do (OR: 0.533). Overall, administrative region appeared the most significant predictor of whether diabetes patients follow a prescribed diet.

Table 8.7 Generalised binary logistic regression odds ratios and 95% confidence interval for prescribed diet (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)	
	OR [95% CI]	p	OR [95% CI]	p	OR [95% CI]	p	OR [95% CI]	p
Education level								
High education (reference)	1		1		1			
Low education	0.603 (0.292 - 1.246)	0.172	0.545 (0.225 - 1.322)	0.180	0.578 (0.224 - 1.489)	0.256		
Middle education	0.699 (0.308 - 1.588)	0.393	0.519 (0.218 - 1.238)	0.139	0.439 (0.167 - 1.153)	0.095		
Age								
60 + (reference)			1		1			
20 – 39			1.435 (0.421 - 4.895)	0.564	1.923 (0.466 - 7.936)	0.366		
40 – 59			1.488 (0.716 - 3.093)	0.287	1.581 (0.726 - 3.445)	0.249		
Sex								
Female (reference)			1		1			
Male			1.801 (0.709 - 4.571)	0.216	2.083 (0.750 - 5.782)	0.159		
Marital status								
Currently married (reference)			1		1			
Not currently married			1.681 (0.627 - 4.511)	0.302	1.948 (0.701 - 5.417)	0.201		
Administrative Region								
Central (reference)			1		1		1	
Eastern			0.111 (0.037 - 0.333)	0.000	0.118 (0.037 - 0.375)	0.000	0.109 (0.035 - 0.339)	0.000
Western			0.296 (0.121 - 0.726)	0.008	0.412 (0.167 - 1.017)	0.054	0.307 (0.131 - 0.719)	0.006
Northern			0.515 (0.186 - 1.425)	0.201	0.715 (0.225 - 2.269)	0.569	0.482 (0.180 - 1.292)	0.147
Southern			0.401 (0.170 - 0.948)	0.037	0.504 (0.209 - 1.215)	0.127	0.368 (0.174 - 0.777)	0.009
Working status								
Currently working (reference)			1		1			
Never worked			2.998 (1.000 - 8.990)	0.050	3.297 (1.102 - 9.860)	0.033		
Not currently working			1.685 (0.694 - 4.090)	0.249	1.693 (0.676 - 4.238)	0.261		
Model 1: unadjusted; Model 2: adjusted by demographic factors (sex, age, and marital status), administrative regions, socio-economic factors (working status and monthly household income). Model 3: includes also self-care behaviour factors (test blood sugar at home and BMI) and self-care factors (and visit for manage diabetes, Use of medications). Model 4: administrative region, P: p-value (significant in BOLD); OR: Odds Ratios; CI: Confidence interval.								

Table 8.7 Generalised binary logistic regression odds ratios and 95% confidence interval for prescribed diet (Household weighted models)

Characteristics	Model (1)		Model (2)		Model (3)		Model (4)	
	OR [95% CI]	p	OR [95% CI]	p	OR [95% CI]	p	OR [95% CI]	p
Monthly household income								
Lowest (reference)			1		1			
Low			0.606 (0.235 - 1.562)	0.300	0.517 (0.197 - 1.355)	0.179		
Middle			1.068 (0.375 - 3.042)	0.903	0.908 (0.304 - 2.708)	0.862		
High			0.764 (0.243 - 2.408)	0.646	0.570 (0.168 - 1.937)	0.368		
Highest			0.504 (0.150 - 1.694)	0.268	0.356 (0.082 - 1.552)	0.169		
Missing			0.295 (0.131 - 0.662)	0.003	0.260 (0.104 - 0.647)	0.004		
Test blood sugar at home								
Yes (reference)					1			
No					0.533 (0.287 - 0.992)	0.047		
Body Mass Index (BMI)								
Under or normal weight (reference)					1			
Overweight					1.682 (0.625 - 4.528)	0.303		
Obese					2.597 (0.997 - 6.766)	0.051		
Visit for manage diabetes in last month								
Yes (reference)					1			
No					0.432 (0.175 - 1.069)	0.069		
Have you ever taken medications								
Yes (reference)					1			
No					0.605 (0.294 - 1.244)	0.172		
Model 1: unadjusted; Model 2: adjusted by demographic factors (sex, age and marital status), administrative regions, socio-economic factors (working status and monthly household income). Model 3: includes also self-care behaviour factors (test blood sugar at home and BMI) and self-care factors (and visit for manage diabetes, Use of medications). Model 4: administrative region, P: p-value (significant in BOLD); OR: Odds Ratios; CI: Confidence interval.								

The effect of following a prescribed diet on the proportion of patients taking medication is presented in **Figure 8.4**. Patients who followed a prescribed diet appeared more likely to be taking medication compared to those who were not on diet. Similarly, patients following a prescribed diet were more likely to be testing blood sugar at home - see **Figure 8.5**.

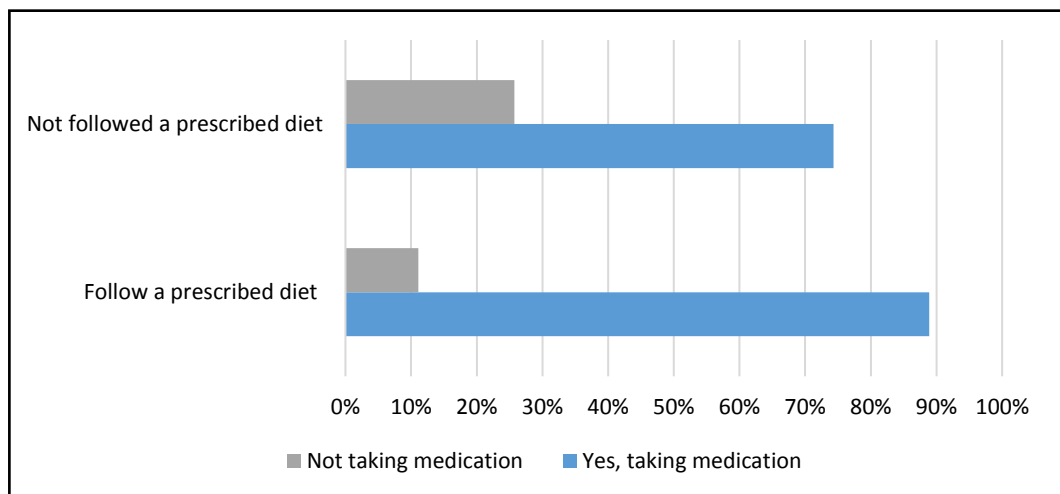


Figure 8.3 Proportion of patients taking medication by prescribed diet status

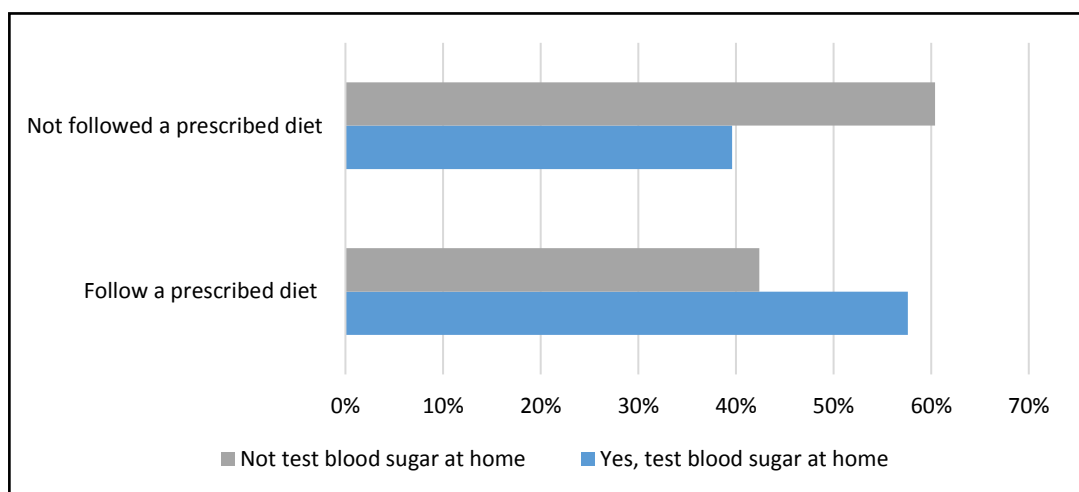


Figure 8.4 Proportion of patients taking medication by home sugar testing status

Intensity of physical activity

The results in **Table 8.8** show the relative risks (RR) from weighted multinomial logit regression models of the association between physical activity level and factors identified as significant after application of the model selection procedure. The results show that low educated patients have a lower likelihood of being physically active compared to highly educated patients. In addition, male patients were more likely to be physically active. **Table 8.9** presents the predicted probabilities of three physical activity levels (high or moderate, low, and none) for these three key factors. It is clear that females, the low educated, and those not currently working are the most likely to display an absence of physical activity. **Figure 8.6** and **Figure 8.7** show the predicted probabilities of the three physical activity levels by education level and working status for diabetes patients. In conclusion, levels of physical activity are clearly affected by socio-economic factors, in particular, patients with lower education levels and who are not currently working appear to be the most physically inactive.

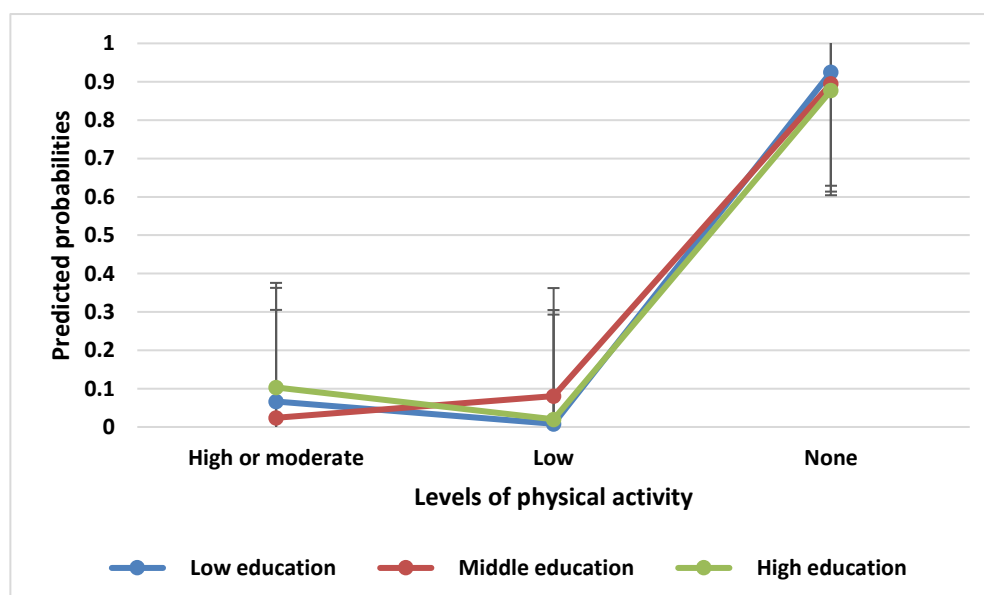
Table 8.8 Relative risk of physical activity levels risk factors

Characteristics	Levels of physical activity			
	High or moderate		Low	
	RR	p-value	RR	p-value
Education level				
Low educated	0.611	0.417	0.396	0.335
Middle educated	0.229	0.006	3.952	0.044
High educated (ref)				
Sex				
Female (ref)				
Male	6.864	0.002	0.216	0.058
Working status				
Currently working	0.850	0.834	0.800	0.831
Not currently working	0.220	0.064	3.053	0.286
Never worked (ref)				
Baseline outcome: None; RR: Relative Risk				

Table 8.9 Predicted probabilities of the likelihood of physical activity levels by different groups

characteristics	Levels of physical activity					
	High or moderate		Low		Not Active	
<i>Education level</i>	PP	p-value	PP	p-value	PP	p-value
Low education	0.0662	0.004	0.0084	0.177	0.9254	0.000
Middle education	0.0240	0.038	0.0809	0.010	0.8949	0.000
High education	0.1028	0.006	0.0201	0.142	0.8771	0.000
<i>Sex</i>						
Female	0.0165	0.062	0.0569	0.069	0.9266	0.000
Male	0.1078	0.000	0.0117	0.073	0.8806	0.000
<i>Working status</i>						
Never worked	0.1122	0.060	0.0129	0.230	0.8748	0.000
Currently working	0.0973	0.003	0.0106	0.144	0.8921	0.000
Not currently working	0.0263	0.020	0.0421	0.035	0.9316	0.000

PP: predicted probability

**Figure 8.5** Predicted probabilities of physical activity levels by education level among Saudi diabetic patients

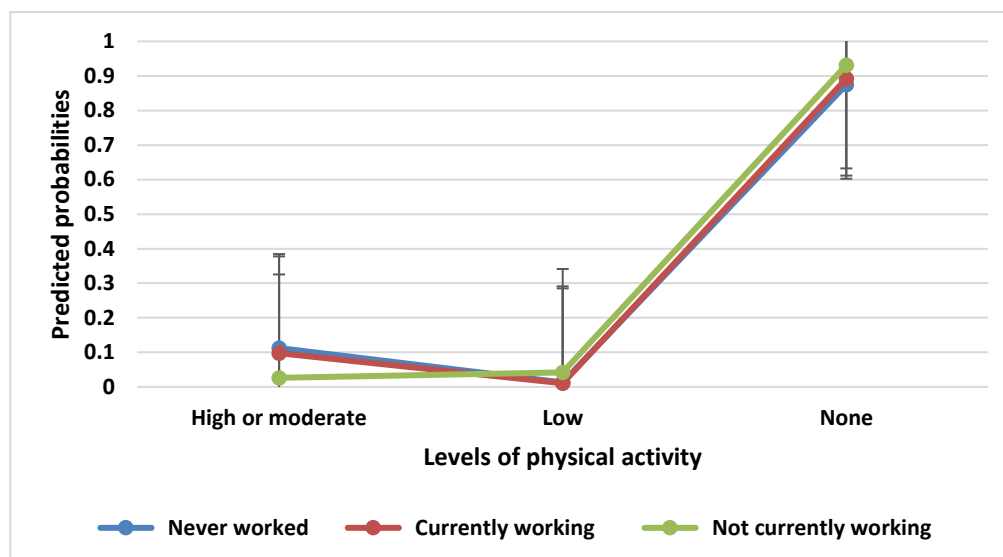


Figure 8.6 Predicted probabilities of physical activity levels by working status among Saudi diabetic patients

8.4 Discussion and conclusions

Understanding the relationship between education level and the 'controlled diabetes' outcome is essential to improve the control and management of diabetes. In this chapter, I aimed to investigate the association between education level and controlled diabetes, in particular how this relationship could be influenced by controlling for demographic, socio-economic (working status and household income), lifestyle, health care access and self-care factors (taking medication, and test blood sugar at home). In doing so, I hoped to gain a better understanding factors affecting the control of diabetes in Saudi Arabia. In addition to socio-economic factors, I also assessed how different patterns of diabetes management affected the 'controlled diabetes' outcome, and how these patterns varied between administrative regions.

This is an important consideration as the administrative region variable can itself be an important predictor of 'controlled diabetes'. This can be rationalised by variation in urbanisation level affecting accessibility to healthy food, availability of public and sport facilities, and the quality of healthcare. Finally, I examined the influence of multiple factors including socio-economic factors on prescribed diet and level of physical activity. The findings from this analysis provided insights as to how socio-economic patterns influence the control and management of diabetes. These will help to expand the evidence base for new prevention programmes and interventions to reduce the high prevalence of uncontrolled diabetes.

8.4.1 Association between education and controlled diabetes

In terms of **controlled diabetes modelling**, although there was no significant association between education level and 'controlled' diabetes, low educated patients had the lowest odds of having controlled diabetes both before and after adjusting the model. This finding is consistent with past studies, which provide evidence that low educated individuals are more likely to exhibit poor glycaemic control compared to highly-educated individuals (Ajlouni et al., 2008, Khattab et al., 2010, Costanian et al., 2014). However, a single Omani study observed the opposite relationship, reporting that highly-educated patients (completed diploma/technical) exhibited poorer glycaemic control than less educated patients (completed high school or lower) (D'Souza et al., 2015). Taken together, the most reasonable explanation of these findings is that in Saudi Arabia, education level does not exert a significant influence on the likelihood of a 'controlled' diabetes outcome.

Considering the effect of **socio-economic factors**, the results indicate that working and income status are not associated with controlled diabetes. This finding is inconsistent with some past studies, which provide evidence that high SES patients are more likely to have controlled diabetes than patients with lower SES (Al-Baghli et al., 2010, Bi et al., 2010, Houle et al., 2016, Alramadan et al., 2018, Fisher et al., 2010). In the Eastern region, for example, the groups exhibiting the poorest control of their diabetes were the widowed and self-employed, whereas the better controlled groups were single with professional employment status (Al-Baghli et al., 2010). Another study in three Saudi cities (Riyadh, Jeddah, and Al-Hofuf) by (Alramadan et al., 2018) observed that patients living in a remote area, housewives and the unemployed were more likely to have inadequately controlled diabetes. However, these studies are not necessarily representative of all Saudi regions.

My results confirm that administrative region is an important predictor of achieving diabetic. Patients who lived in the Western, Eastern or Central regions were more likely to have controlled diabetes than patients in the Northern or Southern regions. These findings are in line with past studies, which provide evidence that the Eastern region, Riyadh (Central region) and Jeddah (Western region) are centres of affluence (AlOmar et al., 2018). These affluent centres are the most urban regions, host the best Universities, and are consequently well-equipped with specialised healthcare services as discussed in the section examining the Saudi Arabian context. Additional characteristics of these regions are that they have better access to healthy food, and more sports and other public facilities for both males and females.

Considering the administrative region is therefore essential, as this variable impacts the control of T2DM: 1) Directly - through significant variation in resources between such as access to diabetes care services and specialised physicians (Grintsova et al., 2014), and 2) Indirectly - through a patient's socio-economic status (Stringhini et al., 2012). Other aspects that vary between regions include healthcare service efficiency and the balance between public and private sector healthcare (Mousa and Aldehayyat, 2018). For example, some less urban regions, such as the Northern borders and Al-Jouf, have better healthcare efficiency than more developed regions such as Eastern region and Makkah (Western region), in both private and public sectors.

One study failed to find an association between region and control of diabetes and is therefore inconsistent with my findings (Alramadan *et al.*, 2018). However, this study focused on three main cities: Riyadh (Central region), Jeddah (Western region), and Al-Hofuf (Eastern region), which are classified as advantaged areas in Saudi Arabia and do not therefore fully represent regional effects. In general, however, glycaemic control is associated with urbanisation level, with higher rates of controlled diabetes in urban areas, for example in the Eastern region (Al-Baghli et al., 2010). I conclude that the administrative region and by implication the level of urbanisation are key factors influencing the control of diabetes. This may be because administrative region is acting as a proxy to the variation in socio-economic factors within and between regions in Saudi Arabia.

In regard to **medication**, the results indicate that patients who are not taking medication are more likely to have controlled diabetes than those who are taking medication. This somewhat paradoxical result is in line with previous studies showing that patients with controlled diabetes are less likely to be on medication than those with uncontrolled disease (Khattab et al., 2010, Alramadan et al., 2018, Al-Hayek et al., 2012). A possible explanation for this finding is that relying on medication alone is suboptimal for managing diabetes among Saudi patients. Effective control of diabetes requires regular monitoring of blood sugar, increased physical activity and following a dietary regimen in addition to medication (Khattab et al., 2010). Given the unexpected relationship between medication adherence and control of diabetes, it was interesting to consider the effect of **dietary intervention** on the control of diabetes. I found that patients who followed a prescribed diet are more likely to have controlled diabetes. This finding is consistent with past studies indicating the importance of diet in achieving glycaemic control (Shamsi et al., 2013). Moreover, diabetic patients who adopt a better diet achieve lower HbA1c levels than patients who rely on insulin and medication alone (Al-Nuaim et al., 1998, Al-Lawati et al., 2012).

One discordant Saudi study reported that patients who used insulin for control of their diabetes had lower HbA1c levels than patients following a combined intervention of diet and medication (Al-Nuaim et al., 1998). One possible explanation for this is that there is a lack of knowledge about management of the disease among Saudi patients. A further factor is the nature of traditional Saudi dishes, such as Kabsah⁶⁰, that rely heavily on rice, meat, and bread. This is compounded by the rapid socio-economic development in the MENA region. The increase in the availability of meat products has contributed to a nutritional transition from the widespread consumption of cereal products, fruit, and vegetables to diets high in fat and protein (Mehio Sibai et al., 2010). Given these effects, it is not surprising that a focus on educating patients to improve their knowledge of diabetes in general and nutritional approaches in particular is an effective management approach resulting in reduced HbA1c values (Redmond et al., 2007, Tayyem et al., 2013). This underlines the importance of prescribed diet as an effective determinant of diabetes control in Saudi Arabia.

8.4.2 Prescribed diet and physical activity level

The results show a clear relationship between the **administrative region** and the likelihood of patients following a prescribed diet. Diabetes patients who lived in the Eastern, Western and Southern regions appeared less likely to follow a prescribed diet compared to the Central region. A key contextual consideration is that only a small proportion of the Saudi population observe dietary recommendations in relation to the consumption of dairy products⁶¹, fruit and vegetables, nuts and fish meat (Moradi-Lakeh et al., 2017). However, this general survey masks considerable regional variation in dietary detail. For example, the Southern and Western regions including Jizan, Asir, Albaha and Makkah are close to the Red Sea and along with the central “Riyadh region” have the highest fish consumption. In contrast, the lowest consumption of fish occurs among individuals who live in the Northern region (Ha’il, Al-Jouf and northern borders) (Moradi-Lakeh et al., 2017). This shows that administrative region can be an important determinant of dietary habits for Saudi individuals.

⁶⁰ **Kabsa**: is traditional Saudi Arabian dish that mainly contained chicken or meat with rice. This dish is frequently serving during week for lunch or dinner meals.

⁶¹ Dairy products: milk, yoghurt

In addition to diet, **physical activity level** is also of critical importance in the management of diabetes. The combination of unhealthy diet and physical inactivity associated with socio-economic change are largely responsible for the rapid increase of diabetes across the GCC region (Davies et al., 2008, Al-Rasheedi, 2015). In general, there is a high prevalence of physical inactivity among all ages and sexes groups in the Saudi Arabia (Alramadan et al., 2018). With respect to the effect of **gender** on physical activity level, the results show that that male patients have a higher level of physical activity, with female patients over-represented in the 'low activity' and 'no physical activity' groups.

These findings are in line with numerous previous studies showing that females are less likely to be physically active (Al-Alfi et al., 2004, Al-Nozha et al., 2004, Midhet et al., 2010, Aljoudi and Taha, 2009, Awadalla et al., 2014, AlQuaiz and Tayel, 2009, Alshaikh et al., 2017). One explanation for this finding is that in Saudi Arabia the weather is warm for most of the year, especially during the summer season; these prolonged high temperatures discourage outdoor physical activity (Al Othaimeen et al., 2007, Al Johani et al., 2015). Furthermore, the urban environment is not well organised for walking and cycling outdoors. As a result there is a very high dependency on car usage, even for short journeys (Giles-Corti et al., 2016).

In addition to these geographical factors, socio-cultural norms and restrictions are powerful factors limiting physical activity (AlQuaiz and Tayel, 2009). This is particularly true for women, as it is less culturally acceptable for them to walk or exercise outdoors. Access to womens' sport facilities are limited in less urban regions and expensive (Al-Lawati and Jousilahti, 2004, Al-Rasheedi, 2015). Another socio-cultural norm that acts against individual physical activity is obesity, traditionally perceived as a sign of beauty and high-income status in females (Al-Nuaim et al., 2012). These two factors are likely to be significant drivers of the high prevalence of obesity and diabetes among Saudi females. In terms of informing policy, it follows that wider provision of public female sports facilities in the GCC region could be a highly cost-effective intervention.

Turning to socio-economic factors in relation to physical activity, the results demonstrate that patients who are highly-educated or who are currently working have a higher probability of being in the 'high' or 'moderate' physical activity groups compared to low educated and other working status groups. In contrast, low educated and not currently working patients have a higher probability of being in the 'low' and 'not physically active' categories. These findings are consistent with previous studies showing that low educated patients also had poor diabetes knowledge compared to highly-educated patients (Ong et al., 2014).

European patients with low SES are also less likely to be physically active, even during leisure-time, compared to those with high SES (Beenackers et al., 2012). It should be stressed, however, that the nature of employment for high SES individuals mean that they may be occupationally less physically active than low SES individuals; such effects could vary between regions depending on employment patterns. The decrease in physical activity with age is more pronounced in individuals who have lost their jobs or are not currently working (Shaw and Spokane, 2008). This lends support to the idea that highly educated patients might be better informed about how to manage their diabetes, including the importance of physical activity. Moreover, patients who are not currently working (i.e. the retired and unemployed) are more likely to be in the lower income levels, this will necessarily limit their access to healthy diet and sports facilities. An additional factor applying to retired patients is that they are more likely to be in the older age groups and less likely to be able to achieve high levels of physical activity. Conversely, however unemployed patients are more likely to be in the younger age groups, they are less likely to be able to afford the cost of access to sports facilities.

In this chapter, I aimed to investigate the association between education level and the control of diabetes. I also examined the interrelated effects of demographic, socio-economic, health care and self-care behaviour factors. In the second analysis, I investigated how lifestyle factors related to socio-economic status and other risk factors. The most important predictors of the controlled diabetes outcome are administrative region, prescribed diet, and taking medication. In addition, I identified administrative region, working status and regular testing for blood sugar as important factors influencing the use of a prescribed diet intervention. This analysis also indicated that education level, gender, and working status are important predictors of physical activity status in Saudi Arabia. Finally, place of residence played a major role influencing lifestyle patterns and the control of diabetes. Overall, the findings highlight the need for a comprehensive region-based regulatory framework for diabetes intervention strategies, necessitating parallel political, social and cultural changes.

Chapter 9 Discussion and conclusions

In this thesis, I aimed to explore the association between T2DM and socio-economic factors in Saudi Arabia, with a particular focus on education level. My plan was to use the SDH framework to examine how different layers of socio-economic characteristics were associated with the prevalence, diagnosis, and control of T2DM. In this way, I hoped to gain a better understanding of the factors behind the increasing prevalence of T2DM in the country. In this chapter I discuss the findings described in the analytical chapters and directly address the research questions presented in Chapter 1. I start by detailing the main contributions of the thesis by analysing the themes that emerged during the study. I then summarise the key findings and discuss their implications in relation to the SDH framework and previous studies. This leads to an appraisal of the strengths and limitations of the study, a consideration of its policy implications and a set of recommendations to address the public health challenge presented by T2DM. I conclude by considering potential areas for future research and present an overall summary of my conclusions.

9.1 Contributions of this thesis

In the last three decades, the GCC region has faced a rapid socio-economic transformation that has resulted in a heavy burden of T2DM, particularly in Saudi Arabia as discussed in the introduction to this thesis. As mentioned in Chapter 3, although the Saudi MOH has addressed this challenge through a national policy plan, this has not been informed by the SDH framework aimed at gaining a better understanding of T2DM in the country. The work described in this thesis begins to fill this gap in our understanding by investigating socio-economic factors in relation to different diabetes outcomes based on a national survey. Additionally, it provides new knowledge about the risk factors for T2DM in Saudi Arabia by exploring how socio-economic factors influence T2DM prevalence, diagnosis, and control. As discussed in Chapter 2, the SDH conceptual framework was used throughout as it provided the most appropriate context for addressing the thesis objectives (Dahlgren and Whitehead, 1991). Based on Saudi Arabian background information and relevant literature reviewed in Chapters two and three, this thesis examined three key outcomes; overall T2DM prevalence, the prevalence of diagnosed as opposed to undiagnosed T2DM, and the extent to which T2DM control is achieved.

These key outcomes were examined in relation to socio-economic factors of T2DM in Saudi Arabia alongside a range of other potential factors. The results of the analytical chapters indicate that there are relationships between T2DM and a set of demographic, socio-economic, lifestyle, healthcare access and self-care behaviour factors. Additionally, administrative region showed significant association with all three diabetes outcomes (Chapters 5-8). The SDH framework guides health outcomes research by considering and addressing SE factors at the micro, meso and macro levels. This thesis can therefore be viewed as an exploratory study into how **education level** affects the prevalence, diagnosis, and control of T2DM using the SDH framework; and the overall applicability of the SDH framework in KSA.

At the end of Chapter 2, I developed three main frameworks based on the SDH framework and previous published findings. Starting with a framework based on the overall prevalence of T2DM, a literature search revealed that socio-economic factors are related to T2DM prevalence in a range of high-income countries. These effects are in addition to well-established demographic, genetic, lifestyle and community factors. However, in this study, I used educational level as a proxy for individual socio-economic status. However, education level is essentially fixed at adulthood, it could still impact on health status in later life. Furthermore, education levels within a large birth cohort might differ depending on the age of the subjects, reflecting economic development (Lynch, 2003). This framework was investigated in Chapter 5. The second framework relates to diagnosed T2DM, and addresses the disparity between overall T2DM prevalence and the prevalence of clinician-confirmed T2DM. Based on previous studies described in Chapters 2 and 3, the framework I adopted examines risk factors that could be related to educational level such as health care access. These factors were explored in Chapter 3 and include healthcare type (private or public), type of clinic, and distance between home and health care facility. Previous studies addressing T2DM susceptibility have concluded that a high proportion of T2DM cases are undiagnosed, and that understanding the determinants of this lack of timely diagnosis is of major importance to public health initiatives.

The final framework - described in Chapter 8 - addressed the relationship between education level and controlled diabetes. This framework includes self-care behavioural factors such regular monitoring of blood sugar. Despite significant improvements in health care in the country, evidence indicates that for many individuals there is still inadequate management of T2DM, as reflected in the high proportion of individuals with uncontrolled diabetes in SA. All these frameworks are SDH based and are aimed at improving our knowledge of how education level relate to the high prevalence of T2DM in Saudi Arabia. This in turn will contribute to the development of policies and interventions by informing strategies based on the identification of individuals who are more likely to develop T2DM or who fail to achieve glycaemic control once diagnosed.

These frameworks were applied to model the impact of education level on the three T2DM outcomes in four analytical chapters. In **Chapter 5**, I explored the relationship between education level and the prevalence of diabetes, and whether this association persists after controlling for demographic, other socio-economic and lifestyle characteristics. I also modelled future trends of diabetes prevalence in Saudi Arabia. **Chapter 6** investigated the association between education level and clinician diagnosed diabetes, and whether this was affected by demographic, socio-economic and health care access factors. This analysis was continued in **Chapter 7**, where I developed a prediction model for undiagnosed diabetes based on whole population data. **Chapter 8** investigated the relationship between education level and controlled diabetes, and how this relationship is controlled by demographic, socio-economic, health care access and self-care behaviour factors. I also examined factors influencing the prescription of diet and physical activity, and adherence to these interventions.

The inspiration for this thesis was to evaluate the impact that education level affected has on T2DM outcomes in the KSA. Through these analyses, I have explored the different pathways that an individual with T2DM might take and highlighted how education level can influence this patient journey. These insights help improve our understanding of which individuals/patients are at higher risk of becoming diabetic, remaining undiagnosed or failing to control their disease after diagnosis. These frameworks are of course linked and directly relevant to Saudi healthcare policies in general and plans for T2DM prevention, early diagnosis, and management. One key factor I identified was the lack of integration between different SDH layers (individual, community, and regional factors).

9.2 Summary of key findings

9.2.1 The association between education level and T2DM outcomes

The following table shows a summary of results for the three main T2DM outcomes and includes the most significant factors. I have identified related to the prevalence, diagnosis, and control of T2DM. Overall, it is obvious from these results that education level was only associated with prevalence of T2DM. However, administrative region emerged as the key risk predictor for all three outcomes. These two factors are high-level in that they reflect a combination of known risk factors such as gender and age, together with important socio-economic cultural influences.

Dissecting the contribution of these various factors will help improve our predictive algorithms for targeting intervention at the level of the individual and developing region-specific policy. The most significant predictors were education, age, employment status, and region and body mass index associated with the prevalence of T2DM. However, for diagnosed diabetes, only age and monthly household income were associated. Finally, region, taking medication and prescribed diet were the only factors associated with controlled diabetes. In fact, I found significant variation in the individual risk of diabetes between regions in all three analytical chapters, emphasising the potential impact of cultural, behavioural and environmental factors that are usually excluded from analysis. This finding underlines the importance of these "hidden" factors, with important implications for integrated diabetes management.

Table 9.1 Summary of three main T2DM outcomes results

T2DM outcomes	Prevalence		Diagnosis		Controlled	
	<i>Bivariate</i>	<i>Final Model</i>	<i>Bivariate</i>	<i>Final Model</i>	<i>Bivariate</i>	<i>Final Model</i>
Study factors						
<i>Education level</i>	+	+	+	-	-	-
Demographic factors						
<i>Age</i>	+	+	+	+	-	-
<i>Sex</i>	+	-	+	-	-	-
<i>Marital status</i>	+	-	+	-	-	-
Socio-economic factors						
<i>Employment status</i>	+	+	+	-	-	-
<i>Monthly household income</i>	-	-	-	+	-	-
<i>Administrative region</i>	+	+	+	-	+	+
Other factors						
<i>Smoking status</i>	+	-				
<i>Physical activity</i>	+	-			-	-
<i>Body mass index</i>	+	+			-	-
<i>Routine check-up</i>			+	-	-	-
<i>Time between home and health care facility</i>			+	-		
<i>Hospital type</i>			+	-		
<i>Taking medication</i>					+	+
<i>Insulin</i>					-	-
<i>Test blood sugar at home</i>					-	-
<i>Prescribed diet</i>					-	+
(+) : significant relationship ; (-) : non-significant relationship						

- ***What is the association between education level and the prevalence of T2DM in Saudi Arabia? More specifically, is this association modified by sex? How is this association affected when controlled for demographic, administrative region, socio-economic and lifestyle factors?***

This was the main question I addressed in chapter five, drawing on the first analytical framework addressing the overall prevalence of T2DM discussed above. Findings from this analysis provided evidence that there is an inverse association between education level and the prevalence of T2DM (as described in conceptual framework). Thus, individuals educated to a low-level have a higher likelihood of being diabetic compared to highly educated individuals. This relationship was expected based on previous studies indicating that individuals educated to a high level are less likely to have diabetes compared to less well-educated individuals (Hwang and Shon, 2014, Murad et al., 2014, Kim et al., 2015). This might be expected, as highly-educated individuals are more likely to be aware of the risks and symptoms of diabetes (Su et al., 2016). It has been argued that increasing awareness of diabetes knowledge would lead to improvements in Saudi public health. The results of Chapter 5 showed that the association between education level and diabetes prevalence was constant through all models even after adjusting for other risk predictors. Additionally, the odds of having T2DM are significantly higher for older age groups, the retired, individuals who were previously married and the obese. The influence of individual age on the prevalence of diabetes has been widely discussed (Al-Rubeaan et al., 2014b). The increased likelihood of diabetes in older-age groups could be due to direct physiological effects of ageing or indirect effects through lack of exercise due to aged-related disability.

Students have a lower likelihood of having diabetes compared to other employment groups. Additionally, obese individuals have a well-documented increased risk of T2DM (Majeed et al., 2014), and it has been argued that this increased risk extends to all individuals with high BMI. The findings also provide evidence that individuals who lived in the Southern region are more likely to be diabetic than those who lived in the Western region (AlOmar et al., 2018). Surprisingly, lifestyle factors such as physical activity and smoking status did not show a significant association with T2DM in this study. This may be due to the fact that in order to understand physical activity and smoking we need longer term data on individuals rather than a snapshot taken at the time of the survey. The factors I tested did not show statistically significant effects in these analyses, potentially due to the small sample size. Nonetheless, some behavioral factors might mediate the association between education level and T2DM outcomes. It has been proposed that individuals with low education level are more likely to be diabetic, and that this might reflect lower levels of physical activity in this group.

Additionally, physical activity itself would be associated with the prevalence of T2DM. Hence, there might be a direct effect of education level on physical activity, and an indirect effect of education level on the prevalence of T2DM mediated by physical activity. However, merging these two hypothetical effects may lead to them cancelling each other out, resulting in a net zero effect. The findings also indicate that BMI, region, and employment status are key factors in mediating the association between education level and the prevalence of T2DM. It has also been suggested that individuals educated to a high level are less likely to belong to older age groups. As older individuals are more likely to be diabetic, age is an important confounding factor for the association between education level and the prevalence of T2DM.

Nonetheless, the significant relationship between education level and prevalence of T2DM underlines the importance of considering socioeconomic status in general and education level in particular when planning T2DM interventions. The existing focus on age, region, and body mass index is supported by the fact that age and BMI were found to be significantly associated with prevalence of T2DM. Consequently, this comprehensive panel of risk factors could be used to highlight specific groups of individuals who are more likely to have diabetes, allowing targeted awareness campaigns to increase knowledge of diabetes and its complications.

- ***What is the association between education level and diagnosed diabetes? Is this association affected by interaction between education level and gender? More specifically, does this relationship persist after controlling for demographic, regional, socio-economic, and healthcare access factors?***

I addressed this question in Chapter 6 using the second framework described in section 9.1 above. This examined the association between education level and the diagnosis of diabetes, and the effect of other risk factors identified from a literature survey. My analysis confirmed that education level was significantly associated with diagnosis of diabetes among individuals in the middle education group (i.e. primary and intermediate school levels). However, this relationship disappeared after adjusting for administrative region and in the fully adjusted model. Individuals in the low education level group were more likely to have diagnosed diabetes than those in the high education level group. As discussed in Chapter 3, educational attainment has dramatically increased among the Saudi population over recent decades.

As a result, stratifying by education level will result in younger adults being over-represented in the high education level group. Thus, less-well educated individuals are more likely to be older and as the odds of having diagnosed diabetes increases with age, so does the likelihood of having diabetic complications. The finding that older-age groups are more likely to have diagnosed T2DM has been reported previously (Bener et al., 2009). This finding indicated that age plays an important role in T2DM diagnosis. This might merely reflect an accumulating risk of disease with time and increased opportunity to identify symptoms and receive a diagnosis. However, it is also likely that the absolute risk of having diabetes will vary between different birth cohorts due to changes in exposure to risk factors over time (Holford, 1991).

In addition to the effect of age, the analysis presented in Chapter 6 identified household income effects on the likelihood of T2DM diagnosis. I found that individuals from low income households are more likely to be diagnosed with diabetes. This effect of low income has been reported previously (Sommer et al., 2015), however it has been argued that individuals with high SES have a higher likelihood of early diagnosis. This is on the basis that household income is associated with the ability to access better healthcare and healthier foods that can improve health and reduce the risk of T2DM. However, this runs contrary to my findings in a study sample focused on diagnosed individuals, as none of the healthcare access factors, I examined significantly influenced the association between education level and diagnosed diabetes. The findings indicate that age and monthly household income are the main predictors of the association between education level and the diagnosis of T2DM. Individuals who are older and from low income households are more likely to have previously diagnosed T2DM compared to younger-aged and high individuals, pointing to the importance of age as a potential confounder in this analysis.

- ***What factors relate to effective control of diabetes in individuals, with specific focus on educational level? How do demographic, administrative region, health care access, self-care, and other socio-economic factors related to controlled diabetes?***

I addressed these questions in Chapter 8. I first carried out a literature review of the known factors affecting control of T2DM and used these as the basis for the final framework investigating the impact of education on the likelihood of T2DM patients achieving and maintaining glycemic control. By analysing the association between education level and controlled T2DM, I observed that administrative region, taking medication and following a prescribed diet are the main factors influencing control of T2DM.

Specifically, patients who live in the Southern region and those who were not following a prescribed diet have higher risk of uncontrolled T2DM compared to other groups. This is also the case for patients who only take medication without undertaking any of the other interventions. Although previous studies have pointed to a relationship between individual socioeconomic factors and inadequate control of T2DM, I could not find any such association. For example, it has been argued that individuals educated to a high-level are more likely to achieve control of their blood sugar. However, my results show no relationship between education level and control of diabetes.

It seems, therefore, that while **control** of T2DM is primarily influenced by regional socioeconomic factors, T2DM **prevalence** and **diagnosis** are determined more by individual socioeconomic factors. The importance of regional factors in determining control of T2DM means that the SMOH and government should target regional needs, particularly in relation to healthcare service provision, access to a healthy diet and facilities for everyday exercise. My analysis revealed that there are significant socioeconomic variations between regions, particularly in terms of education, employment, and income status. These differences are important drivers of inequality between regions (Mousa and Aldehayyat, 2018).

Most patients relied on a combination of prescribed diet and medication for managing their diabetes. However, I was unable to find a significant association between T2DM management approach and control of diabetes, as there were no obvious differences between the controlled and uncontrolled groups with respect to the form of intervention. However, I did discover that intervention strategies varied significantly between regions. Patients who lived in the Central region were more likely to be following a diet, taking medication, and using insulin for managing T2DM than other regions.

In the second analysis described in Chapter 8. I set out to investigate factors influencing the use of behavioral interventions. This revealed that administrative region is the main predictor of whether patients follow a prescribed diet. In relation to physical activity interventions, I found that patients with a low education level, females, and patients who are currently out of work have a higher likelihood of being physically inactive. This observation is in agreement with that of Alshaikh et al. (2017) who found that females in the GCC region have a higher prevalence of physical inactivity.

As mentioned in previous chapters, it should be noted that some of the associations on which these observations are based were not statistically significant due to the influence of sample size and consequent reduction in statistical power (in Chapter 8). Nonetheless my findings highlight the need to gain a better understanding of the impact of socioeconomic and lifestyle factors on T2DM control.

Such understanding will necessarily inform development of the framework for the prevention and control of T2DM across the GCC, particularly in Saudi Arabia. My observations highlight the need for a better understanding of the impact of socioeconomic and lifestyle factors on T2DM control. Such understanding will be key to develop the T2DM framework for the prevention and better control of T2DM across the GCC, particularly in Saudi Arabia.

It is important to highlight the possibility that for the previously reported associations between education level and the three diabetes outcomes, other risk factors could be confounders or mediators. For example, confounding factors that influence both education level and T2DM outcomes include age and gender, while mediating factors could enhance the association. The findings of this chapter demonstrate that administrative region is a key predictor of the association between education level and achieving control of T2DM.

9.2.2 Future projection of T2DM prevalence estimates

- ***What are the future estimates for T2DM total prevalence in Saudi Arabia? How do these estimates vary by age and sex of the individuals?***

Another key aim for this thesis was to predict future trends of T2DM prevalence stratified by age and gender. Based on the modelling presented in Chapter 5, I estimate that the prevalence of T2DM will increase from 16.9% to 21.8% between 2015 and 2030 among the Saudi population aged 25 years and above. In males, the estimated increase was from 18.22% in 2015 to 23.29% in 2030, higher than in females (from 15.88% to 20.34%) over the same period. These findings are in line with those reported by (Whiting et al., 2011), who estimated the global prevalence of diabetes in all age groups for 2011 and 2030. It is important to note that in this study, all estimates of T2DM prevalence in males were higher than for females across nearly all age groups.

This picture of consistently higher T2DM prevalence in females might be a reflection of higher obesity rates among Saudi women. This would be in line with previous studies, which clearly showed that females in the GCC region and EMR have a higher prevalence of obesity (Mehio Sibai et al., 2010). This gender difference in obesity rates has been attributed to a greater tendency for females in the region to adopt an inactive lifestyle, particularly in the GCC countries. This may not be true for high-income groups, where there is a trend for males to display a greater prevalence of physical inactivity, as argued in Chapter 3 (Alshaikh et al., 2017). These trends of inactivity are another example of the unique cultural context provide by the GCC countries, affecting predominantly female citizens.

9.2.3 Prediction models for undiagnosed T2DM, diabetes and pre-diabetes

- ***Which socio-economic factors should be included in a prediction model that yields a risk score to identify individuals who are at risk of having undiagnosed or pre-T2DM? Do socio-economic and health care factors contribute to prediction models of undiagnosed diabetes and high-risk of diabetes in the whole population excluding those with diagnosed diabetes?***

In Chapter 7, I focused on developing prediction models for undiagnosed diabetes and pre-diabetes and comparing the finding with other related studies. Part of this involved testing whether education level, socio-economic and health care access factors contributed to the models. However, my findings highlighted that demographic, health and environmental factors are more significant than education level and other socio-economic factors.

In particular, I showed that regional effects played a key role and argued that consideration of administrative region should be incorporated in public health policy aimed at reducing the impact of undiagnosed diabetes. This builds on my observations in Chapter 3 and 5 that the Northern and Southern regions have the highest diabetes prevalence. In general, the large differences between the results in this thesis and other studies are mainly attributable to the different predictors used to generate the undiagnosed T2DM model. The majority of previous models have relied on demographic, lifestyle and health-related factors, and been validated using the two-sample approach. In contrast, I subjected the models in this thesis to internal validation based on bootstrapped samples. Findings from this analysis provided evidence supporting previous findings that the main predictors of undiagnosed diabetes are age, body mass index (Al-Lawati and Tuomilehto, 2007, Gray et al., 2010, Ko et al., 2010, Sulaiman et al., 2018) and administrative region.

However, the key predictors of the combined undiagnosed diabetes and pre-diabetes outcome are, age, gender, body mass index, administrative region, and hypertension. Thus, individuals who are older, male, and obese, live in the Southern region and are hypertensive have the highest risk of undiagnosed T2DM. I was unable to identify an association between education level and any socio-economic predictor for undiagnosed diabetes and pre-diabetes; however, administrative region was associated with this outcome, and this must be in part a reflection of the health system of the region.

These findings parallel previously identified risk predictors in GCC and high-income countries, particularly in terms of age and body mass index for undiagnosed diabetes. Older and obese individuals have higher risk of undiagnosed diabetes compared to other groups. Additionally, this study highlighted that where an individual lives is an important risk predictor. No previously published models highlighted the importance of administrative region as a predictor for undiagnosed diabetes. However, my findings are in line with previous studies if administrative region is considered as a proxy for SES, with individuals from areas of deprivation and low SES having a higher likelihood of being diabetic compared to those from more privileged areas.

The socio-economic status of a region is important because disadvantaged and poorer regions are likely to have limited health care services (AlOmar et al., 2018). The only differences I identified in the risk predictors for undiagnosed diabetes and pre-diabetes relate to gender and hypertension, with males and hypertensive individuals having a higher likelihood of undiagnosed pre-diabetes. My conclusion is that improvements in the early diagnosis of T2DM in Saudi Arabia should focus on body mass index, age, and administrative region as the three main predictors used to identify individuals for targeted screening and/or interventions.

9.3 Strengths and limitations of the study

This study has several strengths. **First**, this study is one of the few studies that have assessed the association between socio-economic factors and T2DM outcomes in Saudi Arabia. Furthermore, the study used a data from a Saudi national survey that provided an excellent source of information on non-communicable disease in the Kingdom.

Second, in terms of future projections of T2DM prevalence (Chapter 5), previous studies have provided useful information on trends in Saudi Arabia. However, only one of these studies attempted to make future projections of diabetes prevalence, and there are few such studies in the wider GCC region. Furthermore, the current study used different methodology and a different starting point for the projection, producing a valuable independent estimate for comparison with these other studies.

Third, this study explored T2DM trends in Saudi Arabia at a national level, which is important for public health planning and development of prevention strategies. It provides updated projections of T2DM prevalence and covers all regions in the country.

Fourth, the study presents new evidence supporting the use of administrative region as a community socio-economic risk factor to identify individuals at high-risk of developing T2DM or presenting with pre-diabetes. The T2DM risk model described in Chapter 7 is a simple and inexpensive tool for screening high-risk individuals. This has the potential to generate significant population health benefits in Saudi Arabia, where 43% of cases identified from the population survey were unaware of their condition. To my knowledge, this is the first Saudi study to use administrative region for predicting undiagnosed T2DM and pre-diabetes.

Fifth, the study explored the most important predictors of controlled diabetes and the relationship between lifestyle patterns and SES. The findings clearly determined the importance of a patient's place of residence place in determining control of diabetes.

Although this study has produced interesting findings with particular relevance to Saudi Arabia, several limitations should be noted. **First**, the cross-sectional study design limited its ability to identify association between education level and diabetes outcomes as well as other factors. **Second**, high missing data rates and inaccurate Hb1Ac values affected my ability to measure diabetes outcomes and hence the key estimate of diabetes prevalence. Additionally, the monthly household income variable was badly affected by missing data. The availability of measured blood sugar values would have significantly improved the dataset and enabled a better understanding of the factors behind the high burden of diabetes.

The study sample was based on Saudi citizens only and excluded individuals living in the Tabuk region. This was due to the high frequency of missing HbA1c values from this region in the SHIS database. Therefore, the findings of this study were not completely nationally representative due to the omission of the Tabuk region.

Third, some lifestyle variables such as physical activity and diet relied on the household questionnaire and were based on self-reported questions that are likely to be sensitive to a social desirability bias. This limited my ability to analyse and understand the differences in lifestyle patterns between socio-economic groups. **Fourth**, I assumed a linear trend for the projected increase in T2DM prevalence when addressing the second research question. The diabetes trends used in this study were from total population lifetables and population projections, which may lead to error in the estimates for prevalence and number of cases.

In terms of population estimates, I used the Saudi total population lifetable and assumed this applied to the Saudi national population. In addition, I assumed that the 2016 lifetable (the most recent available from the WHO) is applicable to 2017 and 2018. Furthermore, the future projection of T2DM prevalence only covered the Saudi national population due to the lack survey data relating to diabetes trends among non-nationals.

Fifth, the diabetes diagnostic criteria used differed from those of previous prevalence estimates. As a result, some cases might be classified differently due to changes in the cut-off point/threshold and lead to differences in absolute future estimates. However, it is important to highlight that this study was not intended to provide an accurate estimate of the prevalence and number of people with diabetes but rather to predict the scale of the increase in prevalence and potential for societal impact. Without access to longitudinal data, the risk model can only detect association with disease prevalence as it cannot identify incident cases in individuals.

Sixth, health care access factors such as hospital type and health care provider were based on the household questionnaire and were thus self-reported, which might introduce bias in the analysis described in Chapter 6. **Seventh**, there was no other Saudi population data set available with which to externally validate the prediction risk score model described in Chapter 7. Moreover, our database did not include widely used predictors such as waist circumference, and the family history of diabetes variable had missing values. The risk models were limited by the fact that the study used only one measure as the gold standard for participants. Other datasets that may have been useful are not in the public domain for analysis, such as the Saudi World Health Survey Plus. **Last**, for the analysis described in Chapter 8, the sample size was small, not nationally representative, and likely biased by differences between those who completed the all the questions and those who did not.

9.4 Policy implications and recommendations

The research described in this thesis has focused on the socioeconomic, geographical and behavioural factors associated with T2DM outcomes. It is my hope that this will contribute to a better understanding of the risk factors associated with T2DM in the Kingdom. To date, studies on T2DM in Saudi Arabia have mainly concentrated on demographic and health-related factors, and few have included education level and region as key factors as discussed in Chapter 2. This thesis has several policy implications and I list a number of key recommendations here.

As extensively discussed in **Chapter 5**, my findings clearly point to the significance of education level as a factor influencing the prevalence of T2DM, conditioned by other risk predictors such as age, employment status, administrative region, and obesity. The results clearly indicate that education level, age, employment status, obesity and region have significant association with the prevalence of T2DM in Saudi Arabia. The national diabetes prevention and control program should therefore focus on individuals in older age-groups, those with high BMI, students and individuals who live or have lived in high-risk regions as I identified these as key factors contributing to the association between education level and T2DM. In addition to such a targeted approach, there is also an urgent need for comprehensive policy interventions covering the whole population with a view to disease prevention and increasing awareness of T2DM. Promoting intervention within school, university, and workplaces is important to modify behaviour, especially with respect to nutrition education and exercise programs regardless of region (Badran and Laher, 2012). Furthermore, the policies and programme interventions should consider high risk groups defined by employment status, BMI, and administrative region.

This chapter also developed a Saudi Arabian projection for T2DM over the next ten years. This showed that the prevalence of the disease will increase steadily in both males and females across most age groups. This has profound implications for public health policy in Saudi Arabia, requiring an urgent response at national level to plan and implement strategies aimed both at disease prevention and at reducing the impact of the disease. Policy makers will need additional databases and modelling frameworks for accurate projections of diabetes prevalence to guide health policy and decision-making. I therefore recommend that future projections of T2DM prevalence should take account of employment status, obesity, age, and regional variation. Overall, the chapter provides Saudi health policy makers with reliable estimates of the extent of the T2DM problem. These can act as benchmarks against which to measure the effect of targeted interventions to reduce the high prevalence of diabetes in Saudi society and blunt the impact of these projections.

In **Chapter 6** I presented evidence for an association between T2DM diagnosed status and education level after adjusting for demographic, regional, and health care access factors. My findings indicate that age and monthly household income are important factors influencing the diagnosis of T2DM. This implies that intervention programmes should target all socio-demographic groups, especially younger age groups and those with low household income as they are at highest risk of having undiagnosed diabetes. I discussed the importance of supporting a nationwide population health campaign focused on early diagnosis of pre-diabetes and diabetes, including the introduction of a 'produce tool' to help individuals make healthier dietary choices. Furthermore, the prototype screening tool that I have developed based on a combination of predictors could be

applied to reduce the incidence of T2DM in Saudi Arabia through the identification of high-risk individuals.

In **Chapter 7**, the two main outcomes were 1) T2DM, both diagnosed and undiagnosed, defined by Hb1ac levels or 2) prediabetes, defined by a lower range of HbA_{1c} cut-off values. Most previous studies considered only diagnosed diabetes; however, I chose to develop separate prediction models for both undiagnosed diabetes and prediabetes to identify not only individuals at increased risk of undiagnosed diabetes, but also those at high-risk of going on to develop diabetes i.e. individuals with prediabetes. Findings from this chapter indicate that having undiagnosed diabetes is significantly associated with age, BMI, and region. In addition to these predictors gender and hypertension also show association with undiagnosed pre-diabetes. My conclusion is that used appropriately, a combination of these predictors could play a key role in the early detection of diabetes.

Findings from **Chapter 7** clearly identify an important effect of place of residence on the identification of undiagnosed T2DM cases. Introduction of regular screening will therefore be essential to reduce the number of individuals with undiagnosed T2DM. The early identification of high-risk individuals is of critical importance to deflect progression to T2DM and the onset of its expensive and debilitating complications. This screening should not be restricted to primary health care and hospitals. Screening centres should be readily accessible in public places such as schools, universities, shopping centres, and supported by a National information campaign to promote awareness.

Implementation of a screening programme should commence with pilot studies in each region stratified by nationality to ensure that they are representative. However, it may be necessary to develop region-specific prediction models due to variation in socioeconomic status and urbanisation level. Furthermore, model development has revealed the importance of using standardised cut-off points for HbA_{1c} values defining pre-diabetes or confirmed T2DM to ensure consistency of findings between populations. Such a two-pronged T2DM prevention program based on health prevention and targeted screening would need to consider the impact of regional variation on diagnosis of T2DM observed in this study.

Finally, **Chapter 8** explored the association between education level and control of diabetes highlighting the importance of administrative region, and the role of medication and behavioural modification strategies. My findings show that the key environmental and lifestyle factors that impact on T2DM prevalence vary according to the socioeconomic status of a region. It follows that some regions may benefit more from intervention programs. For example, the Southern region

appeared to be most in need of extra interventions such as patient education programs, increased investment in public sports facilities, and campaigns and/or fiscal policies promoting healthy diets.

A key recommendation is therefore that future research should address in more detail the regional effects of socioeconomic factors on lifestyle patterns in Saudi Arabia. This would enable programmes of improvement in environmental and lifestyle factors to be targeted by region (as described above for the Southern region). However, our ability to devise appropriate strategies is hampered by the limited availability of data on dietary patterns and the causes of physical inactivity at community level. Any national initiative must be flexible enough to take account of important differences in context at the regional and community level. Such approaches have proved an effective alternative to dietary advice in the UK and USA (Herman et al., 2006, Burr et al., 2007) and the UAE Ministry of Health has recently adopted such measures through the imposition of a tax on high fat foods and soft drinks⁶². Recently, the MOH drove a change in Saudi national policy leading to the introduction of a 50% sugar tax on all sugary drinks from 1st December 2019⁶³.

Overall, the thesis offers new insights that will aid Ministry of Health planners to consider regional variation across the Kingdom. Further research should address the impact of socioeconomic factors in each region and variation in lifestyle patterns across Saudi Arabia. These studies should examine specialised T2DM clinical services and healthcare access. In parallel, and as a high priority, the MOH should develop a policy for diet and physical activity that encourages behavioural change. As part of such initiatives, a particular priority should be awareness programmes targeted at the less well-educated group as they are at particular risk of T2DM. Finally, these targeted interventions should be in addition to, rather than replacing, nationwide campaigns promoting healthier lifestyles for all.

9.5 Potential area for future research

The work described in this thesis has raised several questions and lines of enquiry that merit further investigation. In the current study, I investigated the association between education level and the prevalence, diagnosis, and control of T2DM based on nationwide data. The findings show the importance of socio-economic factors at both individual and regional levels, and these should be further explored in future national surveys.

⁶² Arab Health Market Series (2018) *Diabetes, the Pulse: ninth Edition*. Available at: <https://www2.colliers.com/en-AE/Research/Dubai/Diabetes-Arab-Health-Market-Series-The-Pulse-9th-Edition> (Accessed 20 December 2018)

⁶³ <https://www.arabnews.com/node/1106726/business-economy> ; <http://saudigazette.com.sa/article/583496>

Comparable national surveys in high-income developed or developing countries have pointed to the value of stratifying populations based on socio-economic criteria. As well as a more comprehensive analysis of such factors, future surveys should aim to be more comprehensive, including Saudi non-nationals and information on ethnicity. One measure of the importance of this is the recent growth in the number of non-national workers from South and Southeast Asia in the country, as they have an inherently higher-risk of developing T2DM (De Bel-Air, 2014). Future research should also address the variation in the quality of healthcare for each region, as the present study could not disentangle health care access factors from demographic, socio-economic and environmental factors due to lack of data and small sample sizes. In particular, further research into the impact of access to health care and employment status on the risk of undiagnosed and diagnosed T2DM could help refine the predictive model and give insights into policy improvements, as would investigating the effect of the non-Saudi population on the validity of the model. A limitation of the current study is the lack of independent population data with which to validate the prediction models. It would therefore be useful to establish collaboration with other GCC countries, with similar socio-economic characteristics, in order to externally test the models.

Finally, place of residence has a significant influence on lifestyle patterns and the likelihood of developing T2DM, failure to achieve a timely diagnosis or failure to achieve adequate control of blood glucose levels. Spatial analysis of risk factors on a regional and community level could therefore be valuable in developing appropriate and effective strategies for the reduction of T2DM diabetes in the country. It would be interesting design a translational research program around pilot behavioural interventions based on diet and physical activity in regions identified as deprived.

9.6 Conclusion

T2DM is a major public health concern in Saudi Arabia that demands a correspondingly large response from individuals and the health care system. The work described in this thesis has confirmed the significance of the relationship between education level and prevalence of diabetes. It has also identified the socioeconomic and lifestyle factors that play a critical role in the pathway connecting education level and the prevalence of T2DM. The increasing prevalence of diabetes and rooted in the growing socioeconomic gap is now an essential issue for Saudi Arabia, and this study is one of the first attempts to forecast T2DM prevalence in the country. My data shows that, without significant government intervention, the burden of diabetes will increase significantly over the next ten years, and that Saudi Arabia will remain among the top ten countries for diabetes prevalence worldwide. Recent Saudi studies have updated estimates of T2DM prevalence. My

estimates are slightly lower than these but are consistent with the global diabetes projections from the International Diabetes Federation.

My expectation is that without significant intervention, ageing and other demographic changes will interact with increases in development, urbanisation and lifestyle changes, adding still further to the already high burden of diabetes. This thesis has also shown that age and monthly income are the most important risk factors for individual being diagnosed T2DM once it has been developed. Furthermore, this thesis was successful in developing prediction risk models for undiagnosed T2DM and pre-diabetes. This aspect of the work identified age, body mass index, and region as the most significant predictors for undiagnosed T2DM. The most important predictors of the composite outcome of undiagnosed T2DM/pre-diabetes are age, body mass index, region, gender, and hypertension. The work presented in this thesis shows that the most important predictors for controlling the diabetes outcome are administrative region, prescribed diet and taking medication. In addition, my analysis revealed associations between prescribed diet and administrative region, working status and regular home-testing for blood-sugar. This analysis also indicated that education level, gender, and working status are important predictors of physical activity level in Saudi Arabia.

Overall, the most obvious finding to emerge from this analysis is that administrative region is a key factor influencing the prevalence, diagnosis, and control of diabetes in the Saudi population. This highlights the need for a national T2DM strategy that is tailored to regional requirements, supported by a comprehensive regulatory framework. This will require political, social, and cultural changes. Limited information is available on future trends for diabetes prevalence in Saudi Arabia, however the estimates from my modelling are that the epidemic of T2DM will continue to grow with profound implications for the healthcare budget and population well-being. It will be important to conduct a regular national survey in order to monitor the trajectory of disease and gain a better understanding of how the risk factors for the disease interact. The findings of this thesis underline the importance of urgent action by government to deflect the societal impact of the current seemingly inexorable rise in the prevalence of T2DM in Saudi Arabia.

Appendix A Saudi Health Interview Survey Questions

Question	Response	Variable name
Module 2: Selected adult questionnaire for full survey		
Section 1: Sociodemographic information		
Sex (Record Male / Female as observed)	Male=1 Female=2	Sex
How old are you?	Age, in years=1 Specify in open text field Don't know=77 Decline to respond=88	Age
What is the highest level of education you have completed?	Can't read or write=1 Can read and write=2 Primary school completed=3 Intermediate school completed=4 High school completed=5 College/University completed=6 Post graduate degree=7 Technical training=8 Don't know=77 Decline to respond=88	Education level
What is your marital status?	Never married =1 Currently married=2 Separated=3 Divorced=4 Widowed=5 Don't know=77 Decline to respond=88	Marital status
Which of the following best describes your main work status over the past 12 months?	Government employee=1 Non-government employee=2 Self-employed=3 Non-paid=4 Student=5 Homemaker=6 Retired=7 Unemployed (able to work)=8 Unemployed (unable to work)=9 Don't know=77 Decline to respond=88	Working status
Can you give an estimate of the monthly household income if I read some options to you?	Less than 3000 Riyal=1 3000 Riyal to less than 5000 Riyal=2 5000 Riyal to less than 7000 Riyal=3 7000 Riyal to less than 10000 Riyal=4 10000 Riyal to less than 15000 Riyal=5 15000 Riyal to less than 20000 Riyal=6 20000 to less than 30000=7 30000 Riyal or more=8 Don't Know=77 Decline to respond=88	Household monthly income (HMI)

Appendix A

Question	Response	Variable label
Section 2: Tobacco Use (Smoking)		
Have you ever smoked any tobacco products, such as cigarettes, cigars, pipes or Shisha?	Yes=1 No=0 If No, go to tobacco_smokeless Don't know=77 If Don't know, go to tobacco_smokeless Decline to respond=88 If Decline to respond, go to tobacco_smokeless	Smoking status One variable of smoking habit will be recoded into three categories: currently , ex-smoker, never
Do you currently smoke any tobacco products, such as cigarettes, cigars, pipes or Shisha?	Yes=1 No=0 If No, go to tobacco_age_start Don't know=77 If Don't know, go to tobacco_age_start Decline to respond=88 If Decline to respond, go to tobacco_age_start	The interaction between both question could be highlighted the ex-smoker category (yes +No) Never (No+No) Current (yes+yes)
Section 3: Diet		
In a typical week, on how many days do you eat fruit?	Number of days=1; Don't Know=77; Decline to respond=88	In a typical week, on how many days do you eat fruit?
How many servings of fruit do you eat on one of those days?	Number of servings=1 Don't Know=77 Decline to respond=88	How many servings of fruit do you eat on one of those days?
In a typical week, on how many days do you eat vegetables?	Number of days=1 Don't Know=77 Decline to respond=88	In a typical week, on how many days do you eat vegetables?
In a typical week, on how many days do you eat fruit?	Number of days=1 Don't Know=77 Decline to respond=88	In a typical week, on how many days do you eat fruit?
How many servings of vegetables do you eat on one of those days?	Number of days=1 Don't Know=77 Decline to respond=88	diet_vegetables_servings
In a typical week, on how many days do you eat red meat?	Number of days=1 Don't Know=77 Decline to respond=88	(diet_redmeat)
How many servings of red meat do you eat on one of those days?	Number of days=1 Don't Know=77 Decline to respond=88	(diet_redmeat_servings)
In a typical week, how many days do you eat poultry meat (including chicken)?	Number of days=1 Don't Know=77 Decline to respond=88	diet_poultry
How many servings of poultry meat (including chicken) do you eat on one of those days?	Number of days=1 Don't Know=77 Decline to respond=88	diet_poultry_servings
In a typical week, on how many days do you eat processed meats such as sausage, or other packaged cold cuts, ...?	Number of days=1 Don't Know=77 Decline to respond=88	(diet_procmeat)
How many servings of processed meats do you eat on one of those days?	Number of days=1 Don't Know=77 Decline to respond=88	diet_procmeat_servings
In a typical week, how many days do you eat other processed foods, such as fast foods, canned foods, packaged entrees, or soup?	Number of days=1; Don't Know=77; Decline to respond=88	diet_procfood
How many servings of other processed foods do you eat on one of those days?		diet_procfood_servings
In a typical week, how many meals do you eat that were not prepared at a home? By meal, I mean breakfast, lunch and dinner.		diet_meals_out
In a typical week, on how many days do you eat fast food?		(diet_fastfood)
What do you think about fast food?	Healthy food=1; Not healthy=2; Don't Know=77; Decline to respond=88	diet_fastfood_healthy

Question	Response	code
Section 5: Physical activity		
<u>Recreational activities:</u>		
Do you do any vigorous-intensity sport, fitness or recreational ...?	Yes=1 No=0 ; If No, go to rec_modactivity Don't know=77 ; If Don't know, go to rec_modactivity Decline to respond=88 ; If Decline to respond, go to rec_modactivity	(rec_vigactivity)
In a typical week, on how many days do you do vigorous-intensity sports, fitness or recreational activities?	Number of days=1; Don't Know=77; Decline to respond=88	(rec_vigactivity_days)
How much time do you spend doing vigorous-intensity sports, fitness or recreational activities?	Hours per day=1 Specify in open text field Minutes per day=2 Specify in open text field Don't know=77 Decline to respond=88	(rec_vigactivity_time)
Do you do any moderate-intensity sport, fitness or recreational ...?	Yes=1 No=0 If No, go to P16 Don't know=77 If Don't know, go to P 16 Decline to respond=88 If Decline to respond, go to P 16	(rec_modactivity)
In a typical week, on how many days do you do moderate-intensity sports, fitness or recreational activities?	Number of days=1; Don't Know=77; Decline to respond=88	(rec_modactivity_days)
How much time do you spend doing moderate-intensity sports, fitness or recreational activities?	Hours per day=1 ; Specify in open text field Minutes per day=2 ; Specify in open text field Don't know=77; Decline to respond=88	(rec_modactivity_time)
<u>Sedentary behaviour:</u>		
How much time do you usually spend sitting or reclining on a typical day?	Hours per day=1 Specify in open text field Minutes per day=2 Specify in open text field Don't know=77 Decline to respond=88	(sitting)
In a typical week, how much time do you usually spend in front of the television or on the computer?	Enter number of hours _____=1 Don't know=77 Decline to respond=88	(tv_comp_time)
Section 6: Access to and utilisation of health care		
In what year did you last visit a doctor or other health professional for a routine check-up?	1=Enter year (YYYY): ____ Specify in open text field 995=Never; 77=Don't know; 88=Decline to respond	Check-up routine year
How far away was the facility you attended from your home?	1=Enter distance (km): ____ Specify in open text field 77=Don't know; 88=Decline to respond.	Check-up distance
How long did it take you to get there?	1=Enter number of hours: ____ Specify in open text field 2=Enter number of minutes: ____ Specify in open text field 77=Don't know; 88=Decline to respond	(hc_checkup_time)
What was the type of the hospital that you last visited for medical attention for any reason?	Governmental(MOH)=1 National Guard=2 Armed Forces=3 Military hospital=4 Private hospital=5 Other =6 Don't know=77 Decline to respond=88	Hospital type
In what year did you last visit a hospital for medical attention for any reason?	1=Enter year (YYYY): ____ Specify in open text field 995=I have never been to a hospital for medical attention. If Never, go to hc_clinic_yr 77=Don't know; 88=Decline to respond	

Appendix A

Question	Response	Variable label
Section 8: History of chronic conditions		
<u>Diabetes mellitus:</u>		
Have you ever had your blood sugar measured by a doctor or other professional?	Yes=1 No=0 Don't know=77 Decline to respond=88	Blood_sugar_measured
Has a doctor or other health professional ever told you that you had diabetes mellitus, otherwise known as diabetes, sugar diabetes, high blood glucose, or high blood sugar?	☐ 1=Yes ☐ 2=Yes, but only during pregnancy (skip to # / predm) ☐ 0=No (skip to # / predm) ☐ 77=Don't know (skip to # / predm) ☐ 88=Decline to respond (skip to # / predm)	Dm_diganoosed
What type of diabetes do you have?	☐ 1=Type 1 ☐ 2=Type 2 ☐ 77=I don't know ☐ 88=decline to repond	Diabetes type
How old were you when you first receive this diagnosis?	☐ 1=Enter age: ____ Specify in open text field ☐ 77=Don't know ☐ 88=Decline to respond	Age
In what year did you first receive this diagnosis?	☐ Enter year (YYYY): ____ Specify in open field text ☐ 77=Don't know ☐ 88=Decline to respond	Year of first diagnosed
During the past 30 years, or since your diagnosis, have you ever taken mediation for this condition?	☐ 1=Yes, currently taking medication ☐ 2=Yes, previously took medication, but not currently ☐ 0=No, never took medication ☐ 77=Don't know ☐ 88=Decline to respond	Medication
In the last month, did you visit a physician or other health professional for the management of your diabetes?	Yes=1; No=0 ; Don't know=77 ; Decline to respond=88	(dm_management)
Insulin	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_insulin)
Drugs(medication) that you have taken in the past two weeks	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_medications)
Special prescribed diet	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_diet)
Have you ever seen a traditional healer for diabetes or raised blood sugar?	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_healer)
Are you currently taking any herbal or traditional remedy for your diabetes?	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_herbalmed)
Do you test your blood sugar at home?	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_bloodsugar_home)
Is there a family history of diabetes, do your parents, children, brothers or sisters suffer from diabetes?	Yes=1; No=0; Don't know=77; Decline to respond=88	(dm_fmhx)

Pre-diabetes			
Question	response	code	Variable label
Has a doctor or other health professional ever told you that you had pre-diabetes mellitus, otherwise known as pre-diabetes, borderline diabetes, impaired fasting glucose, impaired glucose tolerance, or impaired sugar tolerance?	☐ 1=Yes ☐ 2=Yes, but only during pregnancy ☐ 0=No ☐ 77=Don'tknow ☐ 88=Decline to respond	(predm)	Predm_diagnosed
In what year did you first receive this diagnosis?	☐ Enter year (YYYY): ____ Specify in open field text ☐ 77=Don'tknow ☐ 88=Decline to respond	(predm_dx_yr)	
Hypertension (blood pressure):			
Has a doctor or other health professional ever told you that you had hypertension, otherwise known as high blood pressure?	☐ 1=Yes ☐ 2=Yes, but only during pregnancy (skip to # / hce) ☐ 0=No (skip to # / hce) ☐ 77=Don'tknow (skip to # / hce) ☐ 88=Decline to respond (skip to # / hce)	bloodpressure	Bloodpressure_diagnosed
Module 2, part 2: Selected Adult Physical measurements			
Section 1: Anthropometry			
Height measurement:	pe_heightheight		
Height (in centimetres):	(pe_heightcm)		
Weight measurement:	pe_weightweight		
Weight (in kilograms):	(weightkg)		
Section 2: blood pressure readings			
Blood pressure reading	Systolic blood pressure	Diastolic blood pressure	BP_reading
1	pe_bp1sysystolic	pe_bp1diadiastolic	BP_reading1
2	pe_bp2sysystolic	pe_bp2diadiastolic	BP_reading2
3	pe_bp3sysystolic	pe_bp3diadiastolic	BP_reading3
Module 4: Selected adult biochemical measurements			
Section 2: blood glucose			
HgbA1c	Enter HgbA1C (%): Specify in open text field HgbA1C not measured=0	blood_hgba1cB6	HbA1c_measured
Total cholesterol	Enter total cholesterol (mmol/L): Specify in open text field Total cholesterol not measured=0	blood_totcholesterolB9	Total_cholesterol
Triglycerides	Enter triglyceride (mmol/L): Specify in open text field Triglycerides not measured=0	blood_triglycerides	triglycerides
HDL	Enter HDL (mmol/L): Specify in open text field HDL not measured=0	blood_hdl	High_density_lipoprotein
LDL	Enter LDL (mmol/L): Specify in open text field LDL not measured=0	blood_ldl	Low_density_lipoprotein

Appendix B Saudi National Reference for Diabetes Mellitus Guidelines in Primary Health Care

Screening for T2DM: The clinical spectrum of diabetes ranges from low-risk through to higher-risk individuals or to the symptomatic patient who needs immediate treatment.

- Screening for diabetes implies testing for diabetes in individuals without obvious symptoms who are unaware of their condition.
- Screening for diabetes will also detect individuals at increased risk for diabetes (prediabetes) or individuals with less severe states of dysglycemia who may still be at risk for T2DM.
- Screening strategies vary according to the type of diabetes and evidence of effective interventions to prevent progression of prediabetes to diabetes and/or reduce the risk of complications associated with diabetes.
- The growing importance of diabetes screening is undeniable; the same tests would be used for diagnosis of both diabetes and prediabetes.
- To be effective, population-based screening requires wide coverage with the goal of early identification and subsequent intervention to reduce morbidity and mortality. **(P.15)**

Recommendations:

R1: All individuals should be evaluated annually for T2DM risk based on demographic and clinical criteria (Level C).

R2: Screening for diabetes using FPG and/or A1C should be performed every 3 years in individuals >40 years of age. Individuals at high risk using a risk calculator may require more frequent and/or earlier testing.

Individuals with either FPG and/or A1C or 2hPG in a 75 g OGTT should be considered at very high risk using a risk calculator or in people with additional risk factors for diabetes (Level C - Level B).

R3: Testing with 2hPG in a 75 g OGTT should be undertaken in individuals with FPG 5.6-6.9 mmol/L(100-125mg/dl) and/or A1C 5.7%-6.4% in order to identify individuals with diabetes (Level C). (p.16)

Prevention and management of diabetes: Intensive and structured lifestyle modification that results in *loss of approximately 5% of initial body weight* can reduce the risk of progression from impaired glucose tolerance to T2DM by almost 60%.

- Progression from prediabetes to T2DM can also be reduced by pharmacologic therapy with metformin (30% reduction), acarbose (30% reduction).
- Preventing T2DM would result in significant public health benefits, including lower rates of cardiovascular diseases (CVD), renal failure, blindness and premature mortality.
- **Primary approaches to preventing diabetes in a population include the following:**
 1. Programs targeting high-risk individuals in the community (such as those with impaired glucose tolerance [IGT] or obesity)
 2. Programs targeting high-risk subgroups of the population, such as high-risk ethnic groups.
 3. Programs for the general population, such as those designed to promote physical activity and healthy eating in adults or children. (p.17)

Recommendations:

R1: A structured program that emphasises life style changes that include moderate weight loss (7% body weight) and regular physical activity (150 min/week), with dietary strategies including reduced calories and reduced intake of dietary fat; limiting intake of sugar-sweetened beverages can also reduce the risk for developing diabetes (level A).

R2: Metformin therapy for prevention of type 2 diabetes may be considered in those with IGT, IFG, or an A1C of 5.7–6.4%, especially for those with BMI >35 kg/m², aged 60 years, and women with prior GDM (level A) (level C).

R3: At least annual monitoring for the development of diabetes in those with prediabetes is suggested (Level C) (Level B).

R4: Screening for and treatment of modifiable risk factors for CVD is suggested. (p.18)

Monitoring Glycaemic Control

- Glycated haemoglobin (A1C) is a valuable indicator of treatment effectiveness and should be measured every 3 months when glycaemic targets are not being met and when diabetes therapy is being adjusted.
- Awareness of all measures of glycaemia, including Self-Monitoring of Blood Glucose (SMBG) results and A1C, provides the best information to assess glycaemic control.
- The frequency of SMBG should be determined individually, based on the type of diabetes, the treatment prescribed, the need for information about BG levels and the individual's capacity to use the information from testing to modify behaviours or adjust medications. **(p.19)**

Appendix C Sensitivity analysis for weighting methodology

Table C.1 - Scenarios of sensitivity analysis based on all diabetes outcome (Odds ratios and 95% confidence interval)

Education level	Scenarios	Model 1	Model 2
Low education	Unweighted	4.032*** (3.384 - 4.804)	1.795*** (1.368 - 2.355)
Middle education		1.636*** (1.354 - 1.978)	1.361*** (1.075 - 1.724)
High education		1	1
Low education	Lab weight	3.877*** (2.953 - 5.092)	1.564 (1.035 - 2.364)
Middle education		1.469*** (1.105 - 1.953)	1.374 (0.950 - 1.988)
High education		1	1
Low education	New weight	3.944*** (3.007- 5.175)	1.648 (1.093 - 2.483)
Middle education		1.338* (1.001 - 1.787)	1.287 (0.885 - 1.872)
High education		1	1
Low education	Household weight	3.759*** (2.936 - 4.814)	1.815*** (1.247 - 2.641)
Middle education		1.501*** (1.157 - 1.948)	1.448** (1.036 - 2.023)
High education		1	1

Outcome variable “all diabetes” No (0), Yes (1). *p<0.05 **p<0.01 ***p<0.001; **Model 1:** unadjusted; **Model 2:** adjusted by age, sex and marital status, employment status and monthly household income; and Lifestyle factors (smoking, physical activity and BMI)

Appendix D Projection models for T2DM prevalence

- First method: United Nations projection data

Table D.1 - Projection prevalences for the MALE population (2020 – 2025 – 2030)											
Age	1992	2004	2010	2016	2017	2018	Average growth	2015	2020	2025	2030
0 - 4	84.4	82.7	74.1	79.8	79.4	79	0.37	77.9	78.3	76.5	74.6
5 - 9	85.3	84.5	75.2	75.3	74.9	74.4	0.42	73.2	73.6	71.5	69.4
10 - 14	87.6	86	78.3	76.7	76.4	75.9	0.4	74.7	75.1	73.2	71.2
15 - 19	89.2	87.7	82.1	79.2	78.8	78.4	0.36	77.4	77.7	75.9	74.1
20 - 24	67	71.8	70.3	78.8	78.5	78.1	0.37	77	77.4	75.5	73.7
25 - 29	40.4	53.7	53	57.6	57.2	56.6	0.53	55	55.5	52.9	50.3
30 - 34	34	44.5	45.4	48.3	47.9	47.2	0.54	45.6	46.2	43.5	40.8
35 - 39	34.6	42.2	43.8	35.8	35.4	34.9	0.49	33.4	33.9	31.5	29
40 - 44	36	44.6	43.3	33.4	32.9	32.4	0.47	31	31.5	29.1	26.7
45 - 49	47.3	48.1	49.6	36.4	35.9	35.4	0.49	33.9	34.4	31.9	29.5
50 - 54	61	53.5	54.1	40.1	39.7	39.1	0.51	37.5	38.1	35.5	32.9
55 - 59	74.6	63.2	61.6	46.6	46.1	45.5	0.52	44	44.5	41.9	39.3
60 - 64	86	76.2	73.3	54.1	53.7	53.1	0.52	51.5	52	49.4	46.8
65 +	92.5	89.3	86.9	75.7	75.3	74.9	0.36	73.9	74.2	72.4	70.6

Table D.2 – Projection prevalences for the FEMALE population (2020 – 2025 – 2030)											
Age	1992	2004	2010	2016	2017	2018	Average growth	2015	2020	2025	2030
0 - 4	84.5	83	74.8	80	79.6	79.3	0.37	78.2	78.6	76.7	74.9
5 - 9	85.4	84.9	78	75.5	75.1	74.7	0.43	73.4	73.8	71.7	69.5
10 - 14	87.5	87.3	82.2	77.2	76.8	76.4	0.4	75.2	75.6	73.6	71.6
15 - 19	89.3	85.4	85	79.8	79.4	79	0.37	77.9	78.3	76.5	74.6
20 - 24	83.5	83.1	83.5	82.1	81.7	81.4	0.35	80.4	80.7	79	77.2
25 - 29	74.4	74.5	76.6	69.2	68.7	68.3	0.47	66.9	67.3	65	62.6
30 - 34	58.6	68	71.5	67.2	66.6	66.2	0.49	64.7	65.2	62.8	60.3
35 - 39	66.4	72.4	71	58.1	57.5	57	0.53	55.4	56	53.3	50.7
40 - 44	72.1	75.4	74	55.9	55.3	54.8	0.55	53.2	53.7	51	48.2
45 - 49	81.5	79.2	79.6	66.1	65.6	65.1	0.5	63.6	64.1	61.6	59.1
50 - 54	86.5	82.1	81	80.5	80.2	79.8	0.36	78.8	79.1	77.3	75.6
55 - 59	90	87.6	83.9	82.8	82.4	82.1	0.32	81.1	81.5	79.8	78.2
60 - 64	90.1	90	86.5	83.2	82.8	82.6	0.32	81.6	81.9	80.3	78.8
65 +	92.6	92.9	89.5	88.2	87.9	87.7	0.22	87.1	87.3	86.2	85.1

- Projected Saudi population in three timeframes

Table D.3 - Projected population of Saudi males				
Age	2015	2020	2025	2030
25 – 29	844,013	919,498	830,033	794,972
30 – 34	821,217	910,430	859,091	750,340
35 – 39	694,831	707,868	637,335	592,981
40 - 44	568,349	687,728	556,779	518,280
45 – 49	461,712	633,728	623,953	522,294
50 – 54	381,724	490,964	585,026	576,809
55 – 59	289,684	418,673	505,805	586,777
60 – 64	220,511	310,674	434,743	524,053
65 +	375,838	537,091	775,139	1,128,477
total	4,657,879	5,616,654	5,807,904	5,994,983

Table D.4 - Projected population of Saudi females				
Age	2015	2020	2025	2030
25 – 29	858,752	919,301	762,181	721,637
30 – 34	843,798	918,434	889,212	733,625
35 – 39	725,389	724,303	755,035	720,646
40 - 44	562,777	677,299	646,878	669,387
45 – 49	462,190	651,743	754,640	733,610
50 – 54	409,930	535,423	758,714	907,078
55 – 59	268,600	392,532	521,307	750,625
60 – 64	213,018	255,849	372,034	495,594
65 +	395,851	506,060	634,235	858,246
total	4,740,305	5,580,944	6,094,236	6,590,448

Table D.5 – Projected population of Saudi populations - All				
Age	2015	2020	2025	2030
25 – 29	1,702,765	1,838,799	1,592,214	1,516,609
30 – 34	1,665,015	1,828,864	1,748,303	1,483,965
35 – 39	1,420,220	1,432,171	1,392,370	1,313,627
40 - 44	1,131,126	1,365,027	1,203,657	1,187,667
45 – 49	923,902	1,285,471	1,378,593	1,255,904
50 – 54	791,654	1,026,387	1,343,740	1,483,887
55 – 59	558,284	811,205	1,027,112	1,337,402
60 – 64	433,529	566,523	806,777	1,019,647
65 +	771,689	1,043,151	1,409,374	1,986,723
Total (N)	9,398,184	11,197,598	11,902,140	12,585,431

- Second method: Survivorship ratio calculations for Mortality Cohort Component projection

Table D.6 - Saudi male population survival ratios							
Age	2018			2023			2028
15 - 19	926156		SR	number			
20 - 24	1072129	24 - 20	0.993295	919946.6		SR	number
25 – 29	996017	29 - 25	0.994798	1066551	29 - 25	0.994798	915160.7
30 – 34	896790	34 - 30	0.996716	992745.6	34 - 30	0.996716	1063048
35 – 39	786779	39 - 35	0.996084	893278.1	39 - 35	0.996084	988857.9
40 - 44	665841	44 - 40	0.993197	781426.7	44 - 40	0.993197	887201.3
45 – 49	559539	49 - 45	0.989852	659084.1	49 - 45	0.989852	773496.9
50 – 54	446271	54 - 50	0.983808	550478.8	54 - 50	0.983808	648412
55 – 59	348081	59 - 55	0.9734	434400.4	59 - 55	0.9734	535836.2
60 – 64	252157	64 - 60	0.941985	327887.2	64 - 60	0.941985	409198.8
65 +	425553	65 +	0.729181	494173.6	65 +	0.729181	599431.5
	5377028			6200026			6820644
SR = Survivorship Ratio							

Table D.7 - Saudi female population survival ratios

Age	2018			2023			2028
19 - 15	894846		SR	number			
24 - 20	980185	24 - 20	0.997451	892564.6		SR	number
29 - 25	975252	29 - 25	0.997445	977680.2	29 - 25	0.997445	890283.7
34 - 30	881961	34 - 30	0.997511	972824.6	34 - 30	0.997511	975246.8
39 - 35	768617	39 - 35	0.996355	878746	39 - 35	0.996355	969278.4
44 - 40	641244	44 - 40	0.993702	763776.1	44 - 40	0.993702	873211.5
49 - 45	529550	49 - 45	0.990926	635425.6	49 - 45	0.990926	756845.9
54 - 50	421928	54 - 50	0.985788	522023.8	54 - 50	0.985788	626394.6
59 - 55	319715	59 - 55	0.977144	412284.3	59 - 55	0.977144	510092.3
64 - 60	235932	64 - 60	0.956749	305886.9	64 - 60	0.956749	394452.5
65+	445609	65 +	0.757712	516411.6	65 +	0.757712	623065.2
total	5219808			5985059			6618871

SR = Survivorship Ratio

Table D.8 - Saudi total population survival ratios

Age	2018	2023	2028
15 - 19	1821002		
20 - 24	2052314	1812511	
25 - 29	1971269	2044232	1805444
30 - 34	1778751	1965570	2038295
35 - 39	1555396	1772024	1958136
40 - 44	1307085	1545203	1760413
45 - 49	1089089	1294510	1530343
50 - 54	868199	1072503	1274807
55 - 59	667796	846684.7	1045929
60 - 64	488089	633774.2	803651.3
65 +	871162	1010585	1222497
N	10596836	12185085	13439515

Appendix E Internal validation of the undiagnosed diabetes prediction model using the split sample validation method

I internally validated the development model by using a split-sample validation method. I randomly split the data into two parts: 60% (2,228) of the data as the development sample and 40% (1,575) for the validation sample. For each sample of the data, I obtained performance measures to evaluate model development and validation by estimating the AUC. In addition, I evaluated the overall prediction model performance by averaging the AUC estimates between two samples. The performance of the validation sample is similar to the development sample with a minimal decrease in the AUC value. The development sample contains 2,228 individuals of whom 232 (10.4) were undiagnosed of type two diabetes. However, the validation sample contains 1,575 individuals of whom 158 (10.0) were undiagnosed of type two diabetes. The performance of the model is shown in the following table. The AUC of the development model shows a minimal increase with higher sensitivity compared to the validation model. However, both models have similar specificity values 0.62.

Models	observation	AUC	Sensitivity	Specificity	Optimal cut-point	Youden index	LR+	LR-	HL-c	HL-p
Development (60%)	2,228	0.67	0.67	0.62	-2.21	0.285	1.763	0.532	13.02	0.1111
Validation (40%)	1,575	0.64	0.62	0.62	-2.29	0.238	1.632	0.613	14.09	0.0794

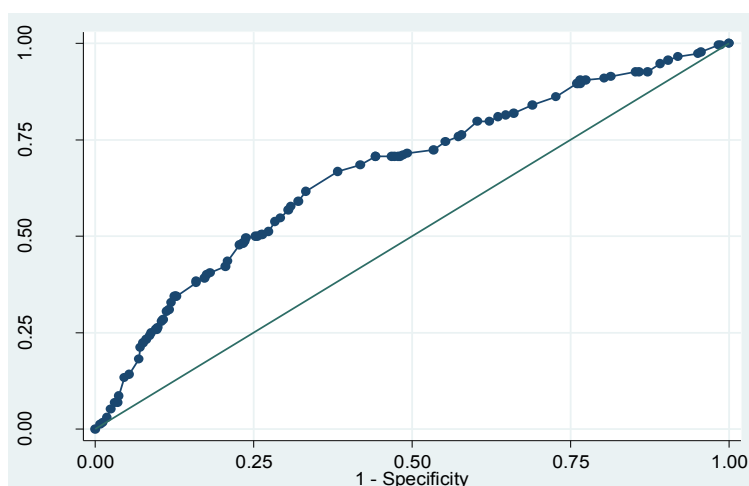


Figure E.1 - Receiver operating characteristics (ROC) curve showing performance of development sample

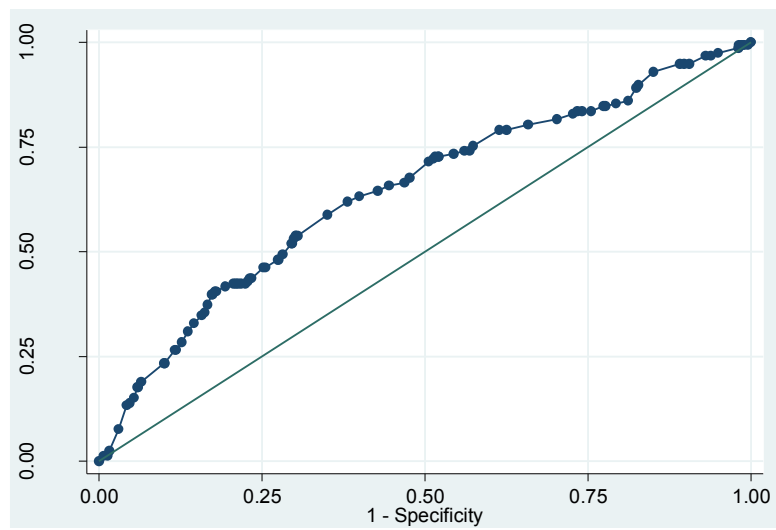


Figure E.2 - Receiver operating characteristics (ROC) curve showing performance of validation sample

Table E.2 - Comparison between prediction models based on different selection procedures and evaluation model measurements

Model	predictors	N	BIC	AIC	AUC	Sens	Spec	Cut-point	Youden	HL-p	HL-c
Household income classification without combine groups (8 groups): 1= Less than 3,000 Riyals; 2=3,000 Riyals to less than 5,000; 3=5,000 Riyals to less than 7,000; 4=7,000 Riyals to less than 10,000; 5= 10,000 Riyals to less than 15,000; 6=15,000 Riyals to less than 20,000; 7=20,000 Riyals to less than 30,000; 8= 30,000 Riyals or more											
Without region	Education +employment+ household income+ bmi + age	3,085	2078.961	1946.206	0.6703	0.64	0.63	-2.2268	0.267	0.6049	6.38
Without Age	Education +employment+ household income +bmi + region	3,085	2082.402	1955.682	0.6577	0.73	0.52	-2.3296	0.244	0.5852	6.56
Without Region and age	Education +employment+ household income +bmi	3,085	2067.644	1965.061	0.6406	0.67	0.55	-2.26307	0.217	0.7036	5.50
With region and age	Education +employment+ household income +bmi + age+ region	3,085	2092.365	1935.473	0.6853	0.67	0.63	-2.23674	0.300	0.1631	11.74
	Age+ region+ household income +bmi	3,085	2049.08	1934.429	0.6764	0.65	0.63	-2.24958	0.278	0.0289	17.12
Household income classification with combine groups (four groups): Low: (1= Less than 3,000 Riyals; 2=3,000 Riyals to less than 5,000); Middle: (3=5,000 Riyals to less than 7,000; 4=7,000 Riyals to less than 10,000) ; high (5= 10,000 Riyals to less than 15,000; 6=15,000 Riyals to less than 20,000) ; highest (7=20,000 Riyals to less than 30,000; 8= 30,000 Riyals or more)											
Without region	Education +employment+ household income+ bmi + age	3,085	2055.734	1947.116	0.6628	0.66	0.60	-2.2466	0.259	0.8847	3.68
Without Age	Education +employment+ household income +bmi + region	3,085	2059.682	1957.099	0.6496	0.71	0.52	-2.313096	0.230	0.7841	4.75
Without Region and age	Education +employment+ household income +bmi	3,085	2044.827	1966.381	0.6290	0.64	0.55	-2.19909	0.191	0.5684	6.71
With region and age	Education +employment+ household income +bmi + age+ region	3,085	2069.311	1936.557	0.6775	0.68	0.59	-2.2615	0.277	0.8401	4.18
	Age+ region+ household income +bmi	3,085	2025.367	1934.853	0.6686	0.55	0.71	-2.07397	0.266	0.8783	3.76
Without individual socio-economic factors											
	Age+BMI+Region	3,085	2008.598	1936.186	0.6653	0.62	0.66	-2.149398	0.277	14.47	0.0702

Appendix F Prediction model for undiagnosed pre-diabetes

Table F.1 - Summary statistics for different prediction models based on model selection and performance measures

Model	predictors	Observations	BIC	AIC	AUC	Sensitivity	Specificity	Cut-point	Youden index	HL-p
1	Age + Sex + BMI+ HTN	3,682	4250.18	4181.85	0.6330	0.66	0.55	-1.059	0.204	0.3047
2	Age + Sex + BMI	3,682	4246.95	4184.83	0.6315	0.60	0.59	-0.983	0.191	0.8679
3	Age + Sex	3,682	4253.72	4210.24	0.6147	0.65	0.53	-1.084	0.181	0.9941
4	Age + HTN	3,682	4252.89	4209.41	0.6170	0.65	0.53	-1.016	0.181	0.5805
5	Age +BMI	3,682	4741.59	4247.97	0.6274	0.65	0.54	-1.149	0.194	0.8554
6	Age + BMI +HTN	3,682	4249.78	4187.67	0.6300	0.63	0.57	-0.994	0.202	0.6759
7	Age + Sex +HTN	3,682	4255.45	4205.76	0.6177	0.68	0.50	-1.186	0.184	0.8723

Table F.2 - Risk factors and final development prediction model for undiagnosed pre-diabetes

Risk factor	Categories	Odds Ratio [95% CI]	β	P-value	Score
	Intercept		-1.359		
Age	15 - 24	1	reference	--	0
	25 - 34	0.98 [0.77, 1.25]	-0.016	0.899	
	35 - 44	1.46 [1.15, 1.86]	0.378	0.002	
	45 - 54	1.77 [1.36, 2.30]	0.569	0.000	
	55 - 64	2.48 [1.79, 3.42]	0.907	0.000	
	65+	2.15 [1.58, 2.91]	0.763	0.000	
Sex	Female	1	reference	--	0
	Male	1.24 [1.07, 1.44]	0.215	0.005	
BMI	Underweight	1	reference	--	0
	Normal	1.79 [1.07, 2.97]	0.579	0.025	
	Overweight	2.47 [1.48, 4.10]	0.903	0.000	
	Obese	2.74 [1.65, 4.55]	1.008	0.000	
Hypertension	No	1	reference	--	0
	Yes	1.23 [1.03, 1.47]	0.206	0.000	

Table F.3 - Confusion matrix for undiagnosed pre-T2DM

		True Disease Status (Gold standard = HbA1c)		
		Abnormal	Normal	Total
Test results (Predicted test)	Positive (+ve)	TP = 663	FP = 1,214	1,805
	Negative (-ve)	FN = 345	TN = 1,460	1,877
	Total	1,008	2,674	3,682
Abnormal: Undiagnosed (HbA1c $\geq$ 5.7%); Normal: (HbA1c < 5.7%); Positive: cut-point $\geq$ -1.059; Negative: cut-point < -1.059				

Table F.4 - Summary statistics for different prediction models based on model selection and performance measures

Models	observation	AUC	Sensitivity	Specificity	Optimal cut-point	Youden index	HL-c	HL-p
Original sample	3,682	0.6330	0.66	0.55	-1.059	0.204	0.3047	9.46
Bootstrapping sample	3,682	0.6330	0.66	0.55	-1.059	0.204	0.3047	9.46

Table F.5 - Observed and predicted values of undiagnosed prediabetes status using test score as a predictive variable in groups defined by Hosmer-Lemeshow Goodness-of-fit test

Group/ quantile	Likelihood/ probability of being undiagnosed pre- diabetes	Present (undiagnosed)		Absent (normal)		Total
		Observed (1)	Expected (1)	Observed (0)	Expected (0)	
1	0.1557	72	65.2	399	405.8	471
2	0.1862	46	56.0	257	247.0	303
3	0.2167	61	68.2	273	265.8	334
4	0.2373	95	86.9	292	300.1	387
5	0.2598	92	95.4	288	284.6	380
6	0.2922	136	124.6	303	314.4	439
7	0.3292	94	89.5	192	196.5	286
8	0.3583	126	136.4	273	262.6	399
9	0.4069	118	125.4	209	201.6	327
10	0.5163	168	160.4	188	195.6	356
Results	H-L Statistic = 9.46	Degree of freedom = 8		Significance (P-value) = 0.3047		

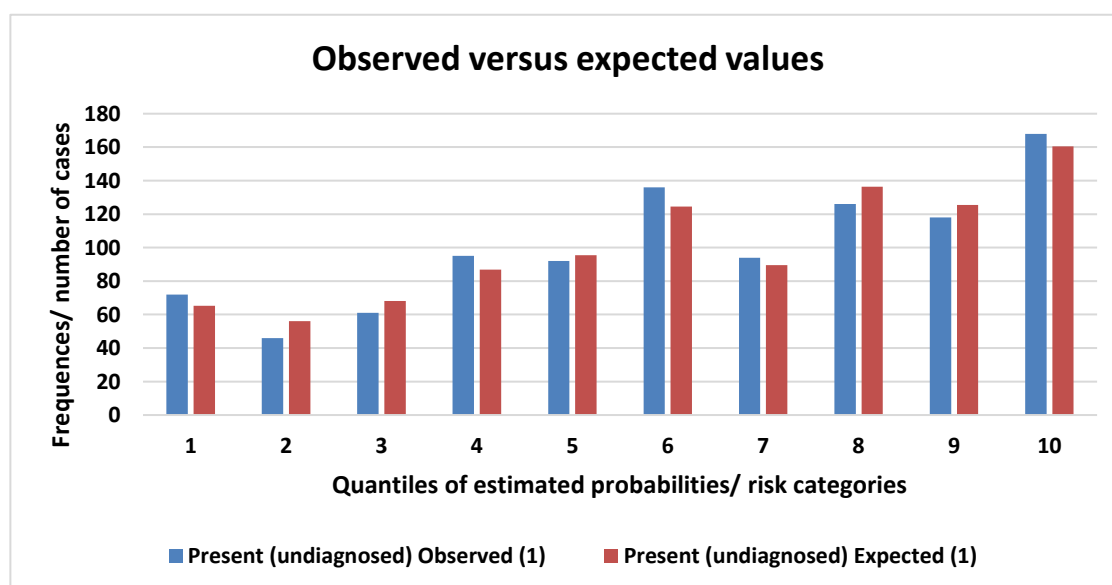


Figure F.1 - Calibration of the test score: observed and expected of undiagnosed prediabetes

Appendix G Characteristics of observed and non- observed diabetes patients

Table G.1 - Characteristics of diabetes patients who observed and missed

Variables	Categories N (%)	Total 538	Observed 462 (85.9)	Missed 76 (14.1)	p-value
Demographic factors					
Sex	Female	225	186 (40.3)	39 (51.3)	0.070
	Male	313	276 (59.7)	37 (48.7)	
Age	20 - 39	41	38 (8.0)	3 (4.2)	0.114
	40 - 59	239	212 (46.0)	27 (37.5)	
	60 +	254	212 (46.0)	42 (58.3)	
Marital status	Currently married	414	355 (76.8)	59 (78.7)	0.931
	Never married	14	12 (2.6)	2(2.6)	
	Divorced, Separated, and widowed	109	95 (20.6)	14 (18.6)	
Socio-economic factors					
Education level	Low education	288	236 (51.1)	52 (68.4)	0.019
	Middle education	131	118 (25.5)	13 (17.1)	
	High education	119	108 (23.4)	11 (14.5)	
Working status	Never worked	135	112 (24.2)	23 (30.7)	0.030
	Currently working	120	112 (24.2)	8 (10.7)	
	Not currently working	282	238 (52.0)	44 (58.6)	
Monthly Household income	Lowest	249	214 (46.3)	35 (46.1)	0.026
	Low	74	65 (14.1)	9 (11.8)	
	Middle	65	57 (12.3)	8 (10.5)	
	High	34	33 (7.1)	1 (1.3)	
	Highest	23	22 (4.8)	1 (1.3)	
	Missing	93	71 (15.4)	22 (29.0)	
Region location					
Administrative Region	Central	163	140 (30.3)	23 (30.3)	0.000
	Eastern	27	24 (5.2)	3 (4.0)	
	Western	88	77 (16.7)	11 (14.5)	
	Northern	93	89 (19.3)	4 (5.3)	
	Southern	167	132 (28.6)	35 (46.1)	

Appendix G

Variables	Categories N (%)	Total 538	Observed 462 (85.9)	Missed 76 (14.1)	p-value
Health care access factors					
Distance between home and health care facility	Less than 60 minutes	397	350 (93.6)	47 (70.2)	0.000
	More than 60 minutes	44	24 (6.4)	20 (29.8)	
Health care provider number (other)	No	143	133 (29.8)	10 (16.1)	0.011
	Yes, there is one place	141	127 (28.5)	14 (22.6)	
	Yes, there are more than two	224	186 (41.7)	38 (61.3)	
Clinic type	Governmental (MOH)	369	314 (81.8)	55 (83.3)	0.637
	Other Governmental clinic	40	36 (9.4)	4 (6.1)	
	Private clinic	41	34 (8.8)	7 (10.6)	
Hospital type	Governmental (MOH)	381	330 (79.5)	51 (76.1)	0.185
	Other Governmental hospital	44	40 (9.6)	4 (5.9)	
	Private hospital	57	45 (10.8)	12 (17.9)	
Self-care behavioural factors					
Taking medication	Yes	454	396 (85.7)	58 (77.3)	0.000
	No	78	66 (14.3)	12 (16.0)	
	Don't know	5		5 (6.7)	
Insulin	Yes	320	278 (60.2)	42 (56.0)	0.000
	No	210	184 (39.8)	26 (34.7)	
	Don't know	7		7 (9.3)	
Herbal medication	Yes	36	31 (6.7)	5 (6.7)	0.002
	No	499	431 (93.3)	68 (90.7)	
	Don't know	2		2 (2.7)	
Visited traditional healer	yes	71	58 (12.6)	13 (17.3)	0.257
	No	466	404 (87.4)	62 (82.7)	
HbA1c status	(uncontrolled)	301	259 (56.1)	42 (55.3)	0.897
	(controlled)	237	203 (43.9)	34 (56.1)	

List of References

- ABDELLA, N., AL AROUJ, M., AL NAKHI, A., AL ASSOUSSE, A. & MOUSSA, M. 1998. Non-insulin-dependent diabetes in Kuwait: prevalence rates and associated risk factors. *Diabetes research and clinical practice*, 42, 187-196.
- ABDELMONEIM, I. & AL HOMRANY, M. 2002. Health education in the management of diabetes at the primary health care level: is there a gender difference?
- ABDELWAHID, H. A., AL-SHAHRANI, S. I., ELSABA, M. S. & ELMORSHEDI, W. S. 2010. Patterns of referral in the Family Medicine Department in Southeastern Saudi Arabia. *Saudi medical journal*, 31, 925-930.
- ABDUL RAHIM, H. F., HUSSEINI, A., GIACAMAN, R., JERVELL, J. & BJERTNESS, E. 2001. Diabetes mellitus in an urban Palestinian population: prevalence and associated factors. *EMHJ-Eastern Mediterranean Health Journal*, 7 (1-2), 67-78, 2001.
- ABDULAZIZ AL DAWISH, M., ALWIN ROBERT, A., BRAHAM, R., ABDALLAH AL HAYEK, A., AL SAEED, A., AHMED AHMED, R. & SULAIMAN AL SABAAN, F. 2016. Diabetes mellitus in Saudi Arabia: a review of the recent literature. *Current diabetes reviews*, 12, 359-368.
- ABUYASSIN, B. & LAHER, I. 2016. Diabetes epidemic sweeping the Arab world. *World journal of diabetes*, 7, 165.
- AGARDH, E., AHLBOM, A., ANDERSSON, T., EFENDIC, S., GRILL, V., HALLQVIST, J. & ÖSTENSON, C. 2007. Socio-economic position at three points in life in association with type 2 diabetes and impaired glucose tolerance in middle-aged Swedish men and women. *International journal of epidemiology*, 36, 84-92.
- AGARDH, E., ALLEBECK, P., HALLQVIST, J., MORADI, T. & SIDORCHUK, A. 2011. Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. *Int J Epidemiol*, 40, 804-18.
- AGARDH, E. E., AHLBOM, A., ANDERSSON, T., EFENDIC, S., GRILL, V., HALLQVIST, J. & ÖSTENSON, C.-G. 2004. Explanations of socioeconomic differences in excess risk of type 2 diabetes in Swedish men and women. *Diabetes care*, 27, 716-721.
- AGHILI, R., POLONSKY, W. H., VALOJERDI, A. E., MALEK, M., KESHTKAR, A. A., ESTEGHAMATI, A., HEYMAN, M. & KHAMSEH, M. E. 2016. Type 2 diabetes: model of factors associated with glycemic control. *Canadian journal of diabetes*, 40, 424-430.
- AGRESTI, A. & KATERI, M. 2011. Categorical data analysis. *International encyclopedia of statistical science*. Springer.
- AJABNOOR, M. & LAAJAM, M. A. 1987. Levels of glucose and glycosylated haemoglobin in Saudi diabetic patients. *Tropical and geographical medicine*, 39, 276-280.
- AJLOUNI, K., KHADER, Y. S., BATIEHA, A., AJLOUNI, H. & EL-KHATEEB, M. 2008. An increase in prevalence of diabetes mellitus in Jordan over 10 years. *Journal of Diabetes and its Complications*, 22, 317-324.
- AL-ABOUDI, I. S., HASSALI, M. A., SHAFIE, A. A., HASSAN, A. & ALRASHEEDY, A. A. 2015. A cross-sectional assessment of health-related quality of life among type 2 diabetes patients in Riyadh, Saudi Arabia. *SAGE open medicine*, 3, 2050312115610129.

List of References

- AL-ALFI, M. A., AL-SAIGUL, A. M., ABED-ELBAST, A. M., SOUROUR, A. M. & RAMZY, H. A. 2007. Quality of primary care referral letters and feedback reports in Buraidah, Qassim region, Saudi Arabia. *Journal of family & community medicine*, 14, 113.
- AL-ALFI, M. A., AL-SAIGUL, A. M., SALEH, M. A., SUROUR, A. M. & RIYADH, M. A. 2004. Audit of structure, process, and outcome of diabetic care at Al Asyah Primary Health Care Centre, Qassim region, Saudi Arabia. *Journal of family & community medicine*, 11, 89.
- AL-BAGHLI, N., AL-GHAMDI, A., AL-TURKI, K., AL ELQ, A., EL-ZUBAIER, A. & BAHNASSY, A. 2010. Prevalence of diabetes mellitus and impaired fasting glucose levels in the Eastern Province of Saudi Arabia: results of a screening campaign. *Singapore medical journal*, 51, 923.
- AL-DAGHRI, N. M., AL-ATTAS, O. S., ALOKAIL, M. S., ALKHARFY, K. M., YOUSEF, M., SABICO, S. L. & CHROUSOS, G. P. 2011. Diabetes mellitus type 2 and other chronic non-communicable diseases in the central region, Saudi Arabia (Riyadh cohort 2): a decade of an epidemic. *BMC medicine*, 9, 1.
- AL-ELQ, A. H. 2009. Current practice in the management of patients with type 2 diabetes mellitus in Saudi Arabia. *Saudi medical journal*, 30, 1551-1556.
- AL-GHAMDI, A. J., AL-TURKI, K. A., AL-BAGHLI, N. A. & EL-ZUBAIER, A. G. 2007. A community-based screening campaign for the detection of diabetes mellitus and hypertension in the Eastern province, Saudi Arabia: methods and participation rates. *Journal of family & community medicine*, 14, 91.
- AL-HABORI, M., AL-MAMARI, M. & AL-MEERI, A. 2004. Type II Diabetes Mellitus and impaired glucose tolerance in Yemen: prevalence, associated metabolic changes and risk factors. *Diabetes research and clinical practice*, 65, 275-281.
- AL-HAMDAN, N., KUTBI, A., CHOUDHRY, A., NOOH, R., SHOUKRI, M. & MUJIB, S. 2005. WHO stepwise approach to NCD surveillance country-specific standard report Saudi Arabia. *World Health Organization*.
- AL-HAYEK, A. A., ROBERT, A. A., ALZAID, A. A., NUSAIR, H. M., ZBAIDI, N. S., AL-EITHAN, M. H. & SAM, A. E. 2012. Association between diabetes self-care, medication adherence, anxiety, depression, and glycemic control in type 2 diabetes. *Saudi medical journal*, 33, 681-683.
- AL-HAZZAA, H. M., ABAHUSSAIN, N. A., AL-SOBAYEL, H. I., QAHWAJI, D. M. & MUSAIGER, A. O. 2011. Physical activity, sedentary behaviors and dietary habits among Saudi adolescents relative to age, gender and region. *International Journal of Behavioral Nutrition and Physical Activity*, 8, 1.
- AL-HERBISH, A. S., EL-MOUZAN, M. I., AL-SALLOUM, A. A., AL-QURACHI, M. M. & AL-OMAR, A. A. 2008. Prevalence of type 1 diabetes mellitus in Saudi Arabian children and adolescents. *Saudi Med J*, 29, 1285-8.
- AL-KAABI, J., AL-MASKARI, F., AFANDI, B., PARKAR, H. & NAGELKERKE, N. 2009. Physical activity and reported barriers to activity among type 2 diabetic patients in the United Arab Emirates. *The review of diabetic studies: RDS*, 6, 271.
- AL-KAABI, J., AL-MASKARI, F., SAADI, H., AFANDI, B., PARKAR, H. & NAGELKERKE, N. 2008. Assessment of dietary practice among diabetic patients in the United Arab Emirates. *The review of diabetic studies: RDS*, 5, 110.

- AL-KHAWALDEH, O. A., AL-HASSAN, M. A. & FROELICHER, E. S. 2012. Self-efficacy, self-management, and glycemic control in adults with type 2 diabetes mellitus. *Journal of Diabetes and its Complications*, 26, 10-16.
- AL-LAWATI, J. & TUOMILEHTO, J. 2007. Diabetes risk score in Oman: a tool to identify prevalent type 2 diabetes among Arabs of the Middle East. *Diabetes research and clinical practice*, 77, 438-444.
- AL-LAWATI, J. A., BARAKAT, M. N., AL-MASKARI, M., ELSAYED, M. K., AL-LAWATI, A. M. & MOHAMMED, A. J. 2012. HbA1c levels among primary healthcare patients with type 2 diabetes mellitus in Oman. *Oman Medical Journal*, 27, 465.
- AL-LAWATI, J. A. & JOUSILAHTI, P. J. 2004. Prevalence and 10-year secular trend of obesity in Oman. *Saudi medical journal*, 25, 346-351.
- AL-LAWATI, J. A., PANDURANGA, P., AL-SHAikh, H. A., MORSI, M., MOHSIN, N., KHANDEKAR, R. B., AL-LAWATI, H. J. & BAYOUMI, R. A. 2015. Epidemiology of Diabetes Mellitus in Oman: Results from two decades of research. *Sultan Qaboos University Medical Journal*, 15, e226.
- AL-MAHROOS, F. & MCKEIGUE, P. M. 1998. High prevalence of diabetes in Bahrainis: associations with ethnicity and raised plasma cholesterol. *Diabetes care*, 21, 936-942.
- AL-MOOSA, S., ALLIN, S., JEMIAI, N., AL-LAWATI, J. & MOSSIALOS, E. 2006. Diabetes and urbanization in the Omani population: an analysis of national survey data. *Population health metrics*, 4, 1.
- AL-NOZHA, M. M., AL-HAZZAA, H. M., ARAFAH, M. R., AL-KHADRA, A., AL-MAZROU, Y. Y., AL-MAATOUQ, M. A., KHAN, N. B., AL-MARZOUKI, K., AL-HARTHI, S. S. & ABDULLAH, M. 2007. Prevalence of physical activity and inactivity among Saudis aged 30-70 years: a population-based cross-sectional study. *Saudi medical journal*, 28, 559-568.
- AL-NOZHA, M. M., AL-MAATOUQ, M. A., AL-MAZROU, Y. Y. & AL-HARTHI, S. S. 2004. Diabetes mellitus in Saudi Arabia.
- AL-NOZHA, M. M., AL-MAZROU, Y. Y., AL-MAATOUQ, M. A., ARAFAH, M. R., KHALIL, M. Z., KHAN, N. B., AL-MARZOUKI, K., ABDULLAH, M. A., AL-KHADRA, A. H. & AL-HARTHI, S. S. 2005. Obesity in Saudi Arabia. *Saudi medical journal*, 26, 824-829.
- AL-NUAIM, A. A., AL-NAKEEB, Y., LYONS, M., AL-HAZZAA, H. M., NEVILL, A., COLLINS, P. & DUNCAN, M. J. 2012. The prevalence of physical activity and sedentary behaviours relative to obesity among adolescents from Al-Ahsa, Saudi Arabia: rural versus urban variations. *Journal of nutrition and metabolism*, 2012.
- AL-NUAIM, A. R., MIRDAD, S., AL-RUBEAN, K., AL-MAZROU, Y., AL-ATTAS, O. & AL-DAGHARI, N. 1998. Pattern and factors associated with glycemic control of Saudi diabetic patients. *Ann Saudi Med*, 18, 109-112.
- AL-QAZAZ, H. K., HASSALI, M. A., SHAFIE, A. A., SULAIMAN, S. A. S. & SUNDRAM, S. 2011. Perception and knowledge of patients with type 2 diabetes in Malaysia about their disease and medication: a qualitative study. *Research in Social and Administrative Pharmacy*, 7, 180-191.
- AL-QUWAIDHI, A., PEARCE, M., CRITCHLEY, J., SOBNGWI, E. & O'FLAHERTY, M. 2014a. Trends and future projections of the prevalence of adult obesity in Saudi Arabia, 1992-2022. *Eastern Mediterranean Health Journal*, 20.

List of References

- AL-QUWAIDHI, A. J., PEARCE, M. S., SOBNGWI, E., CRITCHLEY, J. A. & O'FLAHERTY, M. 2014b. Comparison of type 2 diabetes prevalence estimates in Saudi Arabia from a validated Markov model against the International Diabetes Federation and other modelling studies. *Diabetes Research and Clinical Practice*, 103, 496-503.
- AL-RASHEEDI, A. A. 2015. Glycemic Control among Patients with Type 2 Diabetes Mellitus in Countries of Arabic Gulf. *International journal of health sciences*, 9, 345.
- AL-RUBEAAAN, K. 2015. National surveillance for type 1, type 2 diabetes and prediabetes among children and adolescents: a population-based study (SAUDI-DM). *J Epidemiol Community Health*, 69, 1045-1051.
- AL-RUBEAAAN, K., AL-MANAA, H., KHOJA, T., AHMAD, N., AL-SHARQAWI, A., SIDDIQUI, K., ALNAQEB, D., ABURISHEH, K., YOUSSEF, A. & AL-BATIL, A. 2014a. The Saudi Abnormal Glucose Metabolism and Diabetes Impact Study (SAUDI-DM). *Annals of Saudi medicine*, 34, 465-475.
- AL-RUBEAAAN, K., AL-MANAA, H., KHOJA, T., AHMAD, N., AL-SHARQAWI, A., SIDDIQUI, K., ALNAQEB, D., ABURISHEH, K., YOUSSEF, A. & AL GHAMDIA, A. 2014b. The Saudi Abnormal Glucose Metabolism and Diabetes Impact Study (SAUDI-DM). *Annals of Saudi medicine*, 34, 465.
- AL-RUBEAAAN, K. A., AL-MANAA, H. A., KHOJA, T. A., AL-SHARQAWI, A. H., ABURISHEH, K. H., YOUSSEF, A. M., ALOTAIBI, M. S. & AL-GAMDI, A. A. 2015. Health care services provided to type 1 and type 2 diabetic patients in Saudi Arabia. *Saudi medical journal*, 36, 1216.
- AL-RUBEAAAN, K. A., YOUSSEF, A. M., SUBHANI, S. N., AHMAD, N. A., AL-SHARQAWI, A. H. & IBRAHIM, H. M. 2013. A Web-based interactive diabetes registry for health care management and planning in Saudi Arabia. *Journal of medical Internet research*, 15.
- AL-SHARIF, A. I. & AL-KHALDI, Y. M. 2003. Resource availability for care of hypertensives at primary health settings in Southwestern Saudi Arabia. *Saudi medical journal*, 24, 466-471.
- AL-LAWATI, J., AL RIYAMI, A., MOHAMMED, A. & JOUSILAHTI, P. 2002. Increasing prevalence of diabetes mellitus in Oman. *Diabetic medicine*, 19, 954-957.
- AL-NUAIM, A. 1997. Prevalence of glucose intolerance in urban and rural communities in Saudi Arabia. *Diabetic medicine*, 14, 595-602.
- AL-RUBEAAAN, K., AL-MANAA, H. A., KHOJA, T. A., AHMAD, N. A., AL-SHARQAWI, A. H., SIDDIQUI, K., ALNAQEB, D., ABURISHEH, K. H., YOUSSEF, A. M. & AL-BATEL, A. 2015. Epidemiology of abnormal glucose metabolism in a country facing its epidemic: SAUDI-DM study. *Journal of diabetes*, 7, 622-632.
- AL ALI, R., MZAYEK, F., RASTAM, S., FOUAD, F. M., O'FLAHERTY, M., CAPEWELL, S. & MAZIAK, W. 2013. Forecasting future prevalence of type 2 diabetes mellitus in Syria. *BMC Public Health*, 13, 507.
- AL BALUSHI, K. A., AL-HADDABI, M., AL-ZAKWANI, I. & AL ZA'ABI, M. 2014. Glycemic control among patients with type 2 diabetes at a primary health care center in Oman. *Primary care diabetes*, 8, 239-243.
- AL BASSAM, A. M. A. 2012. *Urbanisation and Migration in Saudi Arabia: the case of Buraydah City*. University of Leicester.

- AL BUSAIDI, N., SHANMUGAM, P. & MANOHARAN, D. 2019. Diabetes in the Middle East: Government Health Care Policies and Strategies that Address the Growing Diabetes Prevalence in the Middle East. *Current diabetes reports*, 19, 8.
- AL JOHANI, K., KENDALL, G. & SNIDER, P. 2015. Self-management practices among type 2 diabetes patients attending primary health-care centres in Medina, Saudi Arabia/Pratiques d'auto-prise en charge chez des patients atteints de diabète de type 2 fréquentant des centres de soins de santé primaires à Médine (Arabie saoudite). *Eastern Mediterranean Health Journal*, 21, 621.
- AL KHALAF, M., EID, M., NAJJAR, H., ALHAJRY, K., DOI, S. & THALIB, L. 2010a. Screening for diabetes in Kuwait and evaluation of risk scores/Depistage du diabete au Koweït et evaluation des scores de risque. *Eastern Mediterranean health journal*, 16, 725-732.
- AL KHALAF, M., EID, M., NAJJAR, H., ALHAJRY, K. & THALIB, L. 2010b. Screening for diabetes in Kuwait and evaluation of risk scores.
- AL OTHAIMEEN, A., AL NOZHA, M. & OSMAN, A. 2007. Obesity: an emerging problem in Saudi Arabia. Analysis of data from the National Nutrition Survey.
- AL QUWAIDHI, A. J. A. 2013. *Epidemiological modelling of type 2 diabetes in Saudi Arabia: predicted trends and public health implications*. Newcastle University.
- AL SAEEDI, M., EL ZUBIER, A., BAHNASSI, A. & AL DAWOOD, K. 2003. Patterns of belief and use of traditional remedies by diabetic patients in Mecca, Saudi Arabia.
- AL SHAHRANI, A. & BARAJA, M. 2014. Patient satisfaction and it's relation to diabetic control in a primary care setting. *Journal of family medicine and primary care*, 3, 5.
- AL WADAANI, H. A. A. & BALAHA, M. H. 2012. Evaluation of medical consultation letters at King Fahd Hospital, Al Hufuf, Saudi Arabia. *Pan African Medical Journal*, 12.
- AL ZURBA, F. I. & AL GARF, A. 1996. Prevalence of diabetes mellitus among Bahrainis attending primary health care centres.
- ALAROUJ, M., BENNAKHI, A., ALNESEF, Y., SHARIFI, M. & ELKUM, N. 2013. Diabetes and associated cardiovascular risk factors in the State of Kuwait: the first national survey. *International Journal of Clinical Practice*, 67, 89-96.
- ALBACHE, N., AL ALI, R., RASTAM, S., FOUAD, F. M., MZAYEK, F. & MAZIAK, W. 2010. Epidemiology of Type 2 diabetes mellitus in Aleppo, Syria. *Journal of diabetes*, 2, 85-91.
- ALBATEL, A. H. 2005. Population growth and economic development in Saudi Arabia. *Scientific Journal of King Faisal University (Humanities and Management Sciences)*, 6, 341-374.
- ALBEJAIDI, F. M. 2010. Healthcare system in Saudi Arabia: An analysis of structure, total quality management and future challenges. *Journal of Alternative Perspectives in the Social Sciences*, 2, 794-818.
- ALBERTI, K. G. M. M. & ZIMMET, P. F. 1998. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diabetic medicine*, 15, 539-553.
- ALDOSSARY, A., WHILE, A. & BARRIBALL, L. 2008. Health care and nursing in Saudi Arabia. *International nursing review*, 55, 125-128.
- ALHAMDAN, A. A., ALSHAMMARI, S. A., AL-AMOUD, M. M., HAMEED, T. A., AL-MUAMMAR, M. N., BINDAWAS, S. M., AL-ORF, S. M., MOHAMED, A. G., AL-GHAMDI, E. A. & CALDER, P. C.

List of References

2015. Evaluation of health care services provided for older adults in primary health care centers and its internal environment. A step towards age-friendly health centers. *Saudi Med J*, 36, 1091-6.
- ALHARBI, N. S., ALMUTARI, R., JONES, S., AL-DAGHRI, N., KHUNTI, K. & DE LUSIGNAN, S. 2014. Trends in the prevalence of type 2 diabetes mellitus and obesity in the Arabian Gulf States: Systematic review and meta-analysis. *Diabetes Research and Clinical Practice*, 106, e30-e33.
- ALHOWAISH, A. K. 2013. Economic costs of diabetes in Saudi Arabia. *Journal of family & community medicine*, 20, 1.
- ALHYAS, L., MCKAY, A., BALASANTHIRAN, A. & MAJEED, A. 2011. Quality of type 2 diabetes management in the states of the Co-operation Council for the Arab States of the Gulf: a systematic review. *PloS one*, 6, e22186.
- ALHYAS, L., MCKAY, A. & MAJEED, A. 2012. Prevalence of type 2 diabetes in the States of the co-operation council for the Arab States of the Gulf: a systematic review. *PLoS One*, 7, e40948.
- ALJEFREE, N. & AHMED, F. 2015. Prevalence of cardiovascular disease and associated risk factors among adult population in the Gulf region: a systematic review. *Advances in Public Health*, 2015.
- ALJOUDI, A. S. & TAHA, A. Z. 2009. Knowledge of diabetes risk factors and preventive measures among attendees of a primary care center in eastern Saudi Arabia. *Annals of Saudi medicine*, 29, 15.
- ALMAJWAL, A. M., AL-BAGHLI, N. A., BATTERHAM, M. J., WILLIAMS, P. G., AL-TURKI, K. A. & AL-GHAMDI, A. J. 2009. Performance of body mass index in predicting diabetes and hypertension in the Eastern Province of Saudi Arabia. *Annals of Saudi medicine*, 29, 437.
- ALMALKI, M., FITZGERALD, G. & CLARK, M. 2011. Health care system in Saudi Arabia: an overview/Aperçu du système de santé en Arabie saoudite. *Eastern Mediterranean health journal*, 17, 784.
- ALOMAR, R. S., PARSLOW, R. C. & LAW, G. R. 2018. Development of two socioeconomic indices for Saudi Arabia. *BMC public health*, 18, 791.
- ALQAHTANI, N., SCOTT, J., ALSHAHRANI, S. & ULLAH, S. 2014. Socio-economic determinants factors of overweight and obesity among Saudi children. *J Obes Weight Loss Ther*, 4, 2.
- ALQUAIZ, A. M. & TAYEL, S. A. 2009. Barriers to a healthy lifestyle among patients attending primary care clinics at a university hospital in Riyadh. *Annals of Saudi medicine*, 29, 30.
- ALQURASHI, K., ALJABRI, K. & BOKHARI, S. 2011. Prevalence of diabetes mellitus in a Saudi community. *Annals of Saudi medicine*, 31, 19.
- ALRAMADAN, M. J., MAGLIANO, D. J., ALMIGBAL, T. H., BATAIS, M. A., AFROZ, A., ALRAMADHAN, H. J., MAHFOUD, W. F., ALRAGAS, A. M. & BILLAH, B. 2018. Glycaemic control for people with type 2 diabetes in Saudi Arabia—an urgent need for a review of management plan. *BMC endocrine disorders*, 18, 62.
- ALSHAIKH, M. K., FILIPPIDIS, F. T., AL-OMAR, H. A., RAWAF, S., MAJEED, A. & SALMASI, A.-M. 2017. The ticking time bomb in lifestyle-related diseases among women in the Gulf Cooperation Council countries; review of systematic reviews. *BMC public health*, 17, 536.

- ALSHARIF, A. Health system and insurance in Saudi Arabia. Board of Healthcare Funders Southern African Annual Conference, Durban, 2008. 3-16.
- ALSULAIMAN, T. A., AL-AJMI, H. A., AL-QAHTANI, S. M., FADLALLAH, I. M., NAWAR, N. E., SHUKERALLAH, R. E., NADEEM, S. R., AL-WEHEEDY, N. M., AL-SULAIMAN, K. A. & HASSAN, A. A. 2016. Control of type 2 diabetes in King Abdulaziz Housing City (Iskan) population, Saudi Arabia. *Journal of family & community medicine*, 23, 1.
- ALTMAN, D. G., VERGOUWE, Y., ROYSTON, P. & MOONS, K. G. 2009. Prognosis and prognostic research: validating a prognostic model. *Bmj*, 338, b605.
- ALZAHAB, R. A. & ALTEMANI, A. H. 2018. The prevalence and determinants of poor glycemic control among adults with type 2 diabetes mellitus in Saudi Arabia. *Diabetes, metabolic syndrome and obesity: targets and therapy*, 11, 15.
- AMARASEKARA, A. A. T. D., FONGKAEW, W., WIMALASEKERA, S. W., TURALE, S. & CHANPRASIT, C. 2015. Cross-sectional study of glycemic control among adults with type 2 diabetes. *Nursing & health sciences*, 17, 223-228.
- AMERICAN DIABETES, A. 2014. Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 37 Suppl 1, S81-90.
- ANARI, R., AMANI, R. & VEISSI, M. 2016. Obesity and poor glycemic control in patients with type 2 diabetes. *International Journal of Research in Medical Sciences*, 4, 584-588.
- ANDERSEN, A., CARSON, C., WATT, H., LAWLOR, D., AVLUND, K. & EBRAHIM, S. 2008. Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study. *Diabetic medicine*, 25, 1462-1468.
- ANDERSSON, T., AHLBOM, A. & CARLSSON, S. 2015. Diabetes prevalence in Sweden at present and projections for year 2050. *PloS one*, 10, e0143084.
- ANOKUTE, C. C. 1990. Epidemiologic Studies of Diabetes Mellitus in Saudi Arabia-Part I—Screening of 3158 Males in King Saud University. *The Journal of the Royal Society for the Promotion of Health*, 110, 201-203.
- ASHWAN, M. S. A., SALAM, A. A. & MOUSELHY, M. A. 2012. Population growth, structure and distribution in Saudi Arabia. *International Journal of Humanities and Social Sciences*, 1.
- ASSOCIATION, A. D. 2016. Standards of medical care in diabetes—2016: summary of revisions. *Diabetes care*, 39, S4-S5.
- AWAD, S. F., O'FLAHERTY, M., CRITCHLEY, J. & ABU-RADDAD, L. J. 2018. Forecasting the burden of type 2 diabetes mellitus in Qatar to 2050: A novel modeling approach. *diabetes research and clinical practice*, 137, 100-108.
- AWADALLA, N., ABOELYAZED, A., HASSANEIN, M., KHALIL, S., AFTAB, R., GABALLA, I. & MAHFOUZ, A. 2014. Assessment of physical inactivity and perceived barriers to physical activity among health college students, south-western Saudi Arabia. *Eastern Mediterranean Health Journal*, 20.
- BAAN, C. A., RUIGE, J. B., STOLK, R. P., WITTEMAN, J. C., DEKKER, J. M., HEINE, R. J. & FESKENS, E. J. 1999. Performance of a predictive model to identify undiagnosed diabetes in a health care setting. *Diabetes Care*, 22, 213-219.
- BACCHUS, R., BELL, J., MADKOUR, M. & KILSHAW, B. 1982. The prevalence of diabetes mellitus in male Saudi Arabs. *Diabetologia*, 23, 330-332.

List of References

- BADAWI, D., SALEH, S., NATAFGI, N., MOURAD, Y. & BEHBEHANI, K. 2015. Quality of type II diabetes care in primary health care centers in Kuwait: employment of a diabetes quality indicator set (DQIS). *PloS one*, 10, e0132883.
- BADRAN, M. & LAHER, I. 2012. Type II Diabetes Mellitus in Arabic-Speaking Countries. *Int J Endocrinol*, 2012, 902873.
- BAE, J., LAGE, M., MO, D., NELSON, D. & HOOGWERF, B. 2016. Obesity and glycemic control in patients with diabetes mellitus: analysis of physician electronic health records in the US from 2009–2011. *Journal of Diabetes and its Complications*, 30, 212–220.
- BAGUST, A., HOPKINSON, P., MASLOVE, L. & CURRIE, C. J. 2002. The projected health care burden of Type 2 diabetes in the UK from 2000 to 2060. *Diabetic medicine*, 19, 1–5.
- BAKHOTMAH, B. A. & ALZHRANI, H. A. 2010. Self-reported use of complementary and alternative medicine (CAM) products in topical treatment of diabetic foot disorders by diabetic patients in Jeddah, Western Saudi Arabia. *BMC research notes*, 3, 254.
- BAMBRA, C., GIBSON, M., AMANDA, S., WRIGHT, K., WHITEHEAD, M. & PETTICREW, M. 2009. Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews. *Journal of Epidemiology and Community health*, jech. 2008.082743.
- BANG, H., EDWARDS, A. M., BOMBACK, A. S., BALLANTYNE, C. M., BRILLON, D., CALLAHAN, M. A., TEUTSCH, S. M., MUSHLIN, A. I. & KERN, L. M. 2009. Development and validation of a patient self-assessment score for diabetes risk. *Annals of internal medicine*, 151, 775–783.
- BANI, I. A. 2015. Prevalence, knowledge, attitude and practices of diabetes mellitus among Jazan population, Kingdom of Saudi Arabia (KSA). *Journal of diabetes mellitus*, 5, 115.
- BARBARA, S. & KRASS, I. 2013. Self management of type 2 diabetes by M altese immigrants in Australia: can community pharmacies play a supporting role? *International Journal of Pharmacy Practice*, 21, 305–313.
- BARENGO, N. C., TAMAYO, D. C., TONO, T. & TUOMILEHTO, J. 2017. A Colombian diabetes risk score for detecting undiagnosed diabetes and impaired glucose regulation. *Primary care diabetes*, 11, 86–93.
- BEAGLEY, J., GUARIGUATA, L., WEIL, C. & MOTALA, A. A. 2014. Global estimates of undiagnosed diabetes in adults. *Diabetes Res Clin Pract*, 103, 150–60.
- BEENACKERS, M. A., KAMPHUIS, C. B., GISKES, K., BRUG, J., KUNST, A. E., BURDORF, A. & VAN LENTHE, F. J. 2012. Socioeconomic inequalities in occupational, leisure-time, and transport related physical activity among European adults: a systematic review. *International journal of behavioral nutrition and physical activity*, 9, 116.
- BENER, A., ZIRIE, M., JANAHI, I. M., AL-HAMAQ, A. O., MUSALLAM, M. & WAREHAM, N. J. 2009. Prevalence of diagnosed and undiagnosed diabetes mellitus and its risk factors in a population-based study of Qatar. *Diabetes research and clinical practice*, 84, 99–106.
- BENNETT, P. 2007. New data, fresh perspectives: diabetes atlas. *Diabetes voice*, 52, 46–48.
- BERNABE-ORTIZ, A., SMEETH, L., GILMAN, R. H., SANCHEZ-ABANTO, J. R., CHECKLEY, W. & MIRANDA, J. J. 2016. Development and Validation of a Simple Risk Score for Undiagnosed Type 2 Diabetes in a Resource-Constrained Setting. *Journal of diabetes research*, 2016.

- BHOPAL, R., HAYES, L., WHITE, M., UNWIN, N., HARLAND, J., AYIS, S. & ALBERTI, G. 2002. Ethnic and socio-economic inequalities in coronary heart disease, diabetes and risk factors in Europeans and South Asians. *Journal of Public Health*, 24, 95-105.
- BI, Y., ZHU, D., CHENG, J., ZHU, Y., XU, N., CUI, S., LI, W., CHENG, X., WANG, F. & HU, Y. 2010. The status of glycemic control: a cross-sectional study of outpatients with type 2 diabetes mellitus across primary, secondary, and tertiary hospitals in the Jiangsu province of China. *Clinical therapeutics*, 32, 973-983.
- BORRELL, L., KUNZEL, C., LAMSTER, I. & LALLA, E. 2007. Diabetes in the dental office: using NHANES III to estimate the probability of undiagnosed disease. *Journal of periodontal research*, 42, 559-565.
- BOYLE, J. P., THOMPSON, T. J., GREGG, E. W., BARKER, L. E. & WILLIAMSON, D. F. 2010. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. *Population health metrics*, 8, 29.
- BROWN, A. F., ETTNER, S. L., PIETTE, J., WEINBERGER, M., GREGG, E., SHAPIRO, M. F., KARTER, A. J., SAFFORD, M., WAITZFELDER, B., PRATA, P. A. & BECKLES, G. L. 2004. Socioeconomic position and health among persons with diabetes mellitus: a conceptual framework and review of the literature. *Epidemiol Rev*, 26, 63-77.
- BURR, M., TREMBETH, J., JONES, K., GEEN, J., LYNCH, L. & ROBERTS, Z. 2007. The effects of dietary advice and vouchers on the intake of fruit and fruit juice by pregnant women in a deprived area: a controlled trial. *Public health nutrition*, 10, 559-565.
- CALIXTO, O. J. & ANAYA, J. M. 2014. Socioeconomic status. The relationship with health and autoimmune diseases. *Autoimmunity Reviews*, 13, 641-654.
- CHATURVEDI, V., REDDY, K., PRABHAKARAN, D., JEEMON, P., RAMAKRISHNAN, L., SHAH, P. & SHAH, B. 2008. Development of a clinical risk score in predicting undiagnosed diabetes in urban Asian Indian adults: a population-based study. *CVD prevention and control*, 3, 141-151.
- CHEN, M.-F., WANG, R.-H., LIN, K.-C., HSU, H.-Y. & CHEN, S.-W. 2015. Efficacy of an empowerment program for Taiwanese patients with type 2 diabetes: A randomized controlled trial. *Applied Nursing Research*, 28, 366-373.
- CHIEN, K.-L., LIN, H.-J., LEE, B.-C., HSU, H.-C. & CHEN, M.-F. 2010. Prediction model for high glycated hemoglobin concentration among ethnic Chinese in Taiwan. *Cardiovascular diabetology*, 9, 59.
- CHO, N., SHAW, J., KARURANGA, S., HUANG, Y., DA ROCHA FERNANDES, J., OHLROGGE, A. & MALANDA, B. 2018. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes research and clinical practice*, 138, 271-281.
- CLARK, M. L. & UTZ, S. W. 2014. Social determinants of type 2 diabetes and health in the United States. *World J Diabetes*, 5, 296-304.
- COLLIER, A., GHOSH, S., HAIR, M. & WAUGH, N. 2015. Impact of socioeconomic status and gender on glycaemic control, cardiovascular risk factors and diabetes complications in type 1 and 2 diabetes: a population based analysis from a Scottish region. *Diabetes & metabolism*, 41, 145-151.
- COLLINS, G. S., MALLETT, S., OMAR, O. & YU, L.-M. 2011. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC medicine*, 9, 103.

List of References

- CONNOLLY, V., UNWIN, N., SHERRIFF, P., BILOUS, R. & KELLY, W. 2000. Diabetes prevalence and socioeconomic status: a population based study showing increased prevalence of type 2 diabetes mellitus in deprived areas. *Journal of Epidemiology and Community Health*, 54, 173-177.
- CONTROL, C. F. D. & PREVENTION 2006. State-specific prevalence of obesity among adults--United States, 2005. *MMWR. Morbidity and mortality weekly report*, 55, 985.
- CORSI, D. J. & SUBRAMANIAN, S. V. 2012. Association between socioeconomic status and self-reported diabetes in India: a cross-sectional multilevel analysis. *BMJ open*, 2, e000895.
- COSTANIAN, C., BENNETT, K., HWALLA, N., ASSAAD, S. & SIBAI, A. M. 2014. Prevalence, correlates and management of type 2 diabetes mellitus in Lebanon: Findings from a national population-based study. *Diabetes Research and Clinical Practice*, 105, 408-415.
- COWIE, C. C., RUST, K. F., BYRD-HOLT, D. D., GREGG, E. W., FORD, E. S., GEISS, L. S., BAINBRIDGE, K. E. & FRADKIN, J. E. 2010. Prevalence of diabetes and high risk for diabetes using A1C criteria in the US population in 1988–2006. *Diabetes care*, 33, 562-568.
- CRAIG, C. L., MARSHALL, A. L., SJORSTROM, M., BAUMAN, A. E., BOOTH, M. L., AINSWORTH, B. E., PRATT, M., EKLUND, U., YNGVE, A. & SALLIS, J. F. 2003. International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise*, 35, 1381-1395.
- D'SOUZA, M. S., KARKADA, S. N., HANRAHAN, N. P., VENKATESAPERUMAL, R. & AMIRTHARAJ, A. 2015. Do perceptions of empowerment affect glycemic control and self-care among adults with type 2 diabetes? *Global journal of health science*, 7, 80.
- DAHLGREN, G. & WHITEHEAD, M. 1991. Policies and strategies to promote social equity in health. *Stockholm: Institute for future studies*.
- DAR-ODEH, N. S. & ABU-HAMMAD, O. A. 2011. The changing trends in tobacco smoking for young Arab women; narghile, an old habit with a liberal attitude. *Harm reduction journal*, 8, 24.
- DARYANI, A., BERGLUND, L., ANDERSSON, Å., KOCTURK, T., BECKER, W. & VESSBY, B. 2005. Risk factors for coronary heart disease among immigrant women from Iran and Turkey, compared to women of Swedish ethnicity. *Ethnicity & disease*, 15, 213-220.
- DAVIES, M. J., HELLER, S., SKINNER, T., CAMPBELL, M., CAREY, M., CRADOCK, S., DALLOSSO, H., DALY, H., DOHERTY, Y. & EATON, S. 2008. Effectiveness of the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cluster randomised controlled trial. *Bmj*, 336, 491-495.
- DE BEL-AIR, F. 2014. Demography, migration and labour market in Saudi Arabia.
- DE SOUSA, A. G. P., PEREIRA, A. C., MARQUEZINE, G. F., DO NASCIMENTO-NETO, R. M., FREITAS, S. N., NICOLATO, R. L. D. C., MACHADO-COELHO, G. L. L., RODRIGUES, S. L., MILL, J. G. & KRIEGER, J. E. 2009. Derivation and external validation of a simple prediction model for the diagnosis of type 2 diabetes mellitus in the Brazilian urban population. *European journal of epidemiology*, 24, 101.
- EFRON, B. & TIBSHIRANI, R. J. 1994. *An introduction to the bootstrap*, CRC press.
- EL-HAZMI, M., AL-SWAILEM, A., WARSY, A., AL-SUDAIRY, F., SULAIMANI, R. & AL-MESHARI, A. 1995a. The prevalence of diabetes mellitus and impaired glucose tolerance in the population of Riyadh. *Annals of Saudi medicine*, 15, 598-601.

- EL-HAZMI, M., AL-SWAILEM, A., WARSY, A., AL-SWAILEM, A., SULAIMANI, R. & AL-MESHARI, A. 1995b. Consanguinity among the Saudi Arabian population. *Journal of medical genetics*, 32, 623-626.
- EL-HAZMI, M., WARSY, A., AL-SWAILEM, A., AL-SWAILEM, A., SULAIMANI, R. & AL-MESHARI, A. 1996a. Diabetes mellitus and impaired glucose tolerance in Saudi Arabia. *Annals of Saudi medicine*, 16, 381-385.
- EL-HAZMI, M., WARSY, A., BARIMAH, N. A., AL-SWAILEM, A., AL-SWAILEM, A., SULAIMANI, R. & AL-MESHARI, A. 1996b. The prevalence of diabetes mellitus and impaired glucose tolerance in the population of Al-Baha, Saudi Arabia. *Saudi medical journal*, 17, 591-597.
- EL-MOUZAN, M. I., AL-SALLOUM, A. A., AL-HERBISH, A. S., QURACHI, M. M. & AL-OMAR, A. A. 2007. Regional variations in the prevalence of consanguinity in Saudi Arabia. *Saudi Medical Journal*, 28, 1881-1884.
- EL BCHERAOU, C., BASULAIMAN, M., TUFFAHA, M., DAOUD, F., ROBINSON, M., JABER, S., MIKHITARIAN, S., MEMISH, Z. A., AL SAEEDI, M., ALMAZROA, M. A. & MOKDAD, A. H. 2014. Status of the diabetes epidemic in the Kingdom of Saudi Arabia, 2013. *Int J Public Health*, 59, 1011-21.
- EL HAZMI, M. A. 1998. Diabetes mellitus as a health problem in Saudi Arabia.
- EL MOUZAN, M., AL SALLOUM, A., AL HERBISH, A., QURACHI, M. & AL OMAR, A. 2008. Consanguinity and major genetic disorders in Saudi children: a community-based cross-sectional study. *Annals of Saudi medicine*, 28, 169.
- EL MOUZAN, M. I., AL HERBISH, A. S., AL SALLOUM, A. A., AL OMAR, A. A. & QURACHI, M. M. 2012. Regional variation in prevalence of overweight and obesity in Saudi children and adolescents. *Saudi journal of gastroenterology: official journal of the Saudi Gastroenterology Association*, 18, 129.
- ERCAN, A. & KIZILTAN, G. 2013. Obesity-related abnormal eating behaviors in Type 2 diabetic patients. *Pakistan journal of medical sciences*, 29, 1323.
- ESCOLAR, A. P. 2009. Social determinants vs. lifestyle in type 2 diabetes mellitus in Andalusia (Spain): difficulty in making ends meet or obesity? *Gaceta sanitaria*, 23, 427-432.
- ESPELT, A., ARRIOLA, L., BORRELL, C., LARRANAGA, I., SANDIN, M. & ESCOLAR-PUJOLAR, A. 2011. Socioeconomic position and type 2 diabetes mellitus in Europe 1999-2009: a panorama of inequalities. *Current diabetes reviews*, 7, 148-158.
- ESPELT, A., BORRELL, C., ROSKAM, A. J., RODRIGUEZ-SANZ, M., STIRBU, I., DALMAU-BUENO, A., REGIDOR, E., BOPP, M., MARTIKAINEN, P., LEINSALU, M., ARTNIK, B., RYCHTARIKOVA, J., KALEDIENE, R., DZUROVA, D., MACKENBACH, J. & KUNST, A. E. 2008. Socioeconomic inequalities in diabetes mellitus across Europe at the beginning of the 21st century. *Diabetologia*, 51, 1971-9.
- EVANS, J., NEWTON, R. W., RUTA, D. A., MACDONALD, T. M. & MORRIS, A. D. 2000. Socio-economic status, obesity and prevalence of Type 1 and Type 2 diabetes mellitus. *Diabetic Medicine*, 17, 478-480.
- EZEAMAMA, A. E., VIALI, S., TUIELE, J. & MCGARVEY, S. T. 2006. The influence of socioeconomic factors on cardiovascular disease risk factors in the context of economic development in the Samoan archipelago. *Social science & medicine*, 63, 2533-2545.

List of References

- EZZATI, M., LOPEZ, A. D., RODGERS, A., VANDER HOORN, S. & MURRAY, C. J. 2002. Selected major risk factors and global and regional burden of disease. *The Lancet*, 360, 1347-1360.
- FALLATAH, M. O. 2018. Knowledge, awareness, and eye care-seeking behavior in diabetic retinopathy: a cross-sectional study in Jeddah, Kingdom of Saudi Arabia. *Ophthalmology and therapy*, 7, 377-385.
- FATANI, H. H., MIRA, S. A. & EL-ZUBIER, A. G. 1987. Prevalence of diabetes mellitus in rural Saudi Arabia. *Diabetes care*, 10, 180-183.
- FIKREE, M., HANAFI, B., HUSSAIN, Z. A. & MASUADI, E. M. 2006. Glycemic control of type 2 diabetes mellitus. *Bahrain Med Bull*, 28, 105-7.
- FISHER, L., GLASGOW, R. E. & STRYCKER, L. A. 2010. The relationship between diabetes distress and clinical depression with glycemic control among patients with type 2 diabetes. *Diabetes care*, 33, 1034-1036.
- FRITSCHI, C., BRONAS, U. G., PARK, C. G., COLLINS, E. G. & QUINN, L. 2017. Early declines in physical function among aging adults with type 2 diabetes. *Journal of Diabetes and its Complications*, 31, 347-352.
- GAO, J., WANG, J., ZHENG, P., HAARDÖRFER, R., KEGLER, M. C., ZHU, Y. & FU, H. 2013. Effects of self-care, self-efficacy, social support on glycemic control in adults with type 2 diabetes. *BMC family practice*, 14, 66.
- GAO, W., DONG, Y., PANG, Z., NAN, H., WANG, S., REN, J., ZHANG, L., TUOMILEHTO, J. & QIAO, Q. 2010. A simple Chinese risk score for undiagnosed diabetes. *Diabetic Medicine*, 27, 274-281.
- GASSASSE, Z., SMITH, D., FINER, S. & GALLO, V. 2017. Association between urbanisation and type 2 diabetes: an ecological study. *BMJ global health*, 2, e000473.
- GEYER, S., PETER, R. & NIELSEN, I. 2004. Health inequalities in different age groups: the case of type 2-diabetes: a study with health insurance and medication data. *Sozial-und Präventivmedizin*, 49, 328-335.
- GILES-CORTI, B., VERNEZ-MOUDON, A., REIS, R., TURRELL, G., DANNENBERG, A. L., BADLAND, H., FOSTER, S., LOWE, M., SALLIS, J. F. & STEVENSON, M. 2016. City planning and population health: a global challenge. *The lancet*, 388, 2912-2924.
- GILSON, L., BUSE, K., MURRAY, S. F. & DICKINSON, C. 2008. Future directions for health policy analysis: a tribute to the work of Professor Gill Walt. Oxford University Press.
- GLÜMER, C., CARSTENSEN, B., SANDBÆK, A., LAURITZEN, T., JØRGENSEN, T. & BORCH-JOHNSEN, K. 2004. A Danish diabetes risk score for targeted screening. *Diabetes care*, 27, 727-733.
- GNAVI, R., KARAGHIOSOFF, L., COSTA, G., MERLETTI, F. & BRUNO, G. 2008. Socio-economic differences in the prevalence of diabetes in Italy: the population-based Turin study. *Nutrition, Metabolism and Cardiovascular Diseases*, 18, 678-682.
- GRAY, L., TAUB, N., KHUNTI, K., GARDINER, E., HILES, S., WEBB, D., SRINIVASAN, B. & DAVIES, M. 2010. The Leicester Risk Assessment score for detecting undiagnosed type 2 diabetes and impaired glucose regulation for use in a multiethnic UK setting. *Diabetic Medicine*, 27, 887-895.
- GREENFIELD, S., BILLIMEK, J., PELLEGRINI, F., FRANCIOSI, M., DE BERARDIS, G., NICOLUCCI, A. & KAPLAN, S. H. 2009. Comorbidity affects the relationship between glycemic control and

- cardiovascular outcomes in diabetes: a cohort study. *Annals of Internal Medicine*, 151, 854-860.
- GRIFFIN, S., LITTLE, P., HALES, C., KINMONTH, A. & WAREHAM, N. 2000. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. *Diabetes/metabolism research and reviews*, 16, 164-171.
- GRINTSOVA, O., MAIER, W. & MIELCK, A. 2014. Inequalities in health care among patients with type 2 diabetes by individual socio-economic status (SES) and regional deprivation: a systematic literature review. *Int J Equity Health*, 13, 43.
- GROUP, U. P. D. S. 1998. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *The lancet*, 352, 837-853.
- GUARIGUATA, L., WHITING, D., WEIL, C. & UNWIN, N. 2011. The International Diabetes Federation diabetes atlas methodology for estimating global and national prevalence of diabetes in adults. *Diabetes research and clinical practice*, 94, 322-332.
- GUARIGUATA, L., WHITING, D. R., HAMBLETON, I., BEAGLEY, J., LINNENKAMP, U. & SHAW, J. E. 2014. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract*, 103, 137-49.
- GUNAID, A. & ASSABRI, A. 2008. Prevalence of type 2 diabetes and other cardiovascular risk factors in a semirural area in Yemen. *EMHJ-Eastern Mediterranean Health Journal*, 14 (1), 42-56, 2008.
- HAMDAN, A. 2005. Women and Education in Saudi Arabia: Challenges and Achievements. *International Education Journal*, 6, 42-64.
- HAMMER, G. P., DU PREL, J.-B. & BLETNER, M. 2009. Avoiding bias in observational studies: part 8 in a series of articles on evaluation of scientific publications. *Deutsches Ärzteblatt International*, 106, 664.
- HANDLOS, L. N., WITTE, D. R., ALMDAL, T. P., NIELSEN, L., BADAWI, S., SHEIKH, A., BELHADJ, M., NADIR, D., ZINAI, S. & VISTISEN, D. 2013. Risk scores for diabetes and impaired glycaemia in the Middle East and North Africa. *Diabetic Medicine*, 30, 443-451.
- HEGAZI, R., EL-GAMAL, M., ABDEL-HADY, N. & HAMDY, O. 2015. Epidemiology of and risk factors for type 2 diabetes in Egypt. *Annals of global health*, 81, 814-820.
- HEIANZA, Y., ARASE, Y., SAITO, K., HSIEH, S. D., TSUJI, H., KODAMA, S., TANAKA, S., OHASHI, Y., SHIMANO, H. & YAMADA, N. 2013. Development of a screening score for undiagnosed diabetes and its application in estimating absolute risk of future type 2 diabetes in Japan: Toranomon Hospital Health Management Center Study 10 (TOPICS 10). *The Journal of Clinical Endocrinology & Metabolism*, 98, 1051-1060.
- HEIKES, K. E., EDDY, D. M., ARONDEKAR, B. & SCHLESSINGER, L. 2008a. Diabetes risk calculator. *Diabetes care*, 31, 1040-1045.
- HEIKES, K. E., EDDY, D. M., ARONDEKAR, B. & SCHLESSINGER, L. 2008b. Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. *Diabetes care*, 31, 1040-1045.
- HELMS, R. B. 1992. Implications of population growth on prevalence of diabetes: a look at the future. *Diabetes Care*, 15, 6-9.

List of References

- HELTBERG, A., ANDERSEN, J. S., SANDHOLDT, H., SIERSMA, V., KRAGSTRUP, J. & ELLERVIK, C. 2018. Predictors of undiagnosed prevalent type 2 diabetes—The Danish General Suburban Population Study. *Primary care diabetes*, 12, 13-22.
- HERMAN, D. R., HARRISON, G. G. & JENKS, E. 2006. Choices made by low-income women provided with an economic supplement for fresh fruit and vegetable purchase. *Journal of the American Dietetic Association*, 106, 740-744.
- HILTUNEN, L. 2005. Are there associations between socio-economic status and known diabetes in an elderly Finnish population? *Central European journal of public health*, 13, 187.
- HINDE, A. 2014. *Demographic methods*, Routledge.
- HOLFORD, T. R. 1991. Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annual review of public health*, 12, 425-457.
- HOLMAN, N., FOROUHI, N., GOYDER, E. & WILD, S. 2011. The Association of Public Health Observatories (APHO) diabetes prevalence model: estimates of total diabetes prevalence for England, 2010–2030. *Diabetic Medicine*, 28, 575-582.
- HONEYCUTT, A. A., BOYLE, J. P., BROGLIO, K. R., THOMPSON, T. J., HOERGER, T. J., GEISS, L. S. & NARAYAN, K. V. 2003. A dynamic Markov model for forecasting diabetes prevalence in the United States through 2050. *Health care management science*, 6, 155-164.
- HOULE, J., LAUZIER-JOBIN, F., BEAULIEU, M.-D., MEUNIER, S., COULOMBE, S., CÔTÉ, J., LESPÉRANCE, F., CHIASSEON, J.-L., BHERER, L. & LAMBERT, J. 2016. Socioeconomic status and glycemic control in adult patients with type 2 diabetes: a mediation analysis. *BMJ Open Diabetes Research and Care*, 4, e000184.
- HUANG, E. S., BASU, A., O'GRADY, M. & CAPRETTA, J. C. 2009. Projecting the future diabetes population size and related costs for the US. *Diabetes care*, 32, 2225-2229.
- HUSSAIN, A., CLAUSSEN, B., RAMACHANDRAN, A. & WILLIAMS, R. 2007. Prevention of type 2 diabetes: a review. *Diabetes Res Clin Pract*, 76, 317-26.
- HWANG, J. & SHON, C. 2014. Relationship between socioeconomic status and type 2 diabetes: results from Korea National Health and Nutrition Examination Survey (KNHANES) 2010–2012. *BMJ open*, 4, e005710.
- ICKS, A., MOEBUS, S., FEUERSENGER, A., HAASTERT, B., JÖCKEL, K.-H. & GIANI, G. 2007. Diabetes prevalence and association with social status—Widening of a social gradient?: German National Health Surveys 1990–1992 and 1998. *Diabetes research and clinical practice*, 78, 293-297.
- JEON, C. Y. & MURRAY, M. B. 2008. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. *PLoS medicine*, 5, e152.
- KADIKI, O. & ROAEID, R. 2001. Prevalence of diabetes mellitus and impaired glucose tolerance in Benghazi Libya. *Diabetes & metabolism*, 27, 647-654.
- KARIM, A., OGBEIDE, D. O., SIDDIQUI, S. & AL-KHALIFA, I. M. 2000. Prevalence of diabetes mellitus in a Saudi community. *Saudi medical journal*, 21, 438-442.
- KEESUKPHAN, P., CHANPRASERTYOTHIN, S., ONGPHIPHADHANAKUL, B. & PUAVILAI, G. 2007. The development and validation of a diabetes risk score for high-risk Thai adults. *Journal-Medical Association of Thailand*, 90, 149.

- KHATIB, O. 2004. Noncommunicable diseases: risk factors and regional strategies for prevention and care.
- KHATIB, O. 2006. Guidelines for the prevention, management and care of diabetes mellitus: World Health Organization. *Regional Office for the Eastern Mediterranean [EMRO Technical Publication Series, No. 32]*.
- KHATTAB, M., KHADER, Y. S., AL-KHAWALDEH, A. & AJLOUNI, K. 2010. Factors associated with poor glycemic control among patients with type 2 diabetes. *Journal of diabetes and its complications*, 24, 84-89.
- KHOJA, A. T., ALJAWADI, M. H., AL-SHAMMARI, S. A., MOHAMED, A. G., AL-MANAA, H. A., MORLOCK, L., AHMED, S. & KHOJA, T. A. 2018. The health of Saudi older adults; results from the Saudi National Survey for Elderly Health (SNSEH) 2006–2015. *Saudi Pharmaceutical Journal*, 26, 292-300.
- KHOJA, T. L. Strategic approaches in combating diabetes mellitus among GCC countries. first BA Regional Workshop on the Epidemiology of Diabetes and Other Non-Communicable Diseases, 2009.
- KHOJA, T., RAWAF, S., QIDWAI, W., RAWAF, D., NANJI, K. & HAMAD, A. 2017. Health care in Gulf Cooperation Council countries: a review of challenges and opportunities. *Cureus*, 9.
- KHRAIF, R. 2007. Urbanization and Growth of Cities in Saudi Arabia. *Kuwaiti Geographical Society, Kuwait*.
- KHRAIF, R. M. 1994. *Residential mobility in the city of Riyadh: a study of its directions, reasons, and characteristics*, Saudi Geographical Society.
- KHRAIF, R. M., SALAM, A. A., ELSEGAIEY, I. & ALMUTAIRI, A. 2015. Changing age structures and ageing scenario of the Arab world. *Social Indicators Research*, 121, 763-785.
- KIM, S., EGERTER, S., CUBBIN, C., TAKAHASHI, E. R. & BRAVEMAN, P. 2007. Potential implications of missing income data in population-based surveys: an example from a postpartum survey in California. *Public health reports*, 122, 753-763.
- KIM, Y. J., JEON, J. Y., HAN, S. J., KIM, H. J., LEE, K. W. & KIM, D. J. 2015. Effect of socio-economic status on the prevalence of diabetes. *Yonsei medical journal*, 56, 641-647.
- KING, H., AUBERT, R. E. & HERMAN, W. H. 1998. Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. *Diabetes care*, 21, 1414-1431.
- KO, G., SO, W., TONG, P., MA, R., KONG, A., OZAKI, R., CHOW, C., COCKRAM, C. & CHAN, J. 2010. A simple risk score to identify Southern Chinese at high risk for diabetes. *Diabetic Medicine*, 27, 644-649.
- KUHA, J. 2004. AIC and BIC: Comparisons of assumptions and performance. *Sociological methods & research*, 33, 188-229.
- KUMARI, M., HEAD, J. & MARMOT, M. 2004. Prospective study of social and other risk factors for incidence of type 2 diabetes in the Whitehall II study. *Archives of internal medicine*, 164, 1873-1880.
- LALKHEN, A. G. & MCCLUSKEY, A. 2008. Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain*, 8, 221-223.

List of References

- LANTING, L. C., JOUNG, I. M., VOGEL, I., BOOTSMA, A. H., LAMBERTS, S. & MACKENBACH, J. P. 2008. Ethnic differences in outcomes of diabetes care and the role of self-management behavior. *Patient education and counseling*, 72, 146-154.
- LARRANAGA, I., ARTEAGOITIA, J., RODRIGUEZ, J., GONZALEZ, F., ESNAOLA, S., PINIES, J. & COUNTRY, S. P. N. O. T. B. 2005. Socio-economic inequalities in the prevalence of Type 2 diabetes, cardiovascular risk factors and chronic diabetic complications in the Basque Country, Spain. *Diabetic Medicine*, 22, 1047-1053.
- LAU, R. S., OHINMAA, A. & JOHNSON, J. A. 2011. Predicting the future burden of diabetes in Alberta from 2008 to 2035. *Canadian Journal of Diabetes*, 35, 274-281.
- LAWLOR, D., PATEL, R., FRASER, A., SMITH, G. D. & EBRAHIM, S. 2007. The association of life course socio-economic position with diagnosis, treatment, control and survival of women with diabetes: findings from the British Women's Heart and Health Study. *Diabetic medicine*, 24, 892-900.
- LEE, Y.-J., SHIN, S.-J., WANG, R.-H., LIN, K.-D., LEE, Y.-L. & WANG, Y.-H. 2016. Pathways of empowerment perceptions, health literacy, self-efficacy, and self-care behaviors to glycemic control in patients with type 2 diabetes mellitus. *Patient education and counseling*, 99, 287-294.
- LFOTOUH, A., SOLIMAN, L., MANSOUR, E., FARGHALY, M. & EL DAWAIATY, A. 2008. Central obesity among adults in Egypt: prevalence and associated morbidity.
- LIDFELDT, J., LI, T. Y., HU, F. B., MANSON, J. E. & KAWACHI, I. 2007. A prospective study of childhood and adult socioeconomic status and incidence of type 2 diabetes in women. *American journal of epidemiology*, 165, 882-889.
- LINDSTRÖM, J., ILANNE-PARIKKA, P., PELTONEN, M., AUNOLA, S., ERIKSSON, J. G., HEMIÖ, K., HÄMÄLÄINEN, H., HÄRKÖNEN, P., KEINÄNEN-KIUKAANNIEMI, S. & LAAKSO, M. 2006. Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finnish Diabetes Prevention Study. *The Lancet*, 368, 1673-1679.
- LITTLE, R. J. & RUBIN, D. B. 2014. *Statistical analysis with missing data*, John Wiley & Sons.
- LYNCH, S. M. 2003. Cohort and life-course patterns in the relationship between education and health: A hierarchical approach. *Demography*, 40, 309-331.
- MABRY, R., KOOHSARI, M. J., BULL, F. & OWEN, N. 2016. A systematic review of physical activity and sedentary behaviour research in the oil-producing countries of the Arabian Peninsula. *BMC Public Health*, 16, 1003.
- MABRY, R., REEVES, M. M., EAKIN, E. G. & OWEN, N. 2010. Evidence of physical activity participation among men and women in the countries of the Gulf Cooperation Council: a review. *Obesity reviews*, 11, 457-464.
- MAGLIANO, D., SHAW, J., SHORTREED, S., NUSSELDER, W., LIEW, D., BARR, E., ZIMMET, P. & PEETERS, A. 2008. Lifetime risk and projected population prevalence of diabetes. *Diabetologia*, 51, 2179-2186.
- MAINOUS, A. G., BAKER, R., KOOPMAN, R. J., SAXENA, S., DIAZ, V. A., EVERETT, C. J. & MAJEED, A. 2007. Impact of the population at risk of diabetes on projections of diabetes burden in the United States: an epidemic on the way. *Diabetologia*, 50, 934-940.
- MAJEED, A., EL-SAYED, A. A., KHOJA, T., ALSHAMSAN, R., MILLETT, C. & RAWAF, S. 2014. Diabetes in the Middle-East and North Africa: an update. *Diabetes Res Clin Pract*, 103, 218-22.

- MALIK, M., BAKIR, A., AL-HAMADI, H. A., PUGH, R., HOSSAIN, M., ABSOOD, G. & GILL, G. 1999. Factors associated with diabetes and its control in a newly developed country. *Practical Diabetes International*, 16, 250-252.
- MALIK, M., BAKIR, A., SAAB, B. A., ROGLIC, G. & KING, H. 2005. Glucose intolerance and associated factors in the multi-ethnic population of the United Arab Emirates: results of a national survey. *Diabetes research and clinical practice*, 69, 188-195.
- MANSOUR, A. A., AL-MALIKY, A. A., KASEM, B., JABAR, A. & MOSBEH, K. A. 2014. Prevalence of diagnosed and undiagnosed diabetes mellitus in adults aged 19 years and older in Basrah, Iraq. *Diabetes, metabolic syndrome and obesity: targets and therapy*, 7, 139.
- MARMOT, M. 2005. Social determinants of health inequalities. *The Lancet*, 365, 1099-1104.
- MARMOT, M. & HEALTH, C. O. S. D. O. 2007. Achieving health equity: from root causes to fair outcomes. *The Lancet*, 370, 1153-1163.
- MARTINSEN, N., JØRGENSEN, M., BJERREGAARD, P., KRASNIK, A., CARSTENSEN, B. & BORCH-JOHNSSEN, K. 2006. Predictions of type 2 diabetes and complications in Greenland in 2014. *International journal of circumpolar health*, 65, 243-252.
- MATTHEWS, S. M., PEDEN, A. R. & ROWLES, G. D. 2009. Patient-provider communication: Understanding diabetes management among adult females. *Patient education and counseling*, 76, 31-37.
- MEHIO SIBAI, A., NASREDDINE, L., MOKDAD, A. H., ADRA, N., TABET, M. & HWALLA, N. 2010. Nutrition transition and cardiovascular disease risk factors in Middle East and North Africa countries: reviewing the evidence. *Annals of Nutrition and Metabolism*, 57, 193-203.
- MEMISH, Z. A., CHANG, J. L., SAEEDI, M. Y., AL HAMID, M. A., ABID, O. & ALI, M. K. 2015. Screening for type 2 diabetes and dysglycemia in Saudi Arabia: development and validation of risk scores. *Diabetes technology & therapeutics*, 17, 693-700.
- MIDHET, F. M., AL-MOHAIMEED, A. A. & SHARAF, F. K. 2010. Lifestyle related risk factors of type 2 diabetes mellitus in Saudi Arabia. *Saudi medical journal*, 31, 768-774.
- MOHAN, V., DEEPA, R., DEEPA, M., SOMANNAVAR, S. & DATTA, M. 2005. A simplified Indian Diabetes Risk Score for screening for undiagnosed diabetic subjects. *JaPi*, 53, 759-763.
- MOKDAD, A. H., JABER, S., AZIZ, M. I. A., ALBUHAIRAN, F., ALGHAITHI, A., ALHAMAD, N. M., AL-HOOTI, S. N., AL-JASARI, A., ALMAZROA, M. A. & ALQASMI, A. M. 2014. The state of health in the Arab world, 1990–2010: an analysis of the burden of diseases, injuries, and risk factors. *The Lancet*, 383, 309-320.
- MOORE, M., GOULD, P. & KEARY, B. S. 2003. Global urbanization and impact on health. *International journal of hygiene and environmental health*, 206, 269-278.
- MORADI-LAKEH, M., EL BCHERAOU, C., AFSHIN, A., DAOUD, F., ALMAZROA, M. A., AL SAEEDI, M., BASULAIMAN, M., MEMISH, Z. A., AL RABEEAH, A. A. & MOKDAD, A. H. 2017. Diet in Saudi Arabia: findings from a nationally representative survey. *Public health nutrition*, 20, 1075-1081.
- MOURSHED, M., HEDIGER, V. & LAMBERT, T. 2006. Gulf cooperation council health care: challenges and opportunities. *Global competitiveness reports*.
- MOUSA, W. & ALDEHAYYAT, J. S. 2018. Regional efficiency of healthcare services in Saudi Arabia. *Middle East Development Journal*, 10, 152-174.

List of References

- MOUSSA, M. A., ALSAEID, M., ABDELLA, N., REFAI, T. M., AL-SHEIKH, N. & GOMEZ, J. E. 2008. Prevalence of type 2 diabetes mellitus among Kuwaiti children and adolescents. *Medical Principles and practice*, 17, 270-275.
- MURAD, M. A., ABDULMAGEED, S. S., IFTIKHAR, R. & SAGGA, B. K. 2014. Assessment of the common risk factors associated with type 2 diabetes mellitus in Jeddah. *International journal of endocrinology*, 2014.
- MUSAIGER, A. O. 2004. Overweight and obesity in the Eastern Mediterranean Region: can we control it?
- MUSAIGER, A. O., HASSAN, A. S. & OBEID, O. 2011. The Paradox of Nutrition-Related Diseases in the Arab Countries: The Need for Action. *International Journal of Environmental Research and Public Health*, 8, 3637-3671.
- NAGAYA, T., YOSHIDA, H., TAKAHASHI, H. & KAWAI, M. 2006. Policemen and firefighters have increased risk for type-2 diabetes mellitus probably due to their large body mass index: A follow-up study in Japanese men. *American journal of industrial medicine*, 49, 30-35.
- NARAYAN, K. V., BOYLE, J. P., GEISS, L. S., SAADDINE, J. B. & THOMPSON, T. J. 2006. Impact of recent increase in incidence on future diabetes burden: US, 2005–2050. *Diabetes care*, 29, 2114-2116.
- NASSER, J. 2007. Evaluation of diabetes care in a primary care setting. *Bahrain Med Bull*, 29, 45-49.
- NG, S. W., ZAGHLOUL, S., ALI, H., HARRISON, G. & POPKIN, B. M. 2011. The prevalence and trends of overweight, obesity and nutrition-related non-communicable diseases in the Arabian Gulf States. *Obesity Reviews*, 12, 1-13.
- NOOR, S., BUSHARA, S., SULAIMAN, A., ELMADHOUN, W. & AHMED, M. 2015. Undiagnosed diabetes mellitus in rural communities in Sudan: prevalence and risk factors. *Eastern Mediterranean Health Journal*, 21.
- OGDEN, C. L., CARROLL, M. D., CURTIN, L. R., MCDOWELL, M. A., TABAK, C. J. & FLEGAL, K. M. 2006. Prevalence of overweight and obesity in the United States, 1999-2004. *Jama*, 295, 1549-1555.
- OGURTSOVA, K., DA ROCHA FERNANDES, J., HUANG, Y., LINNENKAMP, U., GUARIGUATA, L., CHO, N., CAVAN, D., SHAW, J. & MAKAROFF, L. 2017. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes research and clinical practice*, 128, 40-50.
- ONG, W. M., CHUA, S. S. & NG, C. J. 2014. Barriers and facilitators to self-monitoring of blood glucose in people with type 2 diabetes using insulin: a qualitative study. *Patient preference and adherence*, 8, 237.
- WORLD HEALTH ORGANIZATION (WHO), 1998. The World Health Report 1998: Life in the 21st century a vision for all.
- WORLD HEALTH ORGANIZATION (WHO), 1999. Definition, diagnosis and classification of diabetes mellitus and its complications: report of a WHO consultation. Part 1, Diagnosis and classification of diabetes mellitus.
- WORLD HEALTH ORGANIZATION (WHO), 2006a. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WH.

- WORLD HEALTH ORGANIZATION (WHO),2006b. The world health report: 2006: working together for health.
- WORLD HEALTH ORGANIZATION (WHO),2009. *Global health risks: mortality and burden of disease attributable to selected major risks*, Geneva: World Health Organization.
- WORLD HEALTH ORGANIZATION (WHO),2011. Use of glycated haemoglobin (HbA1c) in diagnosis of diabetes mellitus: abbreviated report of a WHO consultation.
- WORLD HEALTH ORGANIZATION (WHO),2014. Global status report on noncommunicable diseases 2014. World Health Organization.
- WORLD HEALTH ORGANIZATION (WHO),2015. Noncommunicable diseases progress monitor 2015.
- WORLD HEALTH ORGANIZATION (WHO),2016a. *Global report on diabetes*, World Health Organization.
- WORLD HEALTH ORGANIZATION (WHO),2016b. World Health Organization obesity and overweight fact sheet.
- PATIL, R. R. 2014. Urbanization as a determinant of health: a socioepidemiological perspective. *Social work in public health*, 29, 335-341.
- PAVLOU, M., AMBLER, G., SEAMAN, S. R., GUTTMANN, O., ELLIOTT, P., KING, M. & OMAR, R. Z. 2015. How to develop a more accurate risk prediction model when there are few events. *Bmj*, 351, h3868.
- PICCOLO, R. S., PEARCE, N., ARAUJO, A. B. & MCKINLAY, J. B. 2014. The contribution of biogeographical ancestry and socioeconomic status to racial/ethnic disparities in type 2 diabetes mellitus: results from the Boston Area Community Health Survey. *Annals of epidemiology*, 24, 648-654. e1.
- POPKIN, B. M., ADAIR, L. S. & NG, S. W. 2012. Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition Reviews*, 70, 3-21.
- PRAMONO, L. A., SETIATI, S., SOEWONDO, P., SUBEKTI, I., ADISASMITA, A., KODIM, N. & SUTRISNA, B. 2010. Prevalence and predictors of undiagnosed diabetes mellitus in Indonesia. *Age*, 46, 100.0.
- PULVIRENTI, M., MCMILLAN, J. & LAWN, S. 2014. Empowerment, patient centred care and self-management. *Health Expectations*, 17, 303-310.
- RABI, D. M., EDWARDS, A. L., SOUTHERN, D. A., SVENSON, L. W., SARGIOUS, P. M., NORTON, P., LARSEN, E. T. & GHALI, W. A. 2006. Association of socio-economic status with diabetes prevalence and utilization of diabetes care services. *BMC Health Serv Res*, 6, 124.
- RAHMAN, M., NAKAMURA, K. & KIZUKI, M. 2015. Socioeconomic differences in the prevalence, awareness, and control of diabetes in Bangladesh. *Journal of diabetes and its complications*, 29, 788-793.
- RAMACHANDRAN, A., SNEHALATHA, C., VIJAY, V., WAREHAM, N. & COLAGIURI, S. 2005. Derivation and validation of diabetes risk score for urban Asian Indians. *Diabetes research and clinical practice*, 70, 63-70.
- REDMOND, E. H., BURNETT, S. M., JOHNSON, M. A., PARK, S., FISCHER, J. G. & JOHNSON, T. 2007. Improvement in A1C levels and diabetes self-management activities following a nutrition

List of References

- and diabetes education program in older adults. *Journal of Nutrition for the Elderly*, 26, 83-102.
- REISIG, V., REITMEIR, P., DÖRING, A., RATHMANN, W., MIELCK, A. & GROUP, K. S. 2007. Social inequalities and outcomes in type 2 diabetes in the German region of Augsburg. A cross-sectional survey. *International journal of public health*, 52, 158-165.
- ROBARDS, J., EVANDROU, M., FALKINGHAM, J. & VLACHANTONI, A. 2012. Marital status, health and mortality. *Maturitas*, 73, 295-299.
- ROBBINS, J. M., VACCARINO, V., ZHANG, H. & KASL, S. V. 2005. Socioeconomic status and diagnosed diabetes incidence. *Diabetes research and clinical practice*, 68, 230-236.
- ROBERTSON, C. C. & RICH, S. S. 2018. Genetics of type 1 diabetes. *Current opinion in genetics & development*, 50, 7-16.
- ROMDHANE, H. B., ALI, S. B., AISSI, W., TRAISSAC, P., AOUNALLAH-SKHIRI, H., BOUGATEF, S., MAIRE, B., DELPEUCH, F. & ACHOUR, N. 2014. Prevalence of diabetes in Northern African countries: the case of Tunisia. *BMC Public Health*, 14, 86.
- ROTHMAN, R. L., MULVANEY, S., ELASY, T. A., VANDERWOUDE, A., GEBRETSADIK, T., SHINTANI, A., POTTER, A., RUSSELL, W. E. & SCHLUNDT, D. 2008. Self-management behaviors, racial disparities, and glycemic control among adolescents with type 2 diabetes. *Pediatrics*, 121, e912-e919.
- RUIGE, J. B., DE NEELING, J. N. D., KOSTENSE, P. J., BOUTER, L. M. & HEINE, R. J. 1997. Performance of an NIDDM screening questionnaire based on symptoms and risk factors. *Diabetes Care*, 20, 491-496.
- RUWAARD, D., HOOGENVEEN, R. T., VERKLEIJ, H., KROMHOUT, D., CASPARIE, A. F. & VAN DER VEEN, E. A. 1993. Forecasting the number of diabetic patients in The Netherlands in 2005. *American journal of public health*, 83, 989-995.
- SAADI, H., CARRUTHERS, S. G., NAGELKERKE, N., AL-MASKARI, F., AFANDI, B., REED, R., LUKIC, M., NICHOLLS, M. G., KAZAM, E. & ALGAWI, K. 2007. Prevalence of diabetes mellitus and its complications in a population-based sample in Al Ain, United Arab Emirates. *Diabetes research and clinical practice*, 78, 369-377.
- SAARISTO, T., PELTONEN, M., KEINÄNEN-KIUKAANNIEMI, S., VANHALA, M., SALTEVO, J., NISKANEN, L., OKSA, H., KORPI-HYÖVÄLTİ, E. & TUOMILEHTO, J. 2007. National type 2 diabetes prevention programme in Finland: FIN-D2D. *International journal of circumpolar health*, 66, 101-112.
- SABRA, A. A., TAHA, A., AL-ZUBIER, A. G. & AL-KURASHI, N. Y. 2010. Misconceptions about diabetes mellitus among adult male attendees of primary health care centres in Eastern Saudi Arabia. *South African Family Practice*, 52, 344-349.
- SACKS, D. B. 2011. A1C versus glucose testing: a comparison. *Diabetes care*, 34, 518-523.
- SALAM, A. A. 2019. Health Impact of Demographic Changes in the Gulf States. *Handbook of Healthcare in the Arab World*, 1-22.
- SALAM, A. A., ELSEGAIEY, I., KHRAIF, R. & AL-MUTAIRI, A. 2014. Population distribution and household conditions in Saudi Arabia: reflections from the 2010 Census. *SpringerPlus*, 3, 1.

- SCHMUHL, H., DEMSKI, H., LAMPRINOS, I., DOGAC, A., PLOESSNIG, M. & HILDEBRAND, C. 2014. Concept of knowledge-based self-management pathways for the empowerment of diabetes patients. *EJBI*, 10, 12-16.
- SEIDELL, J. C. 2000. Obesity, insulin resistance and diabetes—a worldwide epidemic. *British Journal of Nutrition*, 83, S5-S8.
- SHAH, N. M. 2012. Socio-demographic transitions among nationals of GCC countries: Implications for migration and labour force trends. *Migration and Development*, 1, 138-148.
- SHAMSI, N., SHEHAB, Z., ALNAHASH, Z., ALMUHANADI, S. & AL-NASIR, F. 2013. Factors influencing dietary practice among type 2 diabetics. *Bahrain Med Bull*, 35, 1-14.
- SHARAF, F. K. 2013. Impact of Health Education on Hba1c Level Among Diabetic Patients in Al-Qassim Region, Saudi Arabia. *Annals of Alquds Medicine*, 9, 5-12.
- SHAW, B. A. & SPOKANE, L. S. 2008. Examining the association between education level and physical activity changes during early old age. *Journal of aging and health*, 20, 767-787.
- SHAW, J. E., SICREE, R. A. & ZIMMET, P. Z. 2010. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract*, 87, 4-14.
- SICREE, R. & SHAW, J. 2007. Type 2 diabetes: An epidemic or not, and why it is happening. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 1, 75-81.
- SICREE, R., SHAW, J., ZIMMET, P. & TAPP, R. 2003. The global burden of diabetes. *Diabetes atlas*, 2, 15-71.
- SILLES, M. A. 2009. The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education review*, 28, 122-128.
- SKELLY, A. C., DETTORI, J. R. & BRODT, E. D. 2012. Assessing bias: the importance of considering confounding. *Evidence-based spine-care journal*, 3, 9-12.
- SOLAR, O. & IRWIN, A. 2007. A conceptual framework for action on the social determinants of health.
- SOMMER, I., GRIEBLER, U., MAHLKNECHT, P., THALER, K., BOUSKILL, K., GARTLEHNER, G. & MENDIS, S. 2015. Socioeconomic inequalities in non-communicable diseases and their risk factors: an overview of systematic reviews. *Bmc Public Health*, 15, 12.
- STEYERBERG, E. W. 2008. *Clinical prediction models: a practical approach to development, validation, and updating*, Springer Science & Business Media.
- STEYERBERG, E. W., HARRELL, F. E., BORSBOOM, G. J., EIJKEMANS, M., VERGOUWE, Y. & HABBEMA, J. D. F. 2001. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of clinical epidemiology*, 54, 774-781.
- STEYERBERG, E. W., VICKERS, A. J., COOK, N. R., GERDS, T., GONEN, M., OBUCHOWSKI, N., PENCINA, M. J. & KATTAN, M. W. 2010. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, 21, 128.
- STRINGHINI, S., TABAK, A. G., AKBARALY, T. N., SABIA, S., SHIPLEY, M. J., MARMOT, M. G., BRUNNER, E. J., BATTY, G. D., BOVET, P. & KIVIMAKI, M. 2012. Contribution of modifiable risk factors to social inequalities in type 2 diabetes: prospective Whitehall II cohort study. *BMJ*, 345, e5452.

List of References

- SU, R., CAI, L., CUI, W., HE, J., YOU, D. & GOLDEN, A. 2016. Multilevel Analysis of Socioeconomic Determinants on Diabetes Prevalence, Awareness, Treatment and Self-Management in Ethnic Minorities of Yunnan Province, China. *International Journal of Environmental Research and Public Health*, 13, 751.
- SULAIMAN, N., MAHMOUD, I., HUSSEIN, A., ELBADAWI, S., ABUSNANA, S., ZIMMET, P. & SHAW, J. 2018. Diabetes risk score in the United Arab Emirates: a screening tool for the early detection of type 2 diabetes mellitus. *BMJ Open Diabetes Research and Care*, 6, e000489.
- SUTTON, R. 1999. The Policy Process: An Overview ', ODI Working Paper 118. London: Overseas Development Institute.
- TABAEI, B. P. & HERMAN, W. H. 2002. A multivariate logistic regression equation to screen for diabetes. *Diabetes Care*, 25, 1999-2003.
- TAMAYO, T., HERDER, C. & RATHMANN, W. 2010. Impact of early psychosocial factors (childhood socioeconomic factors and adversities) on future risk of type 2 diabetes, metabolic disturbances and obesity: a systematic review. *BMC Public Health*, 10, 1.
- TANG, M., CHEN, Y. & KREWSKI, D. 2003. Gender-related differences in the association between socioeconomic status and self-reported diabetes. *International Journal of Epidemiology*, 32, 381-385.
- TAYLOR, J. 2009. The changing health of the Middle East population through oil and automobiles. *European heart journal*, 30, 1291-1293.
- TAYYEM, R. F., ABU-MWEIS, S. S. & BAWADI, H. A. 2013. Low Self-Efficacy Lead to Failure in Adopting CHO-Counting Approach to Control Diabetes: Pilot Study. *Pakistan Journal of Nutrition*, 12, 651.
- TOL, A., BAGHBANIAN, A., MOHEBBI, B., SHOJAEIZADEH, D., AZAM, K., SHAHMIRZADI, S. E. & ASFIA, A. 2013. Empowerment assessment and influential factors among patients with type 2 diabetes. *Journal of Diabetes & Metabolic Disorders*, 12, 6.
- TRIPLEPI, G., JAGER, K. J., DEKKER, F. W. & ZOCCALI, C. 2010. Selection bias and information bias in clinical research. *Nephron Clinical Practice*, 115, c94-c99.
- TUOMILEHTO, J. 2013. The emerging global epidemic of type 1 diabetes. *Current diabetes reports*, 13, 795-804.
- VAN DER BERG, J. D., STEHOUWER, C. D., BOSMA, H., VAN DER VELDE, J. H., WILLEMS, P. J., SAVELBERG, H. H., SCHRAM, M. T., SEP, S. J., VAN DER KALLEN, C. J. & HENRY, R. M. 2016. Associations of total amount and patterns of sedentary behaviour with type 2 diabetes and the metabolic syndrome: The Maastricht Study. *Diabetologia*, 59, 709-718.
- VLAHOV, D., FREUDENBERG, N., PROIETTI, F., OMPAD, D., QUINN, A., NANDI, V. & GALEA, S. 2007. Urban as a determinant of health. *Journal of Urban Health*, 84, 16-26.
- WALKER, R. J., SMALLS, B. L., CAMPBELL, J. A., WILLIAMS, J. L. S. & EGEDE, L. E. 2014. Impact of social determinants of health on outcomes for type 2 diabetes: a systematic review. *Endocrine*, 47, 29-48.
- WALSTON, S., AL-HARBI, Y. & AL-OMAR, B. 2008. The changing face of healthcare in Saudi Arabia. *Annals of Saudi medicine*, 28, 243.
- WALT, G. & GILSON, L. 1994. Reforming the health sector in developing countries: the central role of policy analysis. *Health policy and planning*, 9, 353-370.

- WÄNDELL, P. E. & GÅFVELS, C. 2004. Patients with type 2 diabetes aged 35–64 years at four primary health care centres in Stockholm County, Sweden: prevalence and complications in relation to gender and socio-economic status. *Diabetes research and clinical practice*, 63, 195-203.
- WANG, H., LIU, T., QIU, Q., KARP, E., DING, P., HE, Y.-H. & CHEN, W.-Q. 2015. Development and validation of a simple risk score for prevalent undiagnosed type 2 diabetes in Southern Chinese population. *International Journal of Diabetes in Developing Countries*, 35, 318-326.
- WARDIAN, J. & SUN, F. 2014. Factors associated with diabetes-related distress: implications for diabetes self-management. *Social work in health care*, 53, 364-381.
- WARSY, A. & EL HAZMI, M. 1999. Diabetes mellitus, hypertension and obesity-common multifactorial disorders in Saudis.
- WELCH, V. L., OSTER, N. V., GAZMARARIAN, J. A., RASK, K., SCHILD, L., CUTLER, C., SPETTELL, C. & REARDON, M. 2006. Impact of a diabetes disease management program by race and ethnicity. *Disease Management & Health Outcomes*, 14, 245-252.
- WHITING, D. R., GUARIGUATA, L., WEIL, C. & SHAW, J. 2011. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Res Clin Pract*, 94, 311-21.
- WHO 2013. Country cooperation strategy for WHO and Saudi Arabia 2012-2016.
- WILD, S., ROGLIC, G., GREEN, A., SICREE, R. & KING, H. 2004a. Global prevalence of diabetes estimates for the year 2000 and projections for 2030. *Diabetes care*, 27, 1047-1053.
- WILD, S., ROGLIC, G., GREEN, A., SICREE, R. & KING, H. 2004b. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes care*, 27, 1047-1053.
- WILDER, R. P., MAJUMDAR, S. R., KLARENBACH, S. W. & JACOBS, P. 2005. Socio-economic status and undiagnosed diabetes. *Diabetes research and clinical practice*, 70, 26-30.
- WILLIAMS, R. 2016. Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, 40, 7-20.
- WINKLEBY, M. A. & CUBBIN, C. 2003. Influence of individual and neighbourhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States. *Journal of Epidemiology and Community Health*, 57, 444-452.
- WRAY, L. A., ALWIN, D. F., MCCAMMON, R. J., MANNING, T. & BEST, L. E. 2006. Social status, risky health behaviors, and diabetes in middle-aged and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 61, S290-S298.
- XIE, J., HU, D., YU, D., CHEN, C.-S., HE, J. & GU, D. 2010. A quick self-assessment tool to identify individuals at high risk of type 2 diabetes in the Chinese general population. *Journal of epidemiology and community health*, 64, 236-242.
- XIN, Z., YUAN, J., HUA, L., MA, Y.-H., ZHAO, L., LU, Y. & YANG, J.-K. 2010. A simple tool detected diabetes and prediabetes in rural Chinese. *Journal of clinical epidemiology*, 63, 1030-1035.
- ZHANG, X., BECKLES, G. L., BULLARD, K. M., GREGG, E. W., ALBRIGHT, A. L., BARKER, L., ZHANG, X., RUIZ-HOLGUÍN, R., CERQUEIRA, M. T. & FRONTINI, M. 2010. Access to health care and undiagnosed diabetes along the United States-Mexico border. *Revista Panamericana de Salud Pública*, 28, 182-189.

List of References

- ZHANG, X., GEISS, L. S., CHENG, Y. J., BECKLES, G. L., GREGG, E. W. & KAHN, H. S. 2008. The missed patient with diabetes: how access to health care affects the detection of diabetes. *Diabetes care*, 31, 1748-1753.
- ZHOU, Z., RAHME, E., ABRAHAMOWICZ, M. & PILOTE, L. 2005. Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. *American journal of epidemiology*, 162, 1016-1023.