Environmental and Genetic Contribution to the Risk of T2D in the United Arab Emirates

Budour Alkaf
(MRes)

2022

Department of Epidemiology and Biostatistics
School of Public Health
Faculty of Medicine
Imperial College London

PhD Thesis

Statement of Originality

I, Budour Alkaf, confirm that the work presented in this thesis is my original work. Information that has been derived from other sources, has been appropriately referenced.

Copyright Declaration

‘The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Noncommercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.’

Dedication

I lovingly dedicate this work to my precious family, my father Khalid, my mother Ebtisam, and my five siblings. Without their unconditional love, support, and the light they bring to my life, this work would have been insurmountable. The long hours and late nights were only possible with the knowledge on the back of my mind that all other aspects of life were taken care of by my parents. Their upbringing instilled an ethic for hard work and nurtured a drive within me to strive for excellence and not perfection. There are no words to express my gratitude for both of them.

Acknowledgements

I am deeply grateful for Professor Järvelin, Professor Blakemore and Dr Nader Lessan for preserving with me for the time it took to complete this research. I cannot thank them enough for their invaluable support and constructive advice and their belief in my ability throughout the last four years. A large proportion of this research has been conducted using data from the UK Biobank, a major biomedical database, which was immensely useful to achieve my objectives. Of course, I am thankful for Professor Bell and Professor Thomas for affording me the opportunity to access data from the UK Biobank database. My thanks and appreciation go to Dr Wielscher for his support, mentorship, and guidance throughout my GWAS analyses. I learned a lot by working with him. No words are enough to thank him. I also would like to acknowledge all members of Imperial College London Diabetes Centre (ICLDC, Abu Dhabi), including Sameera Ahmed, Rayan Khaled, and all nurses for their contribution in participant recruitment and data collection.

I would like to deeply thank Mubadala for funding this study and affording me the opportunity to contribute to research for the United Arab Emirates. This work was partly supported by the Medical Research Council, UK, through MRC Centre for Environment and Health (MR/S019669/1, 2019-2024) and through the MRC/BBSRC PREcisE grant (MR/S03658X/1) (JPI HDHL H2020) for Professor Järvelin.

Abstract

Background

Type 2 diabetes (T2D) has emerged as a major healthcare problem at a global level. Populations in the Arabian Peninsula, including Arab-Emiratis, are repeatedly reported to have some of the highest rates of T2D. I aimed to investigate the contribution of lifestyle and genetic risk factors to this increased burden, that is yet to be determined.

Methods

I analysed publicly-available worldwide estimates to assess the contribution of obesity and physical inactivity towards differences in overall diabetes rates across populations. I investigated two Arab-Emirati sample populations to examine the contribution of adiposity and physical inactivity towards T2D risk, as compared to individuals of Caucasian, South Asian and Afro-Caribbean descent from the UK Biobank study. I performed a genome-wide association study in 1,200 Arab-Emiratis to identify known and novel genetic factors underlying T2D in the population.

Results

Populations experiencing similar environmental shifts are differentially subject to T2D risk. Arab-Emiratis have a 13-fold excess risk of T2D compared to Europeans, independent of BMI. There is a stronger association of central adiposity towards T2D risk in Arab-Emiratis compared to other ethnic groups. Sedentary time explains a bigger proportion of T2D risk compared to overall physical inactivity in Arab-Emiratis. I identified 40 novel loci in association with T2D (*P*-value < $5x10^{-6}$), with one loci reaching

genome-wide significance mapping to *MALRD1* (*P*-value = $1.6x10^{-8}$). Overall, results showed that previously reported T2D risk loci cannot fully account for the increased risk in Arab-Emiratis.

Conclusions

Increased T2D risk in Arab-Emiratis could be more attributable to central adiposity rather than BMI. Findings highlighted the negative impact of physical inactivity on insulin resistance and central adiposity in Arab-Emiratis, but its contribution towards their increased T2D risk is limited. Lastly, I identified 40 novel loci for T2D that potentially provide further insight into the genetic mechanisms underlying T2D in Arab-Emiratis.

Table of Contents

Abstract 6

List of Figures 15

List of Tables 18

Chapter One: Introduction 22

1. INTRODUCTION 22

1.1 Definition of T2D and prediabetes 23

1.1.1 Diabetes classification and T2D cluster 23

1.1.2 Definition of prediabetes 24

1.2 Global burden of diabetes 25

1.3 Pathogenesis of T2D 26

1.3.1 Insulin resistance 26

1.3.2 Pancreatic β-cell dysfunction 29

1.4 T2D is a disease with a heterogeneous spectrum of metabolic disturbance 30

1.4.1 Obesity 33

1.4.2 Physical activity 38

1.4.3 Genetics of T2D 44

1.5 Ethnic predisposition to T2D 55

1.5.1 Variation in T2D rates across ethnic groups 55

1.5.2 Variation in adiposity and T2D susceptibility 58

1.5.3 Variation in insulin resistance and action in T2D susceptibility 60

1.5.4 Ethnicity-specific genetic predisposition to T2D 60

1.5.5 Genetically-isolated populations 61

1.6 Understanding the excess risk of T2D in the United Arab Emirates 62

1.6.1 Demographics: The Arabian Peninsula and the UAE 62

1.6.2 Population substructure of the UAE 64

1.6.3 Global burden of T2D in the UAE 66

1.6.4 Factors contributing to high T2D risk in the UAE: what has been investigated? 67

1.6.5 Obesity and T2D risk in the UAE 67

1.6.6 Physical activity and T2D risk in the UAE 68

1.6.7 Genetic risk factors and T2D in the UAE 69

1.7 Research problem, hypothesis, and objectives 70

1.7.1 Research problem and scope 70

1.7.2 Study primary aim and hypothesis 71

1.7.3 Objectives 71

Chapter Two: Materials and Methods 77

2.1 Objective 1: Use of publicly-available estimates to unravel the contribution of obesity and insufficient physical activity towards diabetes risk across different populations 79

2.1.1 Study design 79

2.1.2 Data retrieval: publicly-available databases for estimates of diabetes, obesity and insufficient physical activity rates 79

2.1.3 Description of the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) data sources 81

2.1.4 Description of the World Health Organisation (WHO) data sources 83

2.1.5 Statistical analyses and approaches 83

2.2 Objective 2: Retrieval and processing of the UK Biobank baseline cross-sectional data to perform trans-ethnic comparisons. 85

2.2.1 Study design 85

2.2.2 UK Biobank Cohort: background 85

2.2.3 Diagnoses and definitions 87

2.2.4 Screening 89

2.3 Objective 3: Establish a cross-sectional sample of Arab-Emiratis that is representative of the general population to investigate the contribution of obesity towards risk of T2D in the population (EM1) 91

2.3.1 Study design 91

2.3.2 Participants 91

2.3.3 Screening 93

2.3.4 Diagnoses and definitions 95

2.4 Objective 4: Conduct a case-control study to collect a sample of Arab-Emirati men and women (EM2) with or without T2D to investigate the contribution of environmental and genetic risk factors for T2D in the UAE 96

2.4.1 Study design 96

2.4.2 Participants 96

2.4.3 Screening 98

2.4.4 Physical activity 99

2.4.5 Diagnosis and definitions 102

2.5 Objective 5: Conducting a genome-wide association analysis to identify novel and known genetic loci influencing, and contributing to the increased risk of T2D in the UAE 103

2.5.1 Study design and participants 103

2.5.2 Extraction of Genomic DNA 103

2.5.3 Genome-Wide SNP Genotyping using the Infinium Global Screening Array -24 v3.0 BeadChip 104

2.5.4 Genotype calling in GenomeStudio 105

2.5.5 Preparing PLINK files 106

2.5.6 GWAS quality control 106

2.5.7 Genotype imputation 110

2.6 Statistical Methods and Approaches 111

Chapter Three: Secondary analyses of global datasets: do obesity and physical activity explain variation in diabetes risk across populations? 119

3.1 OVERVIEW 119

3.2 AIMS 120

3.2.1 Objectives 120

3.3 INTRODUCTION 121

3.4 MATERIALS AND METHODS 122

3.4.1 Data sources, data retrieval and processing 122

3.4.2 Definitions 124

3.4.2 Statistical methods 124

3.5 RESULTS 125

3.5.1 Region- and sex-specific trends in diabetes rates from 1980 to 2014 125

3.5.2 Region- and sex-specific trends in obesity rates from 1975 to 2016 131

3.5.3 Region- and sex-specific estimates of insufficient physical activity rates in 2016 135

3.5.4 Contribution of obesity towards differences in diabetes risk across populations 135

3.5.5 Contribution of insufficient physical activity towards explaining differences in diabetes risk across populations 136

3.5.6 Joint effects of obesity and physical inactivity levels with diabetes 140

3.5.7 The UAE versus the UK: contribution of obesity and insufficient physical activity towards diabetes risk 144

3.6 DISCUSSION 146

Chapter Four: The contribution of adiposity towards increased risk of T2D among Arab-Emiratis 154

4.1 OVERVIEW 154

4.2 AIMS 155

4.2.1 Objectives 155

4.3 INTRODUCTION 156

4.4 MATERIALS AND METHODS 159

4.4.1 Study design and participant characterisation 159

4.4.2 Statistical analyses 162

4.5 RESULTS 164

4.5.1 EM1 sample representativeness of the UAE population 164

4.5.3 Basic characteristics of the EM2 sample population 166

4.5.4 Comparison of T2D prevalence rates between Arab-Emiratis (EM1) and other ethnic groups 167

4.5.5 Insulin, insulin resistance, and β-cell function in the EM2 sample population 168

4.5.6 Adiposity and glycaemic measures in EM2 sample population 168

4.5.7 Contribution of adiposity to increased T2D risk and glycaemic disturbances in Arab-Emiratis compared to other ethnic groups 172

4.5.8 Sensitivity analyses 185

4.6 DISCUSSION 185

Chapter Five: The contribution of physical inactivity and sedentary time towards T2D risk among Arab-Emiratis 220

5.1 OVERVIEW 220

5.2 AIMS 221

5.2.1 Objectives 221

5.3 INTRODUCTION 222

5.4 MATERIALS AND METHODS 224

5.4.1 Study design and participant characterisation 224

5.4.2 Physical activity 225

5.4.3 Diagnosis and definitions 227

5.4.4 Power Calculations 228

5.4.5 Statistical Analysis 228

5.5 RESULTS 230

5.5.1 The contribution of physical inactivity towards risk of T2D, glycaemic disturbances, lipid, and adiposity measures in Arab-Emiratis 230

5.5.2 Cross-ancestry comparisons 238

5.6 DISCUSSION 242

Chapter Six: Identifying Genetic Effects Underlying T2D in Arab-Emiratis 289

6.1 OVERVIEW 289

6.2 AIMS 290

6.2.1 Objectives 290

6.3 INTRODUCTION 294

6.4 MATERIALS AND METHODS 295

6.4.1 Study participants 295

6.4.2 Genotyping and genotype calling 297

6.4.3 Quality control 297

6.4.4 Imputation 297

6.4.5 Regression analyses 299

6.4.6 Conditional analyses 300

6.4.7 Regional plots and Arab-Emirati-specific LD matrix 301

6.4.8 Cross-ancestry comparisons 301

6.4.9 Power calculation 302

6.5 RESULTS 303

6.5.1 The Primary Arab-Emirati GWAS 303

6.5.2 Cross-ancestry comparisons 317

6.6 DISCUSSION 323

Chapter Seven: Discussion 361

7.1 The role of variation in adiposity in type 2 diabetes susceptibility 361

7.1.1 Future work and implications for public health 367

7.2 The relationship of physical inactivity and sedentary time with T2D burden in the UAE 368

7.2.1 Future work and implications for public health 372

7.3 The environment and T2D risk 373

7.4 Genetic risk factors and T2D 374

7.4.1 Future work and implications for public health 380

7.5 Strengths and limitations 381

7.6 Conclusions 383

Bibliography 385

List of Figures

Figure 1.1 29
Figure 1.2 31
Figure 1.3 32
Figure 1.4 56
Figure 1.5 57
Figure 1.6 58
Figure 1.7 63
Figure 1.8 65
Figure 2.1 78
Figure 2.2 80
Figure 2.3 86
Figure 2.4 107
Figure 3.1 123
Figure 3.2 127
Figure 3.3 128
Figure 3.4 130
Figure 3.5 131
Figure 3.6 133
Figure 3.7 134
Figure 3.8 137
Figure 3.9 138
Figure 3.10 139
Figure 3.11 141
Figure 3.12 143
Figure 3.13 145
Figure 4.1 158
Figure 4.2 161

Figure 4.3 166
Figure 4.4 170
Figure 4.5 171
Figure 4.6 173
Figure 4.7 174
Figure 4.8 175
Figure 4.9 177
Figure 4.10 180
Figure 4.11 181
Figure 4.12 183
Figure 4.13 184
Figure 5.1 232
Figure 5.2 233
Figure 5.3 234
Figure 5.4 239
Figure 6.1 292
Figure 6.2 293
Figure 6.3 298
Figure 6.4 305
Figure 6.5 306
Figure 6.6 308
Figure 6.7 309
Figure 6.8 311
Figure 6.9 312
Figure 6.10 313
Figure 6.11 314
Figure 6.12 315
Figure 6.13 316
Figure 6.14 319
Figure 6.15 322
Figure 7.1 363
Figure 7.2 370

Figure 7.3 377

List of Tables

Table 1.1 ..... .273
Table 1.2 73
Table 1.3 74
Table 1.4 75
Table 1.5 76
Table 2.1 113
Table 2.2 114
Table 2.3 115
Table 2.4 117
Table 2.5 118
Table 3.1 151
Table 3.2 152
Table 4.1 191
Table 4.2 192
Table 4.3 195
Table 4.4 198
Table 4.5 201
Table 4.6 203
Table 4.7 205
Table 4.8 207
Table 4.9 211
Table 4.10 214
Table 4.11 217
Table 5.1 248
Table 5.2 251
Table 5.3 254
Table 5.4 260
Table 5.5 264

Table 5.6 269
Table 5.7 272
Table 5.8 274
Table 5.9 276
Table 5.10 279
Table 5.11 280
Table 5.12 282
Table 5.13 285
Table 5.14 287
Table 6.1 329
Table 6.2 332
Table 6.3 336
Table 6.4 338
Table 6.5 340
Table 6.6 344
Table 6.7 350
Table 6.8 355
Table 6.9 358

List of Abbreviations

AC	Afro-Caribbean
ADA	American Diabetes Association
BMI	Body mass index
CAU	Caucasian
CI	Confidence interval
FBG	Fasting blood glucose
GPAQ	Global Physical Activity Questionnaire
GWAS	Genome-wide association study
HbA1c	Glycohemoglobin
HDL	High-density lipoprotein
HOMA	Homeostasis model of assessment
HPA	High physical activity
HRC	Haplotype Reference Panel
ICLDC	Imperial College London Diabetes Centre
IDF	International diabetes federation
IPA	Insufficient physical activity
IPAQ	International Physical Activity Questionnaire
IQR	Interquartile range
LDL	Low-density lipoprotein
LPA	Low physical activity
MET	Metabolic equivalent of task
MPA	Moderate physical activity
NCD-RisC	Non-Communicable Disease Risk factor Collaboration
NGT	Normal Glucose Tolerance
OR	Odds ratios
PAR	Population attributable risk
SA	South Asian
SD	Standard deviation
SNP	Single nucleotide polymorphism
T2D	Type 2 diabetes
TC	Total cholesterol
TPA	Total physical activity
UAE	United Arab Emirates
UKBB	UK Biobank
VAT	Visceral adipose tissue
WC	Waist circumference

WHO	World Health Organisation
WHR	Waist-to-hip ratio

Chapter One: Introduction

This thesis describes an investigation of environmental and genetic risk factors predisposing Arab-Emiratis, the population of the United Arab Emirates, to type 2 diabetes. Through this chapter, I outline and discuss the literature that forms the backbone of my PhD research project, highlight the research gap, and justify my hypothesis, aims and objectives. In short, the overall aim of this thesis is to investigate the contribution of environmental and genetic risk factors to the risk of type 2 diabetes in Arab-Emiratis, compared to other ethnic populations, including Caucasians, South Asians, and Afro-Caribbeans living in the UK (further details can be found in section 1.7.2).

1. INTRODUCTION

Diabetes is one of the largest global public health concerns, and one of the top 10 causes of death globally[1]. Type 2 diabetes (T2D), which accounts for the majority (75-85%) of all diabetes cases, is a metabolic disorder characterised by chronic hyperglycaemia and disturbances in fat, carbohydrate and protein metabolism. These abnormalities are the consequences of inadequate insulin secretion or impaired insulin action, or both[2]. T2D is strongly associated with reduced quality of life, reduced life expectancy, and increased risk of myocardial infarction, stroke, and microvascular events, and mortality.

T2D is a heterogeneous disorder with varying prevalence among different populations. Compared to Europeans, populations in the Middle East, and the Arabian Gulf region in particular, which has over 2-fold higher prevalence rates of T2D[3]. Our knowledge of the contribution of known lifestyle and genetic risk factors to the excess risk of T2D in Middle Eastern populations is very limited.

1.1 Definition of T2D and prediabetes

1.1.1 Diabetes classification and T2D cluster

T2D was first defined in 1980, when the World Health Organisation (WHO) classified diabetes into 2 main distinct types: type-1 diabetes (T1D, Insulin dependent diabetes mellitus [IDDM]) and type-2 diabetes (T2D, non-insulin dependent diabetes mellitus [NIDDM]). The classification of diabetes into those two types primarily relied on the presence (T1D) or absence (T2D) of autoantibodies against pancreatic β-cell antigens and age at diagnosis. With this approach, 75-85% of patients with diabetes are classified as having T2D. T1D is characterised by absolute insulin deficiency as a result of β-cell destruction. Neither its cause nor any means to prevent it are well known. In contrast, T2D results from a progressive insulin secretory defect on a background of insulin resistance. This type of diabetes is very strongly associated with excess body weight and physical inactivity. Symptoms of T2D may be similar to those of T1D, but are often less marked and therefore, in many cases, the disease is diagnosed several years after onset, when complications have already arisen. T2D has historically only been diagnosed in adults, but recently, it is increasingly prevalent in children. An additional (third) sub-group, latent autoimmune diabetes in adults (LADA), affects less than 10% of people with diabetes, and is characterised by the presence of glutamic acid decarboxylase antibodies (GADA). LADA is phenotypically very similar to T2D at diagnosis, but as it progresses, it becomes increasingly similar to T1D over time[4]. Other recognised diabetes classes include gestational diabetes, diabetes due to specific genetic defects in β-cell function, genetic defects in insulin action, diseases of the exocrine pancreas (cystic fibrosis), and drug/chemical induced diabetes (such as in the treatment of AIDS or after organ transplantation).

More recently, the emergence of a refined T2D classification system marked an important milestone[5, 6]. Based on the All New Diabetes in Scania (ANDIS) cohort, Ahlqvist et al. used six clinical-biological

variables, including age at diagnosis, body mass index, haemoglobin A1C, GADA, fasting glucose, and C-peptide levels-based estimates of β-cell function and insulin resistance, to categorise adults with newly diagnosed diabetes into five mutually exclusive clusters[5]. These clusters were distinguished by different underlying pathophysiology and included: severe auto-immune diabetes (SAID), severe insulin-deficient diabetes (SIDD), severe insulin-resistant diabetes (SIRD), mild obesity-related diabetes (MOD), or mild age-related diabetes (MARD). In addition to exhibiting distinct baseline phenotypic characteristics, individuals from different clusters have different diabetic complication trajectories. The distribution of these clusters, and the baseline patient characteristics were first identified in Swedish individuals (ANDIS cohort)[5], but their application was tested in different populations, including China[7], USA[7], and in South Asian[8] and East Asian[9] populations.

Diabetes is traditionally diagnosed via biochemical tests including measurement of fasting blood (plasma) glucose (FBG) levels, an oral glucose tolerance test (OGTT), and/or measurement of glycohemoglobin (HbA1c). WHO criteria for the diagnosis of T2D are described in **Table 1.1**.

1.1.2 Definition of prediabetes

Prediabetes is an intermediate state of hyperglycaemia with glycaemic parameters above the normal range, but below that of clinical T2D[10]. It can be either identified as impaired fasting glucose (IFG) or impaired glucose tolerance (IGT). The diagnosis criteria for prediabetes are not uniform; the WHO defines IFG as fasting plasma glucose levels between 6.1 and 6.9 mmol/l and IGT as 2 h plasma glucose of 7.8-11.0 mmol/l after ingestion of 75g of oral glucose load or the combination of the two based on a 2 h oral glucose tolerance test (OGTT)[11]. On the other hand, the American Diabetes Association (ADA), has

a lower cut-off for IFG, and has additional HbA1c-based criteria, of a level of 5.7% to 6.4%, for the definition of prediabetes[12] (**Table 1.2**).

1.2 Global burden of diabetes

Diabetes is a leading cause of mortality and reduced life expectancy, imposing a heavy global burden on public health and socio-economic development across nations[1]. In 2021, the International Diabetes Federation (IDF) estimated that 537 million adults (20-79 years) are living with diabetes worldwide, with a projected increase to 643 million by 2030 and 784 million by 2045 if no effective prevention measures are adopted[3]. It is estimated that an additional 541 million adults have IGT, which places them at an increased risk of developing T2D. Moreover, the prevalence of both T1D and T2D has increased in children and adolescents below 20 years old, with over one million living with T1D at that age group[13].

Diabetes is considered one of the top 10 causes of death on a global level. Diabetes, cardiovascular disease, cancer and respiratory disease collectively account for over 80% of all premature noncommunicable diseases (NCDs) deaths[14]. Individuals with diabetes have a 2-3-fold increased risk of all cause-mortality, after adjustment for BMI, age, sex and smoking status[15]. Diabetes is also associated with premature deaths both from NCDs[16] and from communicable diseases including those caused by coronaviruses [severe acute respiratory syndrome (SARS) and COVID-19][17, 18]. Despite efforts to promote population health and extend life expectancy, diabetes still has the second largest total effect on reducing global health adjusted life expectancy (HALE), which refers to the equivalent number of years in full health that an individual at a given age can expect to live if exposed to age specific mortality and health status (morbidity) conditions that prevail today[19].

In 2021, according to the IDF, diabetes is responsible for 6.7 million deaths (1 in every 5 seconds)[3]. They

have also estimated that diabetes costs at least 966 billion USD dollars in health expenditure at a global level - a 316% increase over the last 15 years (2006-2021)[3]. In addition, the Global Burden of Diseases, Injuries, and Risk Factors Study estimated that a high fasting level of glucose was the tenth most common global risk factor for disability-adjusted life years (DALYs) in 1990, fourth most common in 2005 and third most common in 2015, accounting for 143 million DALYs in 2015 and a 22% increase in DALYs from 2005 to 2015[20].

1.3 Pathogenesis of T2D

Diabetes mellitus (DM) is defined as a chronic metabolic disease, characterised by the inadequate control of blood glucose levels, which results over time to serious damage to the heart, blood vessels, eyes, nerves, and kidneys[21]. T2D accounts for about 75-85% of all DM cases. It develops in adults (older adults in particular) and is many times left unnoticed and undiagnosed for several years. Diagnosis is often incidental or at the onset of diabetic microvascular and/or macrovascular complications[22]. T2D is a progressive disease that is caused by a combination of two primary factors: defective insulin secretion by pancreatic β-cells and the inability of insulin-sensitive tissues to respond to insulin appropriately[23]. In healthy individuals, molecular mechanisms involved in the synthesis and release of insulin, as well as insulin response in tissues are tightly regulated to precisely meet metabolic needs. A malfunction of the feedback loops between insulin action and insulin secretion results in abnormally high glucose levels in blood[24].

1.3.1 Insulin resistance

Insulin resistance (IR) is a pathophysiological state in which insulin-responsive cells have an impaired/lowered response to circulating insulin, in order to maintain normal glucose levels[25]. Under fasting conditions, insulin response is alleviated by glucagon, glucocorticoids and catecholamines to

prevent insulin-induced hypoglycaemia. On the other hand, in a fed state, the action of insulin is influenced by the interplay of additional molecules including growth hormone (GH) and insulin-like growth factor 1 (IGF-1). IR or insulin-deficient conditions can be categorised into three broad categories: (1) diminished insulin secretion by β-cells; (2) impaired insulin receptors or signalling due to insulin antagonists, counter-regulatory hormones or non-hormonal bodies; and (3) impaired insulin response in target tissues[26]. The latter includes three main insulin-sensitive organs that play major roles in these processes: skeletal muscle, adipose tissue and liver. Defective insulin action in these tissues often precedes the development of systematic IR, progressively leading to T2D development[23]. Target tissues can have differential responses to insulin: skeletal muscle is more sensitive to insulin than the liver and adipose tissue[27]. Therefore, IR in skeletal muscle manifests during the early development of T2D and often prior to IR in the liver and adipose tissue.

1.3.1.1 Skeletal muscle IR

Insulin stimulates muscle glycogen synthesis by enhancing glucose uptake from plasma. In muscle cells, when insulin binds to an insulin receptor (INSR), glucose transporter GLUT4 translocates from intracellular compartments to the plasma membrane. This process allows glucose uptake, lowering circulating glucose levels[28]. Mutations that lead to reduction in *INSR* or *GLUT4* expression, as well as any obstruction in downstream or upstream signalling pathways would result in reduction of glucose intake into the muscle leading to a hyperglycaemic state[25, 29]. Besides the mutations and defective epigenetic regulation, environmental factors can also play an important role in glucose uptake by muscle tissues. Physical activity has been shown to enhance glucose utilisation by muscle tissues through increasing blood flow into skeletal muscle cells[30]. In addition, obesity, which is characterised as a state of chronic inflammation, contributes to IR and T2D. Obesity can lead to myocyte inflammation and impaired myocyte metabolism via increased immune cell infiltration and secretion of proinflammatory molecules.

The different mechanisms linking obesity and/or physical inactivity with T2D are described in sections 1.4.1.4 and 1.4.2.5.

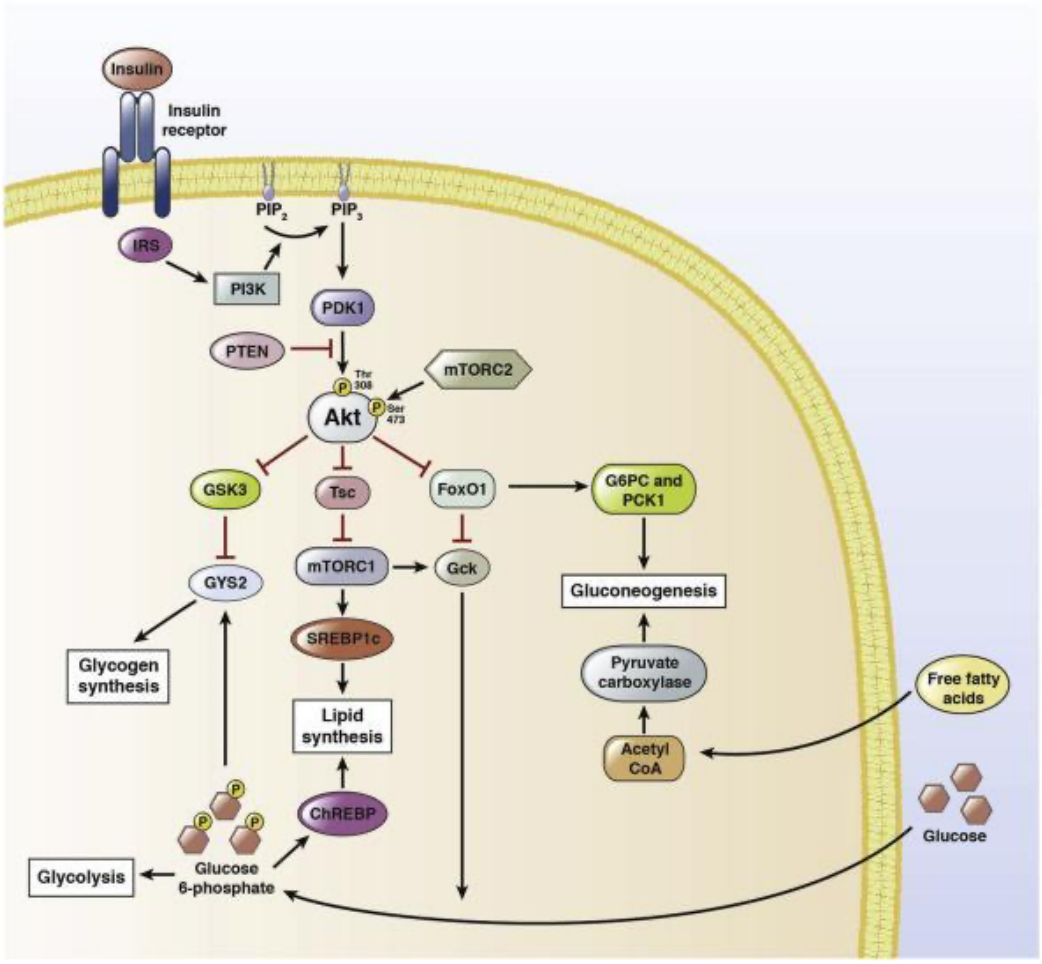
1.3.1.2 Adipose IR

Adipose tissue is an endocrine organ that influences both glucose and lipid metabolism[31, 32] by releasing adipokines, proinflammatory factors, and free fatty acids (FFAs), that impair glucose metabolism and muscle adenosine triphosphate (ATP) synthesis[33], subsequently promoting the synthesis of toxic lipid metabolites, and altering insulin signalling[34, 35]. In adipose tissues, insulin stimulates glucose uptake and triglyceride hydrolysis and release of FFA and glycerol into the circulation[36, 37]. Adipose-IR is defined as the impaired suppression of lipolysis in the presence of high insulin levels. It has been associated with glucose intolerance and elevated FFA levels, both which have been shown to impair muscle insulin signalling, promote hepatic gluconeogenesis, and impair glucose-stimulated insulin response[37-39].

1.3.1.3 Hepatic IR

In the liver, insulin both regulates glucose production/utilisation and affects lipid metabolism[23]. In normal physiological states, glucagon induces hepatic glucose production, whereas insulin acts as a potent inhibitor of glucose production when its blood concentration is elevated[40]. Insulin also inhibits hepatic glucose production, consequently reducing hepatic glucose release through inhibition of Forkhead box protein O1 (FOXO1) (**Figure 1.1**). In states of IR, circulating insulin levels are insufficient to prompt an appropriate insulin response in hepatic cells, resulting in impaired glycogen synthesis, failure to suppress glucose production, increased lipogenesis and proinflammatory protein synthesis[23].

Figure 1.1 Insulin signalling in hepatocytes when insulin binds and activates the insulin receptor after a meal. Adapted from Santoleri D (2018)[41].



1.3.2 Pancreatic β-cell dysfunction

Insulin resistance may initially be compensated by increased insulin secretion (hyperinsulinemia), and progression to T2D is thought to occur when β-cells fail to maintain glucose homeostasis to a sufficient degree, resulting in hyperglycemia[42, 43]. Defective insulin secretion is a product of reduced functional β-cell mass and intrinsic β-cell dysfunction. β-cell dysfunction has been traditionally associated with β-cell death[44]. More recent evidence suggests that it may underlie a more complex network of interactions

between the environment and different molecular pathways implicated in cell biology[45]. An excessive nutritional state is often characterised by hyperglycaemia and hyperlipidaemia, which collectively can lead to insulin resistance and chronic inflammation. Subsequently, in these conditions, β-cells are exposed to toxic pressures including inflammation, inflammatory stress, endoplasmic reticulum stress, metabolic/oxidative stress, and amyloid stress, together potentially leading to a loss of islet integrity/organisation[44].

Impaired optimal cell-to-cell communication within pancreatic islets contributes to poor regulation of insulin and glucagon release[46]. Defects in the synthesis and/or secretion of insulin or any of its precursors, can lead to insulin secretory dysfunction, the primary driver of β-cell failure and a foundation of T2D pathogenesis[46]. It has been shown that pancreatic β-cell dysfunction precedes the onset of hyperglycaemia by approximately 10 years and that β-cell function decreases progressively over time[47].

The relative contribution of insulin resistance and β-cell dysfunction on T2D development varies from one population to another[48, 49], and that is a phenomenon that will be discussed further in the next sections.

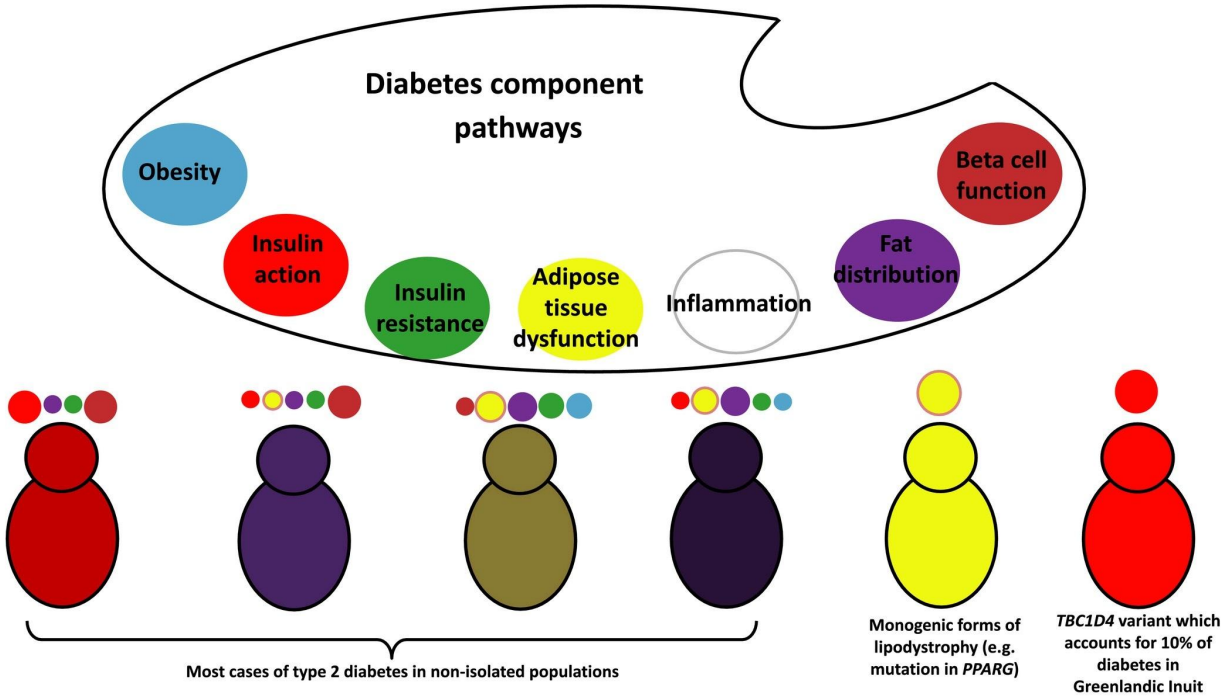
1.4 T2D is a disease with a heterogeneous spectrum of metabolic disturbance

Insulin resistance and insulin secretory failure are two well-established characteristics of T2D (as discussed above), however, T2D is a multisystem disease which involves abnormalities in the pancreas, fat, muscle, liver, and the brain. On the opposing end is the case of rare diseases, where a causal event is uncommon, precise and has high phenotypic impact that usually results in a qualitatively distinct extreme phenotype that can be easily recognised. Whereas T2D features multiple, common, low-effect

genetic risk variants, ubiquitous environmental exposures, and a phenotype that lies on a quantitative spectrum of metabolic disturbances[50].

Individuals seem to develop T2D due to a combination of defects in multiple pathways although to a different degree[50]. According to the palette model of diabetes[50] (**Figure 1.2**), different 'base colours' represent pathophysiological pathways, which in combination albeit different degrees of contribution, drive individuals from health to diabetes. The pathways include β-cell development, size, and function, islet autoimmunity incretin activity, obesity, fat distribution, adipose tissue dysfunction, insulin resistance, and other unknown mechanisms[50].

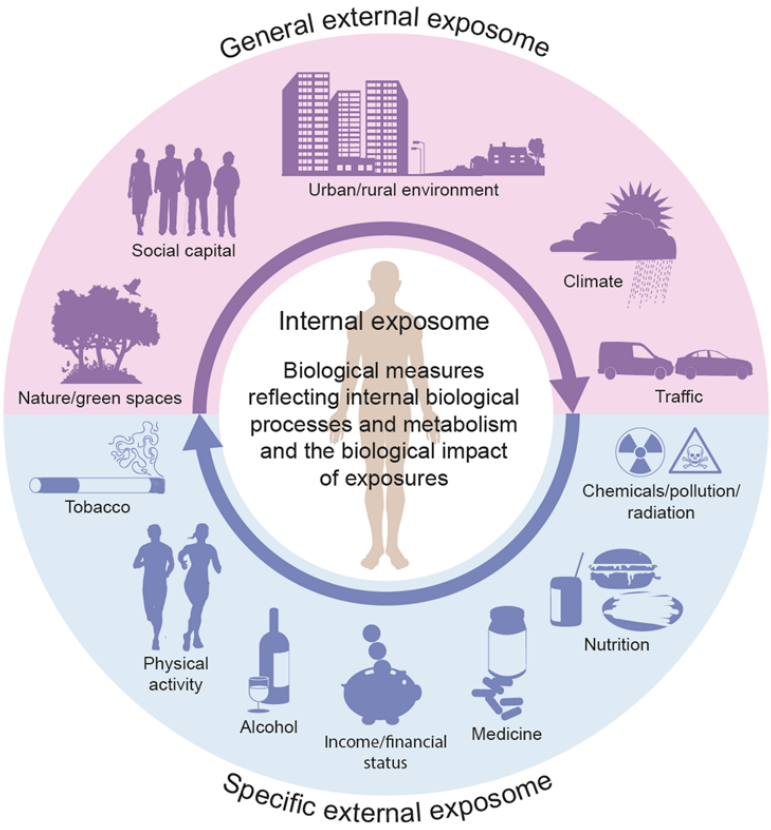
Figure 1.2 The McCarthy 'palette' model of type 2 diabetes. In this model the colour of an individual is painted based on the contribution of the 'base colours' representing the various pathophysiological processes underlying type 2 diabetes. Adapted from McCarthy MI. (2017)[50].



In the next part of this chapter, I considered in turn the role of obesity, physical activity, and genetic

susceptibility to T2D risk. Although my thesis is centred around three primary risk factors of T2D, with general obesity considered as the single most compelling of the traditional risk factors, it is important to note that studying these factors alone cannot fully evaluate the influence of the environment as whole (individual and general environment), as well as gene-environmental, and epigenetic influences on disease risk (**Figure 1.3**). These are considered and discussed in chapter 7, after I present to you the results of my thesis.

Figure 1.3 An overview of the three main domains of an individuals' exposome. These include (1) external influences of the environment (general external), (2) individual's exposure (specific external), and (3) the impact of these exposures on an individual's biological processes and metabolism (internal). Adapted from Beulens J (2021)[51].



1.4.1 Obesity

1.4.1.1 Global burden of obesity

Obesity is a public health concern that arises, at an individual level, as a consequence of how the body regulates energy intake, energy expenditure, and energy storage. A variety of factors are known to be involved in its development, including genetic, behavioural, cultural, environmental, and economic[52]. In some cases, obesity is regarded as a risk factor, and in others as a disease itself. The prevalence of obesity has drastically increased in the last few decades across all nations, both industrialised and developing ones. Worldwide obesity rates have nearly tripled between 1975 and 2016[53]. According to the WHO, in 2016, more than 1.9 billion adults, 18 years and older, were overweight. Of these, over 650 million were obese[53]. Globally, obesity has led to at least 2.8 million deaths per year[53]. In general, obesity is associated with a greater risk of disability or premature death due to T2D and cardiovascular diseases including hypertension, stroke and coronary heart disease, as well as gallbladder disease and certain cancers[54]. It has also been related to obstructive sleep apnea[55], gastro-oesophageal disease[56], and stress incontinence[57], all of which severely impact quality of life. Based on a large body of evidence, the global epidemic of obesity is largely a consequence of societal factors that promote sedentary lifestyles and the consumption of high-fat and energy-dense diets[54].

1.4.1.2 Definitions and measures of obesity and body composition

By definition, obesity is a condition of abnormal or excessive fat accumulation in adipose tissue that presents as a health risk[54]. The amount of excess fat in absolute terms and its distribution in the body - either around the waist and trunk or peripherally around the body - have differential and important health implications[54]. The specific mechanisms that underlie inter-individual differences in body fat distribution are complex and not fully understood, but involve sex hormones[58], use of glucocorticoids[59], genetic and epigenetic mechanisms[60, 61]. T2D risk conferred by differential body fat distribution (i.e.,

visceral, ectopic, subcutaneous) will be discussed later in this chapter (section 1.4.1.4).

Obesity is primarily diagnosed based on measures of weight and height to calculate body mass index (BMI). For adults, a BMI of 18.5 Kg/m^2 to 24.9 Kg/m^2 is considered healthy. Overweight is defined as a BMI between 25.0 Kg/m^2 and 29.9 Kg/m^2; and obesity is defined as a BMI of 30 Kg/m^2 or higher. Morbid obesity, also known as extreme, or severe obesity is defined as a BMI of 40 Kg/m^2 or higher[53].

Although BMI is a simple non-invasive method of assessing obesity and excessive fat stores within a single population, it is less reliable when it comes to assessing body composition[62]. BMI-based classifications cannot distinguish fatness from muscularity, and do not take fat distribution into account, i.e., where and how much of the fat is deposited in different sites of the body. For example, BMI does not accurately reflect visceral fat accumulation, the likely cause leading to most of the metabolic and clinical consequences of obesity[63, 64]. Also, BMI does not account for sex differences in fat distribution or age-related decline in muscle mass[63]. In addition, some individuals despite being obese, are free from the metabolic complications of obesity (metabolically-healthy obese)[65], most likely due to storing less visceral fat and preserved insulin sensitivity[65, 66].

A simple and non-invasive method to assess central adiposity includes waist and hip circumferences and waist-to-hip ratio (WHR). Central adiposity as measured by WHR is defined using gender specific cut-offs; men WHR > 0.9 and women WHR > 0.8. These measures provide useful information[67], particularly given that central obesity is a good surrogate of visceral fat accumulation and metabolic risk syndrome[68]. However, WHR is not a specific indicator of abdominal visceral fat accumulation[63, 69], and gives little information on other fat depots, with no information concerning ectopic fat levels and distribution[69-71]. Given that WHR, like BMI, is a ratio metric, waist circumference alone is considered to

be a better predictor of metabolic and T2D risk[63]. The main drawback of waist measurements seems to be its lack of ability to differentiate subcutaneous from visceral fat deposition[72].

Other relatively simple methods of measuring body fat include skin thickness and bioelectrical impedance analysis[73] ; more sensitive but costly measures include computed tomography, dual-energy X-ray absorptiometry, and magnetic resonance imaging (MRI) [74]. Overall, body composition varies with age, sex, and ethnicity, however, to date, there are insufficient normative sex- and age-specific data that would define obesity.

1.4.1.3 Obesity and T2D risk

Obesity is considered a major risk factor for insulin resistance and T2D[75]. Obesity and T2D frequently occur together, and statistics show that 60–90% of all patients with T2D are, or have been, obese[24, 76]. The relative risks for T2D with rising obesity are among the highest for any risk factor for any disease, with an approximate 50- to 80-fold increase in T2D risk for a BMI of over 35 Kg/m^2 compared with a BMI of less than 23 Kg/m^2 [77, 78]. Given this strong link, any simple measure of adiposity alone (for example BMI, waist circumference or WHR) can predict incidence of T2D[77, 78].

The positive association between obesity and risk of developing T2D has been observed both in cross-sectional studies and in prospective data sets[79-81]. The measure of obesity in nearly all studies has traditionally been BMI. Two of the largest prospective follow-up studies were conducted in the United States, namely the Nurses' Health Study (NHS) in women[81, 82] and Health Professionals Follow-up Study (HPFS) in men [64]. These studies have found obesity to be the single most important predictor of T2D, compared to other risk factors, including physical activity, poor diet, or smoking.

In the NHS, after adjustment for age, family history of T2D, smoking, exercise, and several dietary factors, the relative risk of T2D for the 90th percentile versus the 10th percentile of BMI was 11.2[83]. Similarly, in men (HPFS), a strong association between BMI and risk of T2D was demonstrated[64]. Following adjustment for smoking, family history and age, men with a BMI ≥ 35 Kg/m^2 had a relative risk of 42.1 for T2D, compared to men with a BMI < 23 Kg/m^2.

Interventional studies thereafter demonstrated the negative effect of weight gain on risk of T2D and the positive outcomes from weight loss-based lifestyle intervention as a preventative measure of obesity and T2D [84-86]. Compared to individuals with stable weight, individuals who gained weight had an increased risk of developing T2D, and this risk continued to rise with larger weight increases. On the other hand, weight loss, even of modest amounts, was shown to substantially reduce T2D risk [84-86]. In addition, the Finnish Diabetes Prevention Study (DPS)[87] and the US Diabetes Prevention Program (DPP)[88] independently reported a reduction in T2D incidence of 58%, accompanied with a significant reduction in body weight by 5%-7% as a result of lifestyle intervention. It is important to note, however, that although substantial weight loss is possible, long-term sustenance of lost weight is challenging[88]. A meta-analysis that included 29 long-term weight loss studies demonstrated that 50% of weight lost was regained within two years and 80% was regained within five years[89].

1.4.1.4 Mechanisms linking obesity and T2D

Links between obesity and T2D involve multiple possible mechanisms driven by increased production of adipokines/cytokines, including tumour necrosis factor-α, resistin, and retinol-binding protein-4, that contribute to insulin resistance as well as reduced levels of adiponectin[90], deranged fatty acid metabolism, and cellular processes such as mitochondrial dysfunction and endoplasmic reticulum stress[91]. These interactions are complex, with the relative importance of each not yet clearly defined.

The influence of obesity on T2D risk is not only determined by the degree of obesity but also by the differential allocation of accumulating fat. There are two main adipose tissue depots in the human body; subcutaneous adipose tissue (abdominal and peripheral) and internal adipose tissue (visceral and non-abdominal)[92]. Subcutaneous white adipose tissue (SAT) has a different distribution in men as compared to women. Men are characterised with increased accumulation in the trunk compared to limbs. While women accumulate SAT in abdomen and limbs[93]. Visceral adipose tissue (VAT), which is located between the walls of the abdominal cavity, is generally more prevalent in men than in women[93].

Visceral fat accumulation, as reflected by increased abdominal girth or WHR, is associated with metabolic syndrome, T2D, and cardiovascular disease[94, 95]. Studies that have quantified the amount of fat in subcutaneous and visceral adipose depots with computer tomography (CT) or MRI have demonstrated the larger impact of centrally-located visceral fat on insulin resistance, in comparison to centrally-located subcutaneous fat[96]. In addition, visceral fat, independent of degree of obesity and amount of subcutaneous abdominal fat, has been shown to predict levels of fasting insulin, as well as insulin and glucose responses following oral glucose loads[97].

More recently, another major hypothesis linking obesity with T2D has emerged, indicating the effect of ectopic fat (fat stored in tissues that are supposed to contain only small amounts of fat, such as the liver, skeletal muscle, heart, and pancreas) on insulin resistance via a process termed 'lipotoxicity'[98]. Interestingly, visceral and ectopic fat accumulation may be a downstream consequence of an 'impaired' subcutaneous fat storage capacity[78], for example people with lipodystrophy have an impaired ability to store subcutaneous fat, and consequently with even modest weight gain, they accumulate fat in visceral and ectopic tissues, resulting into marked insulin resistance[99]. At the other end of the spectrum are

individuals, women in particular, who despite having high BMIs (as high as 50-60 Kg/m^2), remain insulin-sensitive, normoglycemic and normolipemic. These have been shown to have high subcutaneous fat content, and lower levels of visceral and ectopic fat[100].

Amongst many mechanisms underlying the aetiology of effective fat storage in subcutaneous adipose tissue is the ability to generate new adipocytes[101]. Post-meals, adipocytes expand in order to store circulating FFAs in the form of triglycerides. This expansion occurs by the enlargement of pre-existing adipocytes (hypertrophy), by the generation of new small adipocytes (hyperplasia), or both[102]. The expansion through hypertrophy or hyperplasia correlates with higher insulin, insulin resistance, and increased T2D risk[103, 104]. In addition, obese subjects with few large adipocytes are more glucose intolerant and hyperinsulinemic than those having the same degree of obesity (as measured by BMI), but many small adipocytes[103, 104].

How differences in adiposity measures and body fat distribution relate to T2D risk across ethnic groups is discussed in more detail in section 1.5.2.

1.4.2 Physical activity

Insufficient physical activity has been recognised as one of the leading risk factors for non-communicable diseases, contributing to approximately 3.2 million deaths per year[104]. In addition, increased physical activity levels have positive effects on mental health, quality of life, and can help maintain a healthy weight[105, 106]. According to the WHO, in 2016, more than a quarter of all adults aged 18 years or older were not getting enough physical activity, with a worldwide prevalence of physical inactivity of 27.5%[105]. This puts over 1.4 billion adults at risk of developing diseases related to inactivity. Given the strong link between physical activity and major non-communicable diseases, member states

of the WHO agreed to aim for a 10% relative reduction in prevalence of insufficient physical activity by 2025, as one of the nine global targets to help prevent and treat non-communicable diseases[107]. If current trends continue, this 2025 global physical activity target will not be met.

1.4.2.1 Definition of physical activity

Physical activity is defined as bodily movement produced by skeletal muscle that requires energy expenditure in excess of resting energy expenditure[108]. Exercise, a type of physical activity, is defined as planned, structured, and repetitive bodily movement performed to improve or maintain one or more components of physical fitness[109]. According to the WHO guidelines, adults aged 18-64 years should do at least 150-300 minutes of moderate-intensity aerobic physical activity, or at least 75-150 minutes of vigorous-intensity aerobic physical activity, or an equivalent combination of moderate- and vigorous-intensity activity throughout the week[110].

1.4.2.2 Quantification methods of physical activity

There are numerous quantifying methods for physical activity. Three key features of physical activity that should be considered when choosing one for a research study: (1) quality of physical activity measured, (2) time and/or effort required to accomplish objectives, and (3) cost/burden to administrator[111].

The doubly-labelled water (DLW) method is the gold standard for measuring total energy expenditure; however, it is expensive, has a high subject burden, is time-consuming, and cannot capture qualitative data[112]. Self-report questionnaires, on the other hand, are the most common method of physical activity assessment[113], but they rely on participants' recall ability. Validation studies comparing results from these questionnaires with data from DLW produced inconsistent results[114]. Self-report questionnaires

vary by what type of data they measure, how data are reported, quality of the data collected (e.g., measures of intensity, inclusion of leisure and non-leisure activity), and how data are obtained[111, 115]. Their advantages include cost effectiveness, accuracy in measuring vigorous/intense activity, and ease of administration[111]. However, their robustness may be limited when measuring light or moderate activity[115], assessing energy expenditure[116], and their sensitivity to external factors including seasonal variation, social desirability and complexity of the questionnaire[117].

More recently, accelerometers gained popularity given their accuracy and their ability to capture large amounts of data[114]. Accelerometers measure acceleration in real time and detect movement in up to three orthogonal planes [118], which are then translated into energy expenditure or physical activity patterns[111, 119]. Strength of accelerometers include minute-by-minute on-line monitoring and capturing intensity level [111]. On the other hand, accelerometers are expensive, require technical expertise, and are considered more effort and time-consuming to participants[111].

Other methods and devices used to measure physical activity include armbands, heart-rate monitors, pedometers, self-report activity diaries/logs, and direct observation by an independent observer [111].

1.4.2.3 Assessing physical activity

Physical activity is commonly classified using a four-category classification, with each domain reflecting the purpose of the activity[120] :

1. Occupational (work-related).
2. Domestic (housework, yard work, physically-active childcare, chores).
3. Transportation (walking or bicycling for the purposes of going somewhere).

4. Leisure-time (discretionary or recreational time for hobbies, sports, and exercise).

In order to accurately assess an individual's overall activity level, four main features of physical activity need to measured[120] :

1. Type of activity.
2. Frequency of the activity. How often is the activity done per day or per week?
3. Intensity: is the activity enough to get the person warm and slightly out of breath (moderate-intensity activity) or out of breath and sweaty (vigorous-intensity activity).
4. Time: for how long is the individual active?

The intensity of an activity is commonly expressed in metabolic equivalents (METs). MET is defined as the ratio of energy expenditure during a specific physical activity to the resting energy expenditure (RMR). One MET is defined as the amount of oxygen consumed while sitting at rest and is equivalent to 3.5 ml O2/kg/min (1.2 Kcal/min for a 70 Kg person)[121]. MET can express the intensity and energy expenditure of physical activities in a way that can be comparable among individuals of different weight. Likewise, actual energy expenditure (e.g., in calories or joules) during an activity depends on an individual's weight and thus, the energy expenditure of the same activity will be different for individuals of different weight.

The *Compendium of Physical Activities* was developed for use in epidemiological studies to standardise the assignment of MET intensities in physical activity studies[122]. Based on the assigned MET values, physical activities are classified as (i) low intensity (< 3 METs), (ii) moderate intensity (3-6 METs) and (iii) vigorous intensity (> 6 METs)[123]. Common physical activities and their assigned MET values are enlisted

in **Table 1.3**.

1.4.2.4 Physical activity and T2D risk

The beneficial impact of physical activity on reducing risk of T2D has been well-established[124, 125]. Cohort studies, which were largely Euro-centric, consistently demonstrated the inverse association of physical activity with T2D in men and women. Like obesity, two of the largest prospective studies investigating the relationship of physical activity and T2D were part of the NHS (in women) and HPFS (in men). Among ~70,000 initially healthy US women participating in the NHS, walking briskly for at least 2.5 hours per week was associated with 25% reduction in T2D over 8 years of follow-up after adjustment for age, BMI, and other risk factors for T2D[126]. In the HPFS, which examined ~8,000 male health professionals aged 40 – 75 years for 10 years, demonstrated a reduction in T2D risk by 11% with each 10 MET-hour increment in weekly walking energy expenditure; an equivalent increase in weekly vigorous activity conferred a 12% reduction in T2D risk[127].

Sedentary behaviour — waking behaviours characterised with low energy expenditure (TV viewing, desk work, etc.) — has also been shown to be associated with increased mortality and morbidity, mostly independently of moderate-to-vigorous physical activity participation[128, 129]. In people with, or at risk for developing, T2D, extended sedentary time is associated with poorer glycaemic control and clustered metabolic risk[128, 130, 131]. Furthermore, interrupting prolonged sitting with 15 min of post meal walking and with 3 min of light walking and simple body-weight resistance activities every 30 min has been shown to improve glycaemic control[128]. Studies also demonstrated the association of sedentary behaviours like TV watching on risk of T2D in both men[127] and women[82], suggesting the importance of reducing sedentary time in disease prevention.

In line with cohort study findings, interventional studies have also demonstrated the beneficial effects of lifestyle modification including increased levels of physical activity on reducing T2D risk. In the US Diabetes Prevention Program, study participants were randomised to lifestyle-modification programs (with the goals of at least a 7% weight loss and ≥150 minutes of physical activity per week), metformin, or placebo[88]. After an average follow-up of 2.8 years follow-up, the incidence of T2D was 11.0, 7.8, and 4.8 cases per 100 person-years in the placebo, metformin, and lifestyle groups, respectively. Compared to placebo, lifestyle intervention reduced the incidence of T2D by 58% and metformin by 31%. Investigators concluded that lifestyle intervention was significantly more effective than metformin at reducing risk of T2D[88]. In the Finnish T2D Prevention Study, 522 middle-aged, overweight men and women with IFG were randomly assigned to either a control group or to an intensive lifestyle intervention designed to promote healthy eating and exercise patterns[87]. Compared to controls, participants in the diet and exercise intervention group lost significantly more weight (3.5 vs. 0.8 kg). In addition, their risk of developing T2D was reduced by 58% over 3 years. Protective effects of exercise on risk of developing T2D were still observed in participants who did not attain the weight loss goal[87].

1.4.2.5 Mechanisms for the protective effect of physical activity on T2D risk

Exercise is routinely advised as an essential part of weight loss strategies for overweight/obese individuals, particularly those who are at high risk of T2D or cardiovascular disease. Despite not having major effects on weight on its own, exercise has several health benefits that occur independently of changes in body weight[132]. These include skeletal muscle adaptations that improve fat and glucose metabolism, and insulin action, enhance endothelial function, have favourable changes in blood lipids, lipoproteins, and homeostatic factors[132]. In addition, it can reduce blood pressure, postprandial lipemia and glycaemia, and proinflammatory markers.

Exercise helps decrease concentrations of fatty acid metabolites, which in turn decreases the risk of fatty acid-induced insulin resistance and stimulates the translocation of GLUT4 to the plasma membrane, thereby improving glucose uptake[133]. Studies have also shown the impact of acute bouts of exercise on insulin action in muscle and liver[134]. Aerobic exercise increases muscle glucose uptake by up to 5-fold, independently of insulin. After exercise, glucose uptake remains elevated by insulin-independent (~2 hours) and insulin-dependent (up to 48 hours) mechanisms if exercise is prolonged[128, 135]. That is linked to muscle glycogen repletion[128]. Individuals with T2D who participate in regular exercise training also experience reductions in HbA1c, triglycerides, blood pressure, and insulin resistance[136]. Moreover, high-intensity interval training (HIIT) promotes rapid enhancement of skeletal muscle oxidative capacity, insulin sensitivity, and glycaemic control in adults with T2D[137]. Conversely, physical inactivity has been closely associated with increasing islets' workload and lowering their efficiency through various pathways, such as endoplasmic reticulum stress, mitochondrial dysfunction, oxidative stress, and promoting the apoptosis of β-cells[138]. Sedentary behaviours can also induce insulin resistance via lowering energy expenditure, dysregulating lipid homeostasis, and enhancing lipid storage[139, 140].

1.4.3 Genetics of T2D

As discussed above, environmental risk factors clearly play a major role in the development of T2D, but they do not impact everyone in the same way. Even with the same environmental exposures, some people are more susceptible to developing T2D than others, and this increased risk appears to be inherited[141]. However, while hereditary factors' role in T2D development is clear, the actual genetic variants involved were completely unknown prior to the advent of advanced genetic technologies.

1.4.3.1 Genetic architecture of T2D risk

T2D is a clinical endpoint and different people reach that by different pathogenic pathways – it is probably a cluster of related conditions rather than a single entity. This was discussed in section 1.4.

With the exception of rare monogenic disorders, the majority of T2D cases result from the interaction of genetic loci at multiple different chromosomal sites with environmental exposures experienced throughout the lifespan[142]. The gene-environmental interactions that result in T2D are still poorly understood. The genetic component of T2D is not concentrated in one region, but it appears to underlie interactions of multiple genes scattered across the genome. It is possible, or even likely, that multiple common genetic variants of small effect (common disease common variant hypothesis) contribute to the genetic component of T2D[143]. It is also a possibility, however, that the effect is due to multiple rare variants or even few rare variants of large effect[143]. Whole-genome sequencing experiments have derived a model of heritability of T2D, which is characterised by a prominent contribution to heritability by common variation, a small contribution of rare variation, and evidence of low selective pressure on predisposition alleles[144].

Rare monogenic disorders make up approximately 5% of non-autoimmune diabetes cases and are classified into different subtypes, namely neonatal diabetes mellitus or monogenic diabetes of infancy, dominantly inherited familial forms of early-onset diabetes (Maturity-Onset Diabetes of the Young, MODY) and rarer diabetes-associated syndromic diseases[145]. These forms of diabetes are very often misdiagnosed as T2D, but they are clinically distinct diseases. Monogenic cases of diabetes are caused by single genes with high penetrance of which mutations in the Hepatocyte nuclear factor-1A (HNF1A) and the glucokinase (GCK) gene are the most common[145]. I have not considered these cases in my thesis, but it should be noted that polymorphisms in or near monogenic diabetes genes may contribute to adult polygenic T2D[145].

1.4.3.2 Identification of T2D risk genes

Linkage analysis, candidate gene approach, large-scale association, and genome-wide association studies (GWAS), have identified over 400 genome-wide significant loci associated with T2D[144, 146-151].

1.4.3.2.1 Linkage studies

Linkage is the tendency for genes and other genetic markers to be inherited together because of their location close to one another on the same chromosome. Linkage studies are used to map genetic loci using observations of related individuals. Through genotyping genetic markers and studying their segregation through family pedigrees, it is possible to map disease or trait loci. Although these methods were successful in detecting rare variants with large effects for classical single gene disorders, they proved rather unsuccessful in identifying genes that are involved in complex polygenic complex disorders. Through linkage studies, only two genes: calpain 10 (*CAPN10*)[152] and transcription factor 7-like 2 (T-cell specific, HMG-box) (*TCF7L2*)[153] were reliably identified to be linked to T2D.

TCF7L2 was first discovered as a T2D susceptibility gene in a Mexican-American population when a strong linkage signal was mapped to chromosome 10q in a study conducted in 1999[153]. Subsequently, it was fine-mapped in the Icelandic population and confirmed in US and Danish cohorts, where the risk locus was found to be located in intron 3 of the gene[154]. Since then, multiple single-nucleotide polymorphisms (SNPs) in the *TCF7L2* gene have been strongly confirmed in several genome-wide association studies (GWAS) across ethnic populations. *TCF7L2* remains the most replicated and the strongest T2D risk signal to date[155]. For example, the OR for TT homozygotes and TC heterozygotes versus CC homozygote of the rs7903146 variant in *TCF7L2*, was estimated to be 1.968 (1.790 - 2.157), 1.406 (1.341-1.476), respectively, and the population attributable risk (PAR) for the TT/TC genotypes of this variant is 16.9% overall[155].

1.4.3.2.2 Candidate gene studies

Candidate gene studies involve focused sequencing of genes selected based on prior hypotheses on their role in T2D pathogenesis, including glucose metabolism, insulin secretion, insulin receptors, and lipid metabolism. Unexpectedly, most genes known to be involved in insulin secretion and action were not found to be associated with T2D. The number of genes identified to be associated with T2D through candidate gene studies were also limited. These included, peroxisome proliferator-activated receptor gamma (*PPARG*), insulin receptor substrate 1 and 2 (*IRS1* and *IRS2*), potassium inwardly-rectifying channel, subfamily J, member 11 (*KCNJ11*), Wolfram syndrome 1 (*WFS1*), HNF1 homeobox A, B, and 4A (*HNF1A*, *HNF1B*, *HNF4A*)[156-158].

1.4.3.2.3 Genome wide association studies

Several key developments have made performing large-scale association studies, including GWAS in many thousands of samples possible. These include the completion of the Human Genome Project in 2003, which in turn led to detailed mapping of common single nucleotide polymorphisms (SNPs) and patterns of linkage disequilibrium (LD) produced by The International HapMap Consortium (HapMap) in 2005[159].

GWAS makes use of a 'hypothesis-free' approach that interrogates hundreds of thousands of common SNPs across the genome, typically defined to have minor allele frequency of at least 5%. In 2007, five genome-wide association studies were published[160, 161]. These studies increased the number of confirmed T2D susceptibility loci to nine, and included *PPARγ*, *KCNJ11*, *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *HHEX/IDE*, *FTO*, and *SLC30A8*[160, 161]. Majority of identified loci were related to β-cell function albeit the exact mechanisms remain largely unknown[162].

In 2008, the understanding of the genetic contribution of T2D from early GWAS was greatly enhanced through international collaborative efforts that increased power to detect disease loci. Zeggini et al. performed a meta-analysis of three GWASs, including DIAGRAM (DIAbetes Genetics Replication And Meta-analysis) consortium, FUSION (Finland-United States Investigation of NIDDM Genetics), and WTCCC (WEllcome Trust Case Control Consortium). This meta-analysis identified at least six novel loci for T2D, including *JAZF1*, *CDC123-CAMK1D*, *TSPAN8-LGR5*, *THADA*, *ADAMTS9*, and *NOTCH2*[163]. Variants in *JAZF1*, *CDC123-CAMK1D*, *TSPAN8-LGR5*, and *THADA* were also reported to affect β-cell function[164, 165]. Tens of studies followed, steadily expanding the list of T2D genetic variants.

Between 2016-2018, a quantum leap occurred with the identification of T2D susceptibility variants, from smaller, cumulative advances to the description of up to around 250 genome-wide significant loci[166]. Several developments enabled such rapid progress. Firstly, the interrogation of large-scale, accessible resources like the UK Biobank, together with the availability of worldwide GWAS results, exome sequencing and genotyping studies, resulted in tripling of effective sample size, with 74,000 cases of T2D and 820,000 controls of European descent were available to study[166]. Secondly, dense and accurate genotype imputation with the sequenced haplotypes of the 1000 Genomes[146] and the Haplotype Reference Consortium[147, 167] has increased the number of genetic variants interrogated to up to 30 million variants, a 100-fold increase compared with early GWAS efforts[166]. Thirdly, direct genotyping efforts in the exonic areas of common and rare alleles have enhanced the ascertainment of coding variants[147], and thus, susceptibility variants for T2D now have minor allele frequencies ranging from 50% to 0.02% and per-allele odds-ratios from 1.04 to 8.05[166]. These findings further justified the model of heritability of T2D as characterised by whole-genome sequencing experiments (described above section 1.4.3.1)[144].

1.4.3.2.4 Glycaemic control and genetic susceptibility to T2D

In parallel with disease-focused discovery GWAS, efforts to understand the genetic contribution of T2D susceptibility were further enhanced by large-scale meta-analyses of genome-wide data for continuous diabetes traits in individuals without diabetes. These commenced in 2010, with the meta-analyses conducted by the MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium) investigators in European ancestry populations[168]. Newly identified loci included those associated with fasting glucose (in or near *ADCY5*, *MADD*, *ADRA2A*, *CRY2*, *FADS1*, *GLIS3*, *SLC2A2*, *PROX1* and *C2CD4B*) and one influencing fasting insulin and insulin resistance (near *IGF1*).

In large studies, like MAGIC, widely available simple measures, including fasting plasma glucose or insulin and glycated haemoglobin have been used[169, 170]. Whereas studies based on gold-standard measures that involve oral glucose tolerance tests or continuous intravenous measurements are smaller in size[171, 172]. Both approaches have enhanced the understanding of underlying mechanisms of common predisposition to T2D and genetic effects on glucose concentrations, insulin secretion, and insulin resistance.

Efforts from numerous studies following the former approach helped expand the genetic landscape of T2D, successfully identified loci influencing glycaemic traits, helped identify insulin secretion as a major driver of associations of T2D susceptibility loci, and enriched for association signals in/near genes involved in cell cycle regulation[173]. Additionally, these studies revealed a multifactorial nature of disease predisposition due to the observed substantial etiological heterogeneity in pathways to T2D[173]. On the other hand, gold-standard-based studies have enabled in-depth physiological characterisation of T2D susceptibility variants, and the identification of novel loci (e.g., *GRB10*[171], *BCL2*[174], *FAM19A2*[174], and *NAT2*[175]), which were not previously identified in much larger meta-analyses based on simple measures

or indices.

1.4.3.2.5 Insights from genetic studies of T2D

During the past years (over a decade), GWASs have dissected the genetic architecture of T2D in a hypothesis-free way. Identified variants have unravelled diverse biological mechanisms and pathways involved in disease pathogenicity. Most of these variants increase T2D risk through insulin secretion, through direct or indirect effects on; (1) islet function (e.g., variants in *KCNJ11*), (2) islet development (e.g., *HNF1A* and *WFS1*), (3) incretin signalling (variants in *GLP1R*)[176]. Conversely, there are numerous T2D risk loci that seem to primarily operate through insulin resistance (variants in *KLF14* and *PPARG*) and obesity (variants in *FTO*)[176].

TCF7L2 is the most intensively studied locus for T2D risk to date. Risk alleles of *TCF7L2* were found to be associated with enhanced gene expression in human islets, as well as impaired insulin secretion (*in vitro* and *in vivo*)[177]. *TCF7L2* variant rs7904146 was also found to affect risk of T2D, at least in part, through modifying the effect of incretins on insulin secretion, independent of reduced secretion of glucose-dependent insulinotropic polypeptide (GIP) and glucagon-like peptide 1 (GLP-1), but rather due to its effects on the sensitivity of β-cells to incretin[178]. In addition, *TCF7L2* rs7903146 variant was also found to be associated with impaired islet function and morphology in individuals without diabetes[179].

The steadily expanding list of genetic variants delivered by these successive waves of genetic discovery have proven GWAS to be extremely successful in identifying loci contributing to T2D susceptibility. Collectively, these associations can explain up to ~20% of the heritability of the disease, in Europeans at least[149]. Comparable analyses in non-European populations have been limited by sample size. Estimates of T2D heritability (median of 40%) suggest that approximately half of the genetic contribution to the

variation in risk can be quantified for a single individual.

1.4.3.2.6 Polygenic risk score

In contrast to classical Mendelian diseases, where the presence of a specific mutation, contingent on penetrance, could determine the onset of disease, the genetic risk for complex, multifactorial diseases is probabilistic where risk variants are used as predictors that quantify discrete increments in overall risk[180]. That is due to the modest effects exerted by each associated risk variant, and their small influence on clinical outcomes. Polygenic risk score (PRS) is an approach that sums the effects of risk alleles associated with a given condition and generates an aggregate estimate of genetic risk[181]. This approach was found justifiable given genetic associations seem to work in an additive fashion, with little to no evidence of epistasis[181].

Early attempts at combining T2D risk alleles utilised 3 well-replicated variants in candidate genes (*KCNJ11*, *PPARG*, and *TCF7L2*) to create the first restricted PRS (rPRS) for T2D[182]. Using this rPRS they could only improve T2D prediction minimally, falling well short of the threshold typically required for clinical implementation[182]. With the start of the GWAS era and the identification of dozens of novel loci, several independent groups attempted to create new sets of rPRS for T2D risk prediction[181]. These studies revealed a beneficial effect of adding T2D SNPs in risk prediction, but that was only 1-2% better than a model including clinical covariates alone[181].

Given the limited prognostic value of rPRS for T2D, more recently, a different approach that involves constructing a "global" PRS (gPRS) was utilised[183]. Through the use of a bioinformatic machine-learning tool, genetic inheritance patterns (LD) are leveraged in Europeans populations to create a gPRS that accounted for almost all common variations across the genome, rather than only including a subset[183].

Their findings revealed a 2.75 fold increase in odds of T2D in individuals at the top 5% of the gPRS, compared to the remaining 95%, a risk similar to that observed in some monogenic diseases[183].

To date, the clinical use of PRS in predicting T2D has been limited. For example, the use of rPRS for T2D prediction has revealed that those with higher genetic risk had a greater reduction in HbA1c at 1-year follow-up in response to therapy with sulfonylureas[181]. In an independent study, an rPRS based on 123 T2D variants has demonstrated that each increase in SD increased risk of developing an insulin dependence by 7%[181]. Further work is highly required to elucidate PRS as an effective tool in clinical context.

1.4.3.3 Missing heritability of T2D and the role of the Epigenome

A variety of population, family, and twin-based studies have estimated heritability of T2D to be ranging from 20 to 80%[184, 185]. The lifetime risk of developing T2D in individuals who have one parent with T2D is 40%, and that increases to 70% if both parents are affected[186]. Additionally, compared to individuals without a positive family history of T2D, first-degree relatives of individuals with T2D are about 3 times more likely to develop the disease[180]. Concordance rates for T2D in dizygotic twins were observed to be 20%-30%, and reached up to 70% in monozygotic twins[187]. Familial risk observed in studies restricted to parents in the 35-60-year age range is higher than those including older individuals: this suggests a greater role played by environmental factors in those who develop T2D at a later stage in life[188].

Given that commonly occurring genetic variants do not account for as much of T2D susceptibility as anticipated, it is thought that the remaining genetic susceptibility could be accounted for by rare genetic variants and, importantly, gene-environment interactions. This led to the hypothesis that there is an

additional layer in the genome that remains to be discovered, with epigenetic modifications being one possible explanation[189].

Epigenetic regulation refers to heritable changes in gene expression and phenotype that are not determined by changes in DNA sequence[190]. It has been implicated as a natural integrator of external environmental exposure and genetic susceptibility in common disease[191]. Thereby, these mechanisms may potentially explain the link between rapid environmental changes and observed obesity and T2D epidemic[192]. In pancreatic islets, epigenetic modifications regulates β-cell maturation and function, and thereby insulin secretion[189]. The central role of epigenetic modifications, particularly DNA methylation, in T2D is supported by a number of case-control studies, which demonstrated extensive changes in the islet methylome in islets of people with T2D. Some of the initial epigenetic studies identified altered methylation patterns at the promotor regions of key T2D candidate genes in pancreatic islets of people with T2D compared to non-diabetic controls[193, 194]. For example, a whole-genome bisulfite sequencing study identified ~26,000 differentially methylated regions (DMRs) in islets from people with T2D, seven of which were annotated to *PDX1*[195]. These DMRs could probably contribute to the observed reduction in *PDX1* expression in islets of people with T2D[196-198]. T2D epigenetic studies also extended to investigating the methylome of insulin-targeted tissues including skeletal muscle, liver and adipose tissue[189].

Epigenetic modifications could also mediate the effects of environmental exposures on disease risk. Several studies support this hypothesis. For example, they demonstrated the association of lifestyle intervention in sedentary men with altered DNA methylation and gene expression in skeletal muscle and adipose tissue[199, 200]. These epigenetic alterations were annotated to some T2D candidate genes, such as *TCF7L2*, *KCNQ1*, *FTO* and *THADA*, which showed both increased and decreased methylation in response

to exercise[199, 200]. Furthermore, the molecular mechanisms underlying the effect of nutrient surplus or caloric restriction on insulin resistance and islet function were also linked to epigenetic modifications. For example, in adipose tissue, overfeeding altered methylation of over 600 regions, including increased methylation of *SLC2A4*, which encodes *GLUT4.* Caloric restriction was also shown to alter DNA methylation of *NFATC2IP* in blood samples. This DMR increases expression of *NFATC2IP*, and is associated with greater weight loss in response to high fat rather than low fat diet[201].

There is a growing body of literature that highlights the association of epigenetic markers with obesity and weight change. For example, methylation of *HIF3A* in blood was found to be associated with BMI in multiple studies, and its methylation in cord blood was associated with weight and adiposity at birth, and in peripheral blood with childhood obesity.[202, 203] Studies do not support causal effects of *HIF3A* methylation on BMI , but rather suggest that BMI affects/regulates *HIF3A* methylation[204].

Overall, although we know that the environment can affect genetic function and thereby T2D pathogenesis through epigenetic mechanisms, we still need to consider the origin of these mechanisms and distinguish heritable epigenetic modifications from those which are environmentally induced. Furthermore, like GWAS studies, findings from epigenome studies were often based on relatively homogeneous western populations. Hence, more diverse populations living in different environments need to be included in future studies.

1.5 Ethnic predisposition to T2D

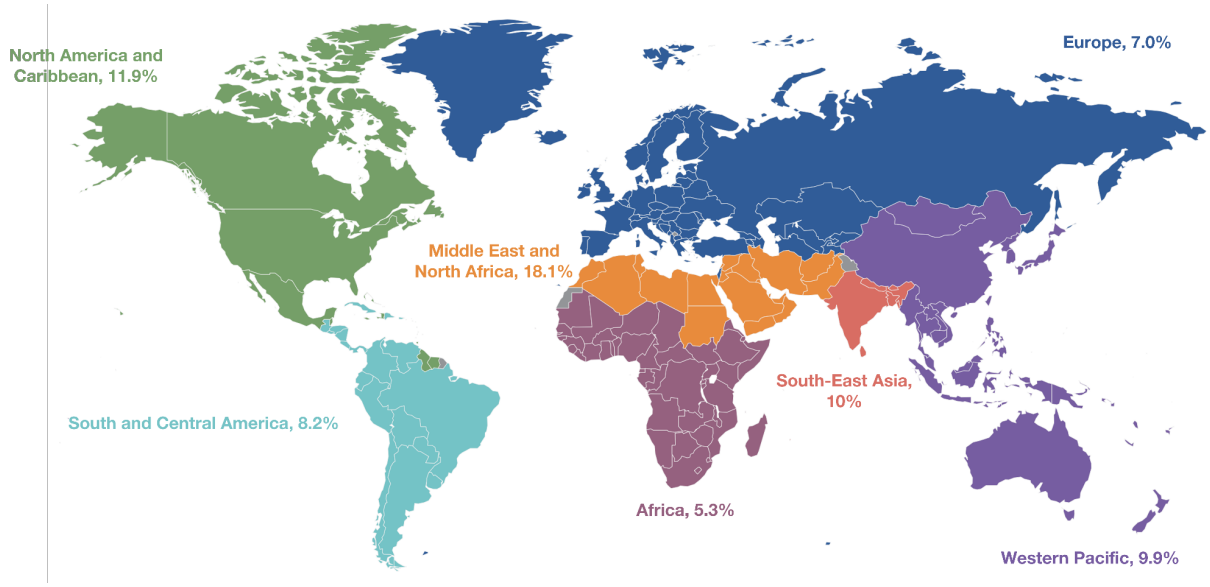
1.5.1 Variation in T2D rates across ethnic groups

In this thesis, the terms "ethnicity" and "ancestry" are used interchangeably. An ethnic group is a group of people who share cultural ideas and beliefs that have been a part of their community for generations, including language, religion, history, types of food and celebrations. People from the same ethnic group do not always share the same religion, and on the other hand, people from different ethnic groups may share the same religion despite being from different cultures.

Although T2D is a common disease and a global pandemic, its rates vary significantly between countries, geographic regions, and ethnic groups[205]. According to the IDF estimates in 2021, the Middle-East and North Africa (MENA) region had the highest rates of diabetes worldwide, with 1 in 6 adults living with diabetes (73 million) and an age-adjusted prevalence of 18.1%[3]. Africa, on the other hand, was estimated to have the lowest rates of diabetes in 2021, with 1 in 22 adults living with diabetes (24 million) and an age-adjusted prevalence of 5.3%[3]. **Figure 1.4** is a map summarising diabetes estimates around the world according to IDF in 2021. Note that studies on individuals of African descent do not support the low diabetes rates IDF estimates for Africa (discussed more in this section below).

These variations may be driven by differences in environmental exposures related to urbanisation/westernisation. A wealth of epidemiological data that compared diabetes prevalence rates between ethnic groups living in different environments, along with other studies comparing multi-ethnic populations, support this view. In 1972 and 1973, for example, a survey conducted in Japan reported a prevalence of diabetes of about 4% in a population aged 40-69 years residing in rural areas, whereas a slightly higher prevalence was found in Tokyo for the same age group[206].

Figure 1.4 Global rates of diabetes by region. World map showing age-adjusted comparative prevalence rates of diabetes according to the International Diabetes Federation (IDF) in 2021. The seven regions presented were defined according to the IDF.



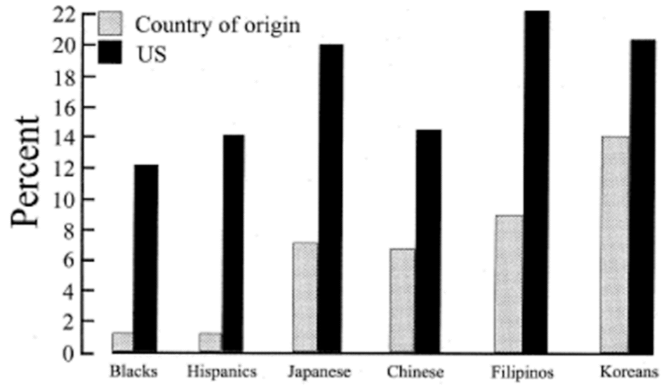
Diabetes prevalence among Japanese living in Japan has gradually increased with time, but the same ethnic group (Japanese) were found to have higher rates of diabetes when living outside Japan; in Hawaii and in continental US[207]. Epidemiological observations in other ethnic populations, including Chinese, Filipinos, African-American, and Hispanics, have shown similar results[207].

Figure 1.5 summarises the variation in diabetes prevalence rates within the same ethnic group under different environmental conditions (i.e., living in different countries/areas). The figure demonstrates the impact of an adoption of "western" lifestyle on increasing risk of diabetes across all ethnic groups. Yet, environmental exposures appear to have a heterogeneous impact on people of different ethnicities, indicating the differential predisposition to diabetes across ethnic groups, even when exposed to similar

environmental challenges. This is demonstrated in **Figure 1.6**, where Hispanics, African-Americans, and Asians, appear to have a much higher predisposition than individuals of European ancestry[207].

More recently, in adults aged ≥ 75 years living in the US, prevalence rates of diabetes reported ranged from 17.7% in White European individuals to 22.5% in Asians, 30.6% in Hispanics and Africans, and 45.2% in American Indians/Alaska Natives[208]. This diversity was also evident in ethnic subgroups; Indians with the highest risk of disease (38.8%) amongst Asians, and Mexicans (33.4%) amongst Hispanics[209]. In addition to ethnic disparity in prevalence rates, T2D starts at a younger age and at lower BMI cut-offs in Africans and Asians than in White Europeans[101].

Figure 1.5 Prevalence of diabetes for various ethnic groups living in the United States compared to their countries of origin [Figure adapted from Abate 2003 et al.][207].

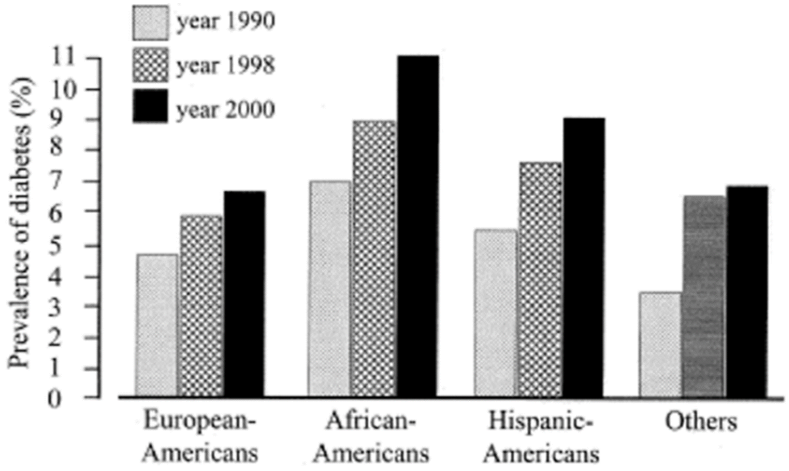


As discussed earlier, T2D is a collection of metabolic disorders, some of which may play a causal role in the ethnic differences in disease prevalence rates, including body fat distribution, adipose tissue function and different levels of insulin secretion and insulin sensitivity. Studying ethnic differences in T2D risk could lead us to better understand disease mechanisms and eventually have implications for clinical and public health efforts in targeting the prevention of the disease especially in the era of precision medicine[101].

1.5.2 Variation in adiposity and T2D susceptibility

Part of this differential ethnic predisposition to T2D can be explained by differences in body composition and fat distribution[207, 210].

Figure 1.6 Prevalence rates and trends of diabetes for various ethnic groups living in the United States [Figure adapted from Abate 2003 et al.][207].



At any given BMI, compared to Europeans, South Asians have higher waist circumference, WHR and 5-7% higher total body fat[101]. Whereas lean mass and lean-to-fat-mass ratios are lower in South Asians compared to White Europeans and East Asians. In addition, South Asian new-borns have a combination of low muscle mass and high truncal fat, which are presumed to predispose to insulin resistance later in life[101]. South Asian populations also exhibit greater levels of VAT, ectopic fat (intrahepatic lipid content)[211], and larger adipocytes[212] at similar levels of whole-body adiposity compared to Europeans. This has commonly been used to explain the greater risk of T2D in South Asians; their reduced capacity to store excess energy in subcutaneous adipose tissue can lead to the observed increase in ectopic fat deposition[213].

Like South Asians, East Asians have lower obesity rates compared to White Europeans, despite having a less favourable fat distribution, with higher abdominal obesity[101]. Fat distribution patterns in East Asians seem to arise early since prepubescent East Asian boys and girls have lower gynoid fat and fat in extremities compared to white children at pubertal age[214, 215].

On the other hand, individuals of African descent are recognised to be at higher risk of T2D, yet exhibit favourable body fat distribution, as characterised by significantly lower levels of visceral fat compared to other ethnic groups[216]. A large, pooled analysis of black women demonstrated a lesser increase in visceral fat per unit increase in waist circumference, compared to European women, highlighting the questionable role of visceral fat in the development of T2D in black populations[216, 217]. In addition, despite being at a greater risk of T2D, they have lower levels of ectopic fat compared to their European counterparts[216]. This phenomenon has previously been coined the African ectopic fat paradox.

More recently, genetic predisposition to favourable adiposity has been implicated in explaining the ethnic differences in risk of T2D. Large-scale GWASs identified several genetic variants that implicate a paradoxically inverse relationship between adiposity and risk of cardiometabolic disease[172, 218-220]. These adiposity-increasing alleles, named 'favourable adiposity alleles', were found to be associated with a favourable metabolic profile and lower risk of T2D, heart disease and hypertension. Further MRI-based imaging data highlighted the association of these alleles with higher subcutaneous and lower liver fat, with no differences in visceral fat [218]. Trans-ethnic comparisons revealed that South Asians living in the UK have fewer of these protective alleles compared to Europeans [101]. This genetic difference could in turn partly explain the differences in body fat distribution and its relationship with disease in South Asians.

1.5.3 Variation in insulin resistance and action in T2D susceptibility

Numerous epidemiological studies indicated that although degrees of adiposity differ across ethnic groups, these differences cannot fully explain the excess risk[101]. Some populations are characterised with increased insulin resistance and/or impaired insulin action, independent of adiposity. Non-diabetic African American and Hispanic men for example, have higher insulin resistance compared with non-diabetic White Europeans after adjusting for adiposity measures. These observations suggest that higher insulin resistance, independent of obesity, may be in part responsible for the higher prevalence of T2D in these groups[221].

On the other hand, in Asians, insulin secretory defects play a more predominant role than insulin resistance in the pathogenesis of T2D[222]. For example, β-cell function has been shown to be the driving force behind the high T2D rates in Indian Asians[223]. Other studies of β-cell function amongst individuals of Chinese, Malay, and Indian Asians living in Singapore reported the highest early-phase insulin secretion in Chinese, followed by Indian Asians and lowest among Malays. However, Chinese were the most insulin sensitive whilst Indian Asians were the least insulin sensitive[224].

1.5.4 Ethnicity-specific genetic predisposition to T2D

An alternative explanation for ethnic differences in risk of T2D is related to differential genetic predisposition. Although GWA studies were predominantly carried out amongst people of European ancestry, efforts to extend these to include East Asian[225, 226], South Asian[227], Hispanic/Latino[228], and African American[229] populations have been at an increase in the last few years. These have been successful in identifying ancestry-specific signals that are rare or non-existent in most populations but have large effects in their studied populations. **Table 1.4** summarises some of the genetic variants identified in GWAS of T2D in populations of non-Europeans. These variants successfully provided novel

mechanistic insights and potential drug targets for T2D; however, their contribution to the ethnic disparities in disease risk are limited. Together trans-ethnic meta-analyses have demonstrated a remarkable overlap in T2D susceptibility loci between ethnicities and provided evidence for the transferability of European-derived genetic risk scores to other ethnic groups[230, 231]. They also revealed minimal evidence of heterogeneity in allelic odds-ratios on T2D susceptibility at lead SNPs, despite considerable differences in allele frequencies[230, 231]. These findings imply that the majority of common causal variants at T2D susceptibility loci are shared across ancestry groups, and likely arose prior to human migration out of Africa[173].

Increasing the population diversity in genetic studies is valuable for several reasons. Firstly, susceptibility variants might be present at appreciable allele frequencies in a certain ethnic group or even population but not others, as shown by the discovery of associations at *HNF1A*[232] and *IGF2*[233] loci via exome sequencing, and *SLC16A11* locus via GWAS[234] in Latin American populations. Secondly, due to the homogeneity in T2D susceptibility loci across ancestry groups, trans-ethnic studies increase statistical power for new discoveries[231]. Thirdly, the diverse linkage disequilibrium patterns across ancestries are beneficial for localising causal variants underlying association signals[231]. Lastly, when coupled with the study of diverse environmental exposure, it could help infer the effect of susceptibility loci in different settings (gene-environmental mechanisms).

1.5.5 Genetically-isolated populations

The genetic composition of isolated populations differs greatly from large, outbred populations with respect to the amount of variation and allele frequency spectrum[235]. An isolated population can be founded by a subpopulation migrating to a new geographical area, or by certain lifestyle or religious reasons that lead to cultural isolation[235]. These scenarios can introduce a bottleneck, where a small

subpopulation exhibits only a subset of the genetic diversity of the original population. In this setting, genetic drift has a profound effect on the genetic architecture of the population, and thus deleterious variants with large phenotypic effects could arise randomly to higher frequencies[236].

Identification of disease association in such populations is easier compared to admixed populations, in which these variants might not be present or might be extremely rare. For example, in Inuit populations in Greenland, homozygous carriers of a loss-of-function variant in *TBC1D4* was found to confer muscle insulin resistance and T2D risk[237]. This variant segregates at high frequency in Inuit people (minor allele frequency 17%) but is rare or monomorphic in other populations[237]. *TBC1D4* variants were found to account for 10% of T2D in Greenlandic Inuit, where 14% of people carry at least one copy of the risk allele and those who have two copies, are at tenfold higher risk of T2D[237]. Another example is the identification of the loss-of-function variant in *ADCY3*, which is also associated with obesity and T2D in the Greenland Inuit population[238].

1.6 Understanding the excess risk of T2D in the UAE

1.6.1 Demographics: The Arabian Peninsula and the United Arab Emirates

The United Arab Emirates (UAE) is in the South-eastern corner of the Arabian Peninsula (AP). Apart from the UAE, the Arabian Peninsula comprises Kuwait, Bahrain, Oman, Qatar, Saudi Arabia, Yemen, and parts of Iraq and Jordan (**Figure 1.7**). It is ethnically diverse and has arisen from a series of social and cultural influences. Inhabitants of the AP are collectively referred to as “Arabs”; a major pan-ethnic group that inhabits Western Asia, North Africa, and parts of the Horn of Africa. Although they inhabit different parts of the world, Arabs share common linguistic, cultural, religious, and political traditions[239]. The AP sits at the crossroads of significant human migration between the African, European, and Asian

continents, and it represents the epicentre of multidirectional flow between populations of the three continents. This, coupled with local migration that would have occurred between settlements within the peninsula, shaped a diverse admixture group of people[239].

Contemporary UAE was formed by the union of seven emirates or sheikhdoms in 1971 led by Sheikh Zayed bin Sultan Al Nahyan. The Trucial states that formed the UAE were once located on the southern routes of human migration within the peninsula. The UAE population is approximately 10 million with approximately 15% citizens of the country (November 2020). Most other residents are expatriates, with approximately 30% being of South Asian origin.

Figure 1.7 Map of the Arabian Peninsula.

Adapted from: https://www.mappr.co/thematic-maps/arabian-peninsula-history-facts-map/



1.6.2 Population substructure of the UAE

The AP populations are characterised into three distinct primary ancestral groups, namely the Bedouins, the Persians from Central and South Asia, and Africans, with the Bedouins predominating the ancestral pool[240]. Bedouins, a heterogeneous population that dates back to at least 850 BC, are nomadic, desert-dwelling Arab peoples of the Middle East, ranging from camel nomads to more sedentary tribes who combined shepherding with scattered cultivation[240, 241].

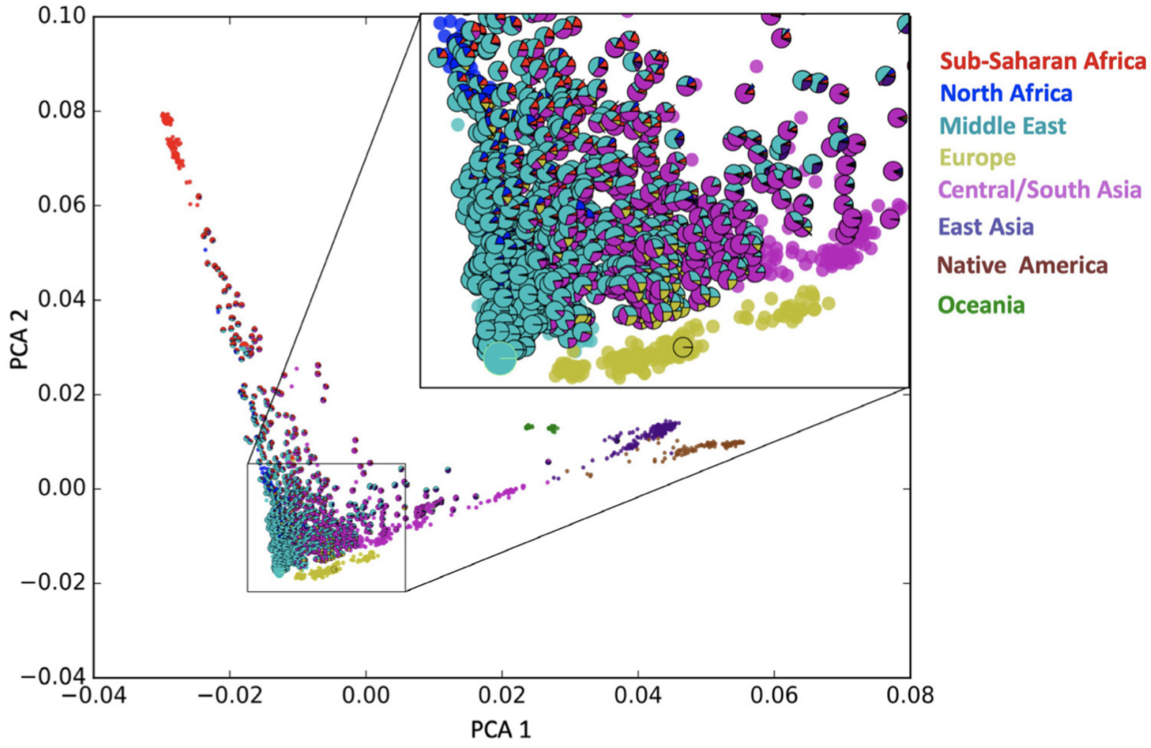
A more recent comprehensive, large-scale population stratification study of the UAE population challenges the notion of these three clusters[239]. Instead, they defined a continuous spectrum of genotype distribution within the population, with the Middle Eastern component dominating in the contemporary gene pool, together with an influence from regions that are juxtaposed to the AP, specifically from Central/South Asia, Sub-Saharan Africa, and to a lesser extent Europe and North Africa[239]. These findings indicate continuous gradients of ancestral population contributions, and the gradual occurrence of admixture along the Eastern seaboard and South-eastern tip of the AP without strong social stratification borders (**Figure 1.8**)[239]. In addition, there was no clear correlation between birth location throughout the UAE and ancestral background. Similar overall results describing a diverse admixed population were observed in Qatar[242] and Kuwait[243].

The UAE is a relatively new nation, but the diversity seen in the population arose over generations before the current political borders were drawn. The observed diversity is thought to have arisen from the contribution of ancestral populations who were already diverse at the time of immigration out of Africa[239]. Alternatively, the contemporary population of the UAE could have been influenced by populations from surrounding regions; migration occurred across the Arabian Gulf into South Asia and where the Red Sea and the Gulf of Aden meet into Africa[239]. Therefore, diversity observed in the UAE is

potentially a function of this migratory pattern of people moving back and forth between the African and Asian continents[239].

Over the last 10,000 years, sequential climatic changes that included aridification and long droughts led to regional abandons and geographic relocation of groups[244]. This induced bottlenecks in the Middle East as observed in archaeology[245] and genetic studies[246].

Figure 1.8 Admixture informed principal component analysis (PCA) plot for 1,000 Arab-Emiratis. Each outlined sample denoting an individual studied is represented as a pie chart of ancestral composition. In addition, HGDP (Human Genome Diversity Project) samples from model populations are shown as circles in their respective colours (no outline). Adapted from Tay G. (2020)[239].



Movements to AP were also promoted by sequential trade hubs, such as pearl diving, or dependence on natural resources like the destruction of Marib dam in Yemen[244]. This has led those groups of people to

adopt nomadic lifestyles and consolidate kinship structures[244]. The impact of such events on the genetic diversity of the local population is not fully understood.

On the other hand, in recent history, with the discovery of oil in 1959, the population of the UAE has grown from 87,000 to 9.9 million, predominantly due to the immigration of expatriates from the Indian subcontinent. This influx would be expected to promote genetic admixture, but against it are social and cultural factors that are thought to have restricted gene flow and diversity[247]. These include the practice of consanguineous marriages for example, which is estimated to be as high as 20-50% of all marriages in the Middle East[248, 249]. The population of the UAE also observes tribal endogamy and marriage with non-Emiratis is discouraged.

1.6.3 Global burden of T2D in the UAE

According to IDF estimates, five countries in the AP region including the UAE are among the top 20 countries with some of the highest rates of diabetes worldwide[3]. Several nationally-representative studies conducted in the region support these estimates[250-257]. Despite this increased risk in the region, these populations remain largely understudied. In 2021, the UAE was estimated to have an age-adjusted prevalence of 16.4% among adults aged 20-79 years, with approximately 991 thousand people living with diabetes in the country[3]. In addition, age-adjusted comparative prevalence rates of IGT and IFG were estimated to be 18.3% and 8.8%, respectively[3]. In 2021, over 4,000 deaths were attributable to diabetes in the UAE, and diabetes-related health expenditure per person was estimated to exceed 2,000 US dollars[3]. According to NCD-RisC, between 1980 and 2014, diabetes rates in the UAE increased from 9.6% to 15.0% in men, and from 10.4% to 15.4% in women[205].

1.6.4 Factors contributing to high T2D risk in the UAE: what has been investigated?

The UAE, like other Gulf Cooperation Council (GCC) member states, has undergone an unprecedented socioeconomic growth in the post-oil era. The increase in *per capita* income has resulted in major demographic and epidemiological transition, with a rise in the prevalence of obesity, diabetes, and other chronic diseases[258]. Factors underpinning this increasing prevalence of obesity and T2D are multifactorial and are thought to be primarily related to economic development in the region, and the parallel shift in culture, lifestyle, and dietary habits.

Overall, the Arab pan-ethnic group is largely under-represented in trans-ethnic studies investigating the role of risk factors (e.g., BMI, central adiposity, fat distribution, and genetic risk factors) in explaining T2D predisposition. *NB: in this section of my thesis, I am focussing on studies investigating the relationship of obesity, physical inactivity, and genetic risk factors with T2D in the AP region and the UAE in particular.*

1.6.5 Obesity and T2D risk in the UAE

The concomitant increase in prevalence rates of T2D and obesity observed in the AP region and GCC states[259], highlights the strong link of obesity with T2D in the region. These trends are in line with obesity and diabetes trends reported by the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) throughout the last 35 years for the region[260]. According to NCD-RisC, from 1975 to 2016, obesity rates in the UAE increased from 9.6% to 28.4% in men, and from 25.1% to 42.5% in women[260]. Despite these observed trends, the contribution of obesity towards increased disease risk has been scarcely studied.

Studies among the GCC states have investigated the relationship of obesity with T2D risk in cross-sectional or case-control studies[250, 251, 253, 254, 261]. The strong association of obesity with T2D risk was consistently demonstrated across all studies in the region. One of the largest studies was conducted in Oman. Al-Riyami and colleagues[253] performed a community based cross-sectional survey in 7,003 individuals, to examine the relationship of obesity with risk of diabetes, hypertension, and hypercholesterolemia. In their sample, individuals with BMI of 30 Kg/m^2 had more than a 2-fold increased risk of diabetes compared to those with normal BMI. In another study, central obesity as measured by waist circumference or WHR, was shown to have a stronger association with diabetes risk compared to that observed for BMI[261].

1.6.6 Physical activity and T2D risk in the UAE

According to the WHO, insufficient physical activity rates in the UAE were estimated to be 39% in males and 49.3% in females[105]. Like obesity, the number of studies investigating the prevalence of physical inactivity and its relationship with T2D risk in the AP region and the GCC states are limited. Studies conducted in Saudi Arabia and Kuwait, which used reliable physical activity instruments, defined sufficiently active based on internationally recognised definitions, and used acceptable population-based samples, reported prevalence rates of insufficient physical activity participation to range between 58% and 61% for men and 71.6% and 73.7% in women[262]. These estimates were higher than those estimated by the WHO[105]. Furthermore, studies conducted in Oman[263], Saudi Arabia[264], and Qatar[254], which examined the impact of physical activity on T2D risk, demonstrated a significant association of lack of physical activity with disease risk.

1.6.7 Genetic risk factors and T2D in the UAE

Although T2D GWASs have recently extended to other non-European populations, including East and South Asians, Africans, and Hispanics/Latinos, there is still an underrepresentation of populations of the Middle East in general. This has created a serious gap in understanding the contribution of genetic factors to risk of T2D within these populations.

The majority of, if not all, studies investigating the contribution of common genetic variants with T2D risk in populations on the AP have examined SNPs that were previously reported to confer increased risk of T2D in other populations. Al-Daghri and colleagues[265] conducted one of the largest targeted analyses of 38 previously established SNPs for T2D, in a sample population from Saudi Arabia (1,166 T2D cases and 1,235 controls free from diabetes). Their findings demonstrated a significant association of 8 out the 38 SNPs with risk of T2D[265]. In addition, they found that the direction of effect of the identified T2D SNPs to be consistent with previously reported results in other ethnic groups. Interestingly, as in other populations, the *TCF7L2* variant (rs790146) showed the strongest association with T2D. Other studies, however, that also investigated the association of this variant (rs790146) with risk of T2D in Saudi Arabia[266], Oman[267], and UAE[268], reported contradictory results.

In the UAE, Al Safar and colleagues[268] studied an extended family of 66 T2D cases and 112 controls free from diabetes and reported novel associations of variants at *GABRA4* and *PRKD1* with T2D (**Table 1.5**). In another study of the general population of the UAE, Osman and colleagues[269] studied 464 Arab-Emirati T2D cases and 415 controls free from diabetes and reported novel associations of *KCNK3* and *RARB* with risk of T2D (**Table 1.5**). These studies are some of the first attempts to discover novel genetic variants conferring risk of T2D in Arab people; however, further studies to confirm these findings, with larger sample sizes and higher power are still highly required.

1.7 Research problem, hypothesis, and objectives

1.7.1 Research problem and scope

There is a global increasing concern about the rising tide of T2D around the world, particularly in populations of the AP/Gulf region. Variation in T2D rates between geographical regions are sometimes assumed to be driven entirely by the differential distribution of environmental triggers. Global and regional estimates highlight the concurrence of increased obesity and lack of physical activity with increased rates of T2D in populations of the Gulf region, although various other regions in the world with similar environmental exposures, have much lower rates of disease. Despite the severity of the health situation in the Gulf region, their populations remain understudied, and their inclusion in trans-ethnic studies investigating differential contribution of environmental and/or genetic risk factors towards T2D risk is non-existent. Given the unique genetic architecture of populations in the Gulf region, together with the observed high rates of T2D, it is likely that population-specific genetic variants or undiscovered shared common genetic polymorphisms, underlying distinct disease relationships, may explain their increased risk.

As already indicated in the literature review, my study focusses on the contribution of three primary risk factors (obesity, physical activity, and genetic susceptibility) towards T2D risk in Arab-Emiratis. It is important to reiterate that studying these factors alone cannot fully evaluate the influence of other lifestyle/environmental, gene-environmental, and epigenetic influences on disease risk. I utilised additional data from the UK Biobank, DIAGRAM, and South Asian GWAS meta-analyses to perform my trans-ethnic analyses and comparisons. It is also crucial to mention that I used epidemiological approaches to achieve the objectives of this study, and causality analyses using genetic tools (Mendelian Randomisation) for example were beyond the scope of my PhD.

1.7.2 Study primary aim and hypothesis

I aim to investigate the contribution of environmental and genetic risk factors towards the risk of T2D in Arab-Emiratis (population of the UAE), compared to other ethnic populations, including Caucasians, South Asians, and Afro-Caribbeans living in the UK. My generic hypothesis is that Arab-Emiratis have an increased genetic predisposition to T2D, making them exceptionally susceptible to westernised environments. More specific hypotheses are described in relevant chapters.

Results of this study will help elucidate whether environmental exposures, including obesity and lack of physical activity can explain the increased risk of T2D in the UAE population, or whether Arab-Emiratis respond differently to environmental exposures, potentially due to an underlying [potentially increased] genetic susceptibility.

1.7.3 Objectives

1. To use publicly-available estimates to unravel the contribution of obesity and insufficient physical activity towards diabetes risk across different populations.
2. To establish a cross-sectional sample of Arab-Emiratis that is representative of the general population to investigate the contribution of obesity towards increased risk of T2D among Arab-Emiratis (EM1), compared to individuals of Caucasian (CAU), South Asian (SA), and Afro-Caribbean (AC) descent from the UK Biobank (UKBB) baseline cross-sectional data.
3. To collect a sample of Arab-Emirati men and women (**EM2**) with or without T2D in a case-control study to enable the investigation of:
    a. Trans-ethnic comparison of the relationship of central adiposity as measured by waist circumference, waist-to-hip ratio, and anthropometry-derived visceral adipose tissue volume with T2D risk among Arab-Emiratis and other ethnic groups in the UKBB cohort.

b. The association of adiposity measures with insulin resistance and β-cell function as measured by the homeostatic model assessment in Arab-Emiratis.

c. Investigate the contribution of physical inactivity and sedentary time towards risk of T2D and its related metabolic disturbances in Arab-Emiratis, compared to other ethnic groups in the UKBB cohort.

4. To carry out a genome-wide association analysis on 1,200 Arab-Emiratis from the EM2 sample population to identify genetic effects underlying T2D in the population through:

a. Identifying known and novel loci that are associated with T2D in Arab-Emiratis

b. Evaluating heterogeneity and the strength of association of cosmopolitan genetic effects underlying T2D across ethnic groups (Arab-Emiratis, Europeans and South Asians), and identifying additional potentially novel loci through trans-ethnic meta-analysis.

Table 1.1 Criteria for the diagnosis of type 2 diabetes according to the American Diabetes Association (ADA) and the World Health Organisation (WHO).
FBG ≥126 mg/dL (7.0 mmol/L). Fasting is defined as no caloric intake for at least 8 h.*
OR
2-h PG ≥200 mg/dL (11.1 mmol/L) during OGTT. The test should be performed as described by WHO, using a glucose load containing the equivalent of 75 g anhydrous glucose dissolved in water. *
OR
HbA1c ≥6.5% (48 mmol/mol). The test should be performed in a laboratory using a method that is NGSP certified and standardised to the DCCT assay. *
OR
In a patient with classic symptoms of hyperglycaemia or hyperglycaemic crisis, a random plasma glucose ≥200 mg/dL (11.1 mmol/L).
DCCT, Diabetes Control and Complications Trial; FBG, fasting blood (plasma) glucose; OGTT, oral glucose tolerance test; WHO, World Health Organisation; 2-h PG, 2-h plasma glucose.
* In the absence of unequivocal hyperglycaemia, diagnosis requires two abnormal test results from the same sample or in two separate test samples.

Table 1.2 Definition of prediabetes according to WHO and ADA criteria	
	Criteria for diagnosis of prediabetes
World Health Organisation (WHO)	
Fasting blood (plasma) glucose (mmol/l)	6.1-6.8
2 h oral glucose tolerance test (mmol/l)	7.8-11.0
HbA1c (%)	6.0-6.4
American Diabetes Association (ADA)	
Fasting plasma glucose (mmol/l)	5.6-6.9
2 h oral glucose tolerance test (mmol/l)	7.8-11.0
HbA1c (%)	5.7-6.4

Table 1.3 Metabolic Equivalent (MET) values estimates for commonly performed physical activities, from the Compendium of Physical Activities.	
Physical activity	**MET**
Light intensity activities	**< 3**
Sleeping	0.9
Watching television	1.0
Writing, desk work, typing	1.5
Walking, 2.7 Km/h, level ground, strolling, very slow	2.3
Walking, 4 Km/h	2.9
Moderate intensity activities	**3 to 6**
Cycling, stationary, 50 watts, very light effort	3.0
Walking 4.8 Km/h	3.3
Callisthenics, home exercise, light or moderate effort	3.5
Walking 5.5 Km/h	3.6
Cycling < 16 Km/h, leisure, to work or for pleasure	4.0
Cycling, stationary, 100 watts, light effort	5.5
Moderate intensity activities	**> 7**
Jogging, general	7.0
Callisthenics (e.g., push-ups, sit-ups, pullups, jumping jacks), heavy, vigorous effort	8.0
Running jogging, in place	8.0
Rope jumping	10.0

Table 1.4 Examples of genetic variants identified in genome wide association studies of type 2 diabetes in non-European populations.

SNP ID	Mapped gene	Function	Population	Effect on risk of T2D in odds ratio
rs2283228	*KCNQ1*	Insulin secretion	Japanese	1.26
chr2:136064024	*ZRANB3*	β-cells production or maintenance	Africans	1.05
rs13342232	*SLC16A11*	hepatic lipid metabolism	Mexicans	1.29
rs9470794	*ZFAND3*	β-cell function	East Asians	1.16
rs373863828	*CREBRF*	Adipogenesis	Samoans	0.60
rs137853240	*HNF1A*	β-cells production or maintenance	Oji-Cree Indians	2.00
rs870992	*ITGA1*	β-cell function	Greenlandic Inuit	2.79
rs61736969	*TBC1D4*	Regulates the insulin-dependent uptake of glucose by skeletal muscle and fat tissues	Greenlandic Inuit	10.30

Table 1.5 Genome-wide association studies investigating the association of genetic risk variants with type 2 diabetes in the Gulf region (both studies in Arab-Emiratis).

Reference	Subjects	Chromosome	Position	SNP ID	Mapped gene
Al Safar HS (2013)	Discovery sample: 66 T2D cases and 112 controls; Replication sample: 116 T2D cases and 199 controls.				
		4	44049621	rs7675224	*KCTD8*
		4	44076716	rs4407541	*KCTD8*
		4	44107694	rs4695718	*KCTD8*
		4	45570356	rs7692404	*GABRG1*
		4	45790488	rs135642	*GABRG1*
		4	4871585	rs7662743	*GABRG1\GABRA2*
		4	46012680	rs279856	GABRA2
		4	46582873	rs7679715	COX7B2
		4	46662807	rs2055942	GABRA4
		4	46797750	rs10517178	GABRAB1
		4	46804117	rs1372491	GABRAB1
		14	29300389	rs1953722	PRKD1
		14	29342060	rs10144903	PRKD1
Osman W (2018)	422 T2D cases and 455 controls				
		10	94463945	rs1977833	*HHEX*
		18	60845884	rs12454712	*BCL2*
		3	64703665	rs4411878	*ADAMTS9*
		6	160770312	rs474513	*SLCC22A3*
		11	92708710	rs10830963	*MTNR1B*

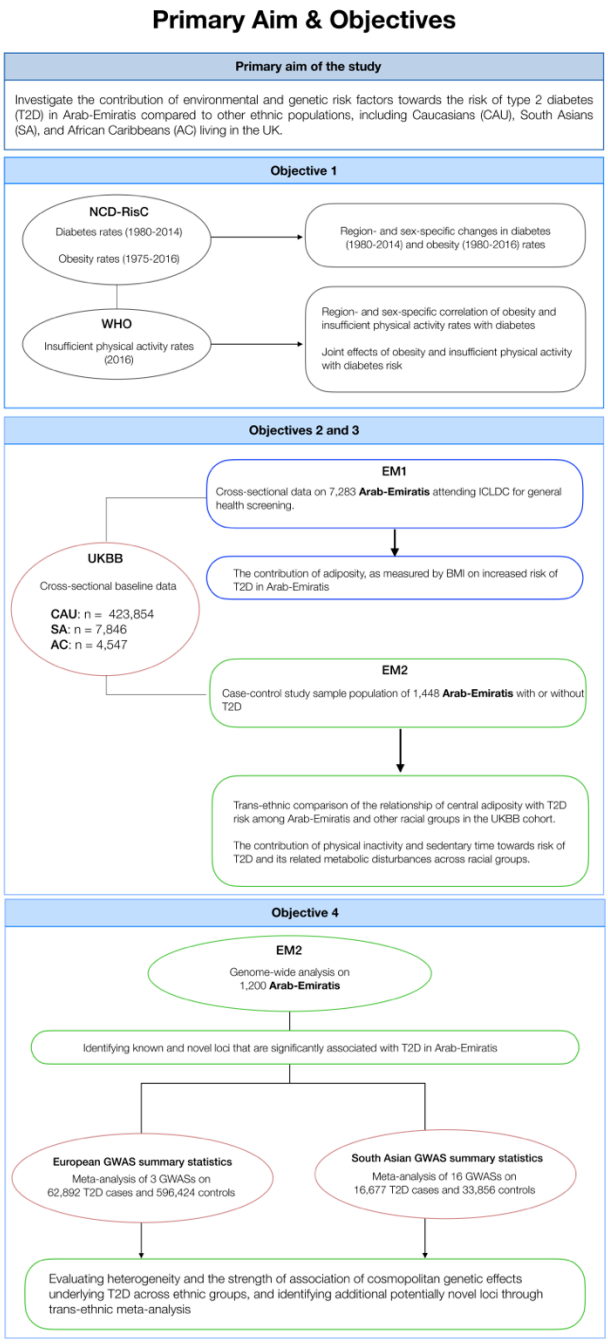
Chapter Two: Materials and Methods

In this chapter, I describe the materials and methods I used to conduct this study. These include study design, participants, screening, diagnoses and definitions, and statistical methods. The chapter is divided into five sections according to the main objectives of my PhD:

1. Section 2.1: Use of publicly-available estimates to unravel the contribution of obesity and insufficient physical activity towards diabetes risk across different populations.
2. Section 2.2: Retrieval and data processing of the UK Biobank (UKBB) baseline cross-sectional data to perform trans-ethnic comparisons.
3. Section 2.3: Establish a cross-sectional sample of Arab-Emiratis that is representative of the general population (EM1) to investigate the contribution of obesity towards increased risk of T2D among Arab-Emiratis compared to individuals of Caucasian (CAU), South Asian (SA), and Afro-Caribbean (AC) descent from the UKBB.
4. Sections 2.4 and 2.5: Conduct a case-control study to collect a sample of Arab-Emirati men and women (EM2) with or without T2D to:
    a. Investigate the contribution of adiposity and physical inactivity towards risk of T2D across ethnic groups.
    b. Conduct a genome-wide association analysis, GWAS, on 1,200 Arab-Emiratis from the EM2 sample population to identify genetic variation underlying T2D in the population.

Figure 2.1 summarises the primary aim and the specific objectives of my study.

Figure 2.1. Flow diagram summarising the primary aim and the specific objectives for my PhD. Data from different sample populations used to successfully reach the objectives of this study are highlighted.



2.1 Objective 1: Use of publicly-available estimates to unravel the contribution of obesity and insufficient physical activity towards diabetes risk across different populations

2.1.1 Study design

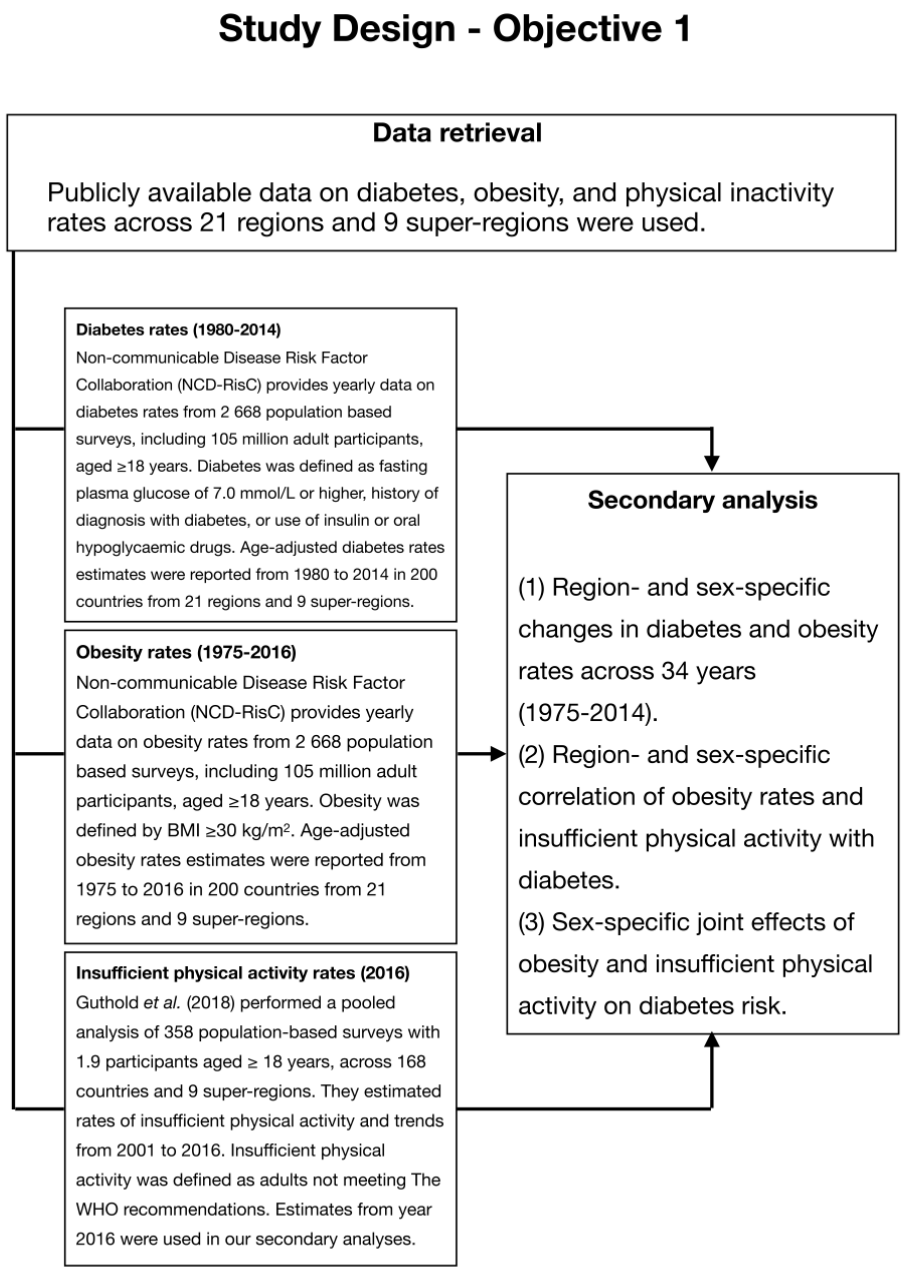
I performed secondary analyses to unravel the relative contribution of obesity and physical inactivity to diabetes risk across different populations, at a global level. I used publicly-available estimates of diabetes, obesity and insufficient physical activity rates retrieved from the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) and the World Health Organisation (WHO) in 9 regions and 200 countries. **Figure 2.2** illustrates the study design for this part of the study.

2.1.2 Data retrieval: publicly-available databases for estimates of diabetes, obesity and insufficient physical activity rates

Region-, country-, and sex-specific estimates of, (1) diabetes prevalence rates from the year 1980 to 2014, and (2) obesity prevalence rates from 1975 to 2016, were downloaded from www.ncdrisc.org . All estimates downloaded and used in this study covered males and females aged ≥ 18 years. Region-, country-, and sex-specific estimates of insufficient physical activity prevalence rates in 2016 were retrieved from a recent publication by the WHO[105].

Estimates of diabetes, obesity, and insufficient physical activity rates were available for 200 (diabetes and obesity) and 168 (insufficient physical activity) countries, respectively. Estimates were available as age-standardised prevalence rates with upper and lower 95% uncertainty intervals, and arranged by sex, country, region, and year.

Figure 2.2. Flow diagram summarising data retrieval, data type, and analysis used for objective 1.



The NCD-RisC and the WHO organised countries into 21 regions and 9 super-regions (**Table 2.1, 2.2**), mostly based on geography and national income. The exception was the High-income Western countries

region: this region comprises English-speaking countries that are geographically separated but exhibit similar trends in cardiometabolic risk factors and outcomes [205]. Region definitions by NCD-RisC and WHO were equivalent. Region-specific analyses here were based on the 9 super-regions, namely: (1) Central Asia, Middle East, and North Africa, (2) Central and Eastern Europe, (3) East and Southeast Asia, (4) High income Asia Pacific, (5) High income Western countries, (6) Latin America and Caribbean, (7) Oceania, (8) South Asia, and (9) Sub-Saharan Africa.

2.1.3 Description of the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) data sources

NCD-RisC is a worldwide network of health researchers and practitioners who aim to document systematically the worldwide trends and variations in non-communicable diseases (NCDs). In my study, I used yearly data for diabetes and obesity prevalence rates.

2.1.3.1 Diabetes and obesity definitions

Diabetes was defined by the NCD-RisC as having fasting plasma glucose of 7.0 mmol/L or higher, history of diagnosis with diabetes, or use of insulin or oral hypoglycaemic drugs[205]. It is important to note that survey data did not separate T1D and T2D, and this limitation is discussed/considered in the respective results chapter (Chapter 3, section 3.6). Obesity was defined as BMI ≥ 30 Kg/m^2.

2.1.3.2 NCD-RisC data inclusion and exclusion

Detailed methods on how the NCD-RisC identifies and accesses data sources was described previous publications[205]. Overall, data on diabetes included were representative of a national, subnational, or community population. History of diabetes diagnosis and the use of diabetes drugs were established with survey-specific questions and/or verified information using visual inspection or medical records. Data sources that relied entirely on self-reported history of diagnosis were excluded, as this approach would lead to high rates of missing undiagnosed diabetes, which primarily forms a subnational share of

all people with diabetes, especially in communities with limited access to health care[270, 271]. In addition, to estimate obesity prevalence rates and trends, the NCD-RisC also included data sources based on samples of national, subnational, or community populations that had measured weight and height. Data sources that were based on self-reported weight and height were excluded, due to being subject to biases that vary by geography, time, age, sex, and socioeconomic characteristics[205].

Here, I summarise their inclusion and exclusion criteria of data sources used to estimate diabetes prevalence rates.

1. Data sources were included in the NCD-RisC database if:

- Measured levels fasting blood glucose, 2-h plasma glucose in an oral glucose tolerance test (2hOGTT), or haemoglobin A1c (HbA1c) were available.
- Study participants were ten years of age and older.
- Data were collected using a probabilistic sampling method with a defined sampling frame.
- Data were representative of the general population at the national, subnational, or community level.
- Data collection in or after 1950.

2. Data sources were excluded from the NCD-RisC database if:

- Relied only on self-reported diabetes diagnosis history or medication status without measurement of at least one of the above-mentioned biomarkers.
- Did not use a standard glucose load to carry out the 2hOGTT.
- HbA1c data measured before the year 2000, to minimise the use of non-standard assays.
- Relied on population subgroups whose glycaemic status may differ systematically from the general population.

2.1.4 Description of the World Health Organisation (WHO) data sources

The WHO estimated prevalence rates of insufficient physical activity among adults aged 18 years and older in 168 countries for three World Bank income groups, and nine regions (**Table 2.2**), from year 2001 to 2016. Data were published in Lancet Global Health journal in 2018[105]. Insufficient physical activity was defined as adults not meeting the WHO recommendations on physical activity for health—i.e., at least 150 min of moderate-intensity, or 75 min of vigorous-intensity physical activity per week, or any equivalent combination of the two.

Data sources were included if they fulfilled the following criteria: (1) the survey questionnaire explicitly included physical activity across four key domains—i.e., work in the house-hold (paid or unpaid), transport to get to and from places (i.e., walking and cycling), leisure time activities (i.e., sports and active recreation); (2) data were collected through random sampling with a sample size of at least 200, and were representative of a national or defined subnational population; (3) prevalence of insufficient physical activity was reported by age and sex, according to the current WHO[110], or former physical activity recommendations[272]. Physical activity data collected using wearable devices were not included due to limited comparability with self-reported data.

2.1.5 Statistical analyses and approaches

Statistical models used by the NCD-RisC to estimate diabetes and obesity prevalence rates by country, year, sex, and age are described in detail in their statistical paper and related substantive papers[273, 274]. Statistical models used by the WHO to derive global and regional estimates of insufficient physical activity rates were also described in detail in their paper[105]. Here I describe both briefly, as they are important to consider when interpreting my results.

The NCD-RisC uses a validated hierarchical probit model in which diabetes or obesity levels and trends in countries were nested in regional levels and trends, which in turn nested in those of super-regions and worldwide[205, 273]. With this structure, estimates for each country and year were informed by the country's own data, when available, and by data from other years in the same country and in other neighbouring countries, especially those in the same region, with available data for similar time periods. To correct for incomparability of diabetes definitions, NCD-RisC used regression models that converted prevalence from those sources into the measured primary outcome (prevalence of fasting plasma glucose ≥ 7.0 mmol/L or history of diabetes diagnosis or use of diabetes drugs)[205].

The WHO applied four key adjustments to physical activity survey data to improve comparability. Data from surveys using former recommendations on physical activity and health only, were converted to the main outcome (current physical activity and health recommendations)[105]. Surveys that estimated physical activity rates in urban samples only, were adjusted to include estimations of prevalence in rural areas using information from surveys reporting both urban and rural prevalence rates. Sex- and age-specific prevalence estimates were derived using the WHO Standard Population[275]. In addition, to produce global and regional estimates, they first created population-weighted sex-specific and age-specific estimates for each subgroup and year and then age-standardised these estimates [105]. Using the bootstrap method, they drew 1000 samples, which contain 80% of all survey data to produce uncertainty intervals for these estimates[105].

Statistical approaches that I used to perform a secondary-analyses of the NCD-RisC, and the WHO data are described in detail in Chapter 3 section 3.4.2. Briefly, region- and sex-specific differences in diabetes (from 1980 to 2014) and obesity (from 1975 to 2016) rates and trends were assessed. Correlation of diabetes with obesity and insufficient physical activity was tested separately using Pearson Correlation

tests. Country- and sex-specific joint effects of obesity and insufficient physical activity prevalence rates with diabetes were based on estimates in 163 countries in years 2014 (diabetes) and 2016 (obesity and insufficient physical activity).

2.2 Objective 2: Retrieval and processing of the UK Biobank baseline cross-sectional data to perform trans-ethnic comparisons.

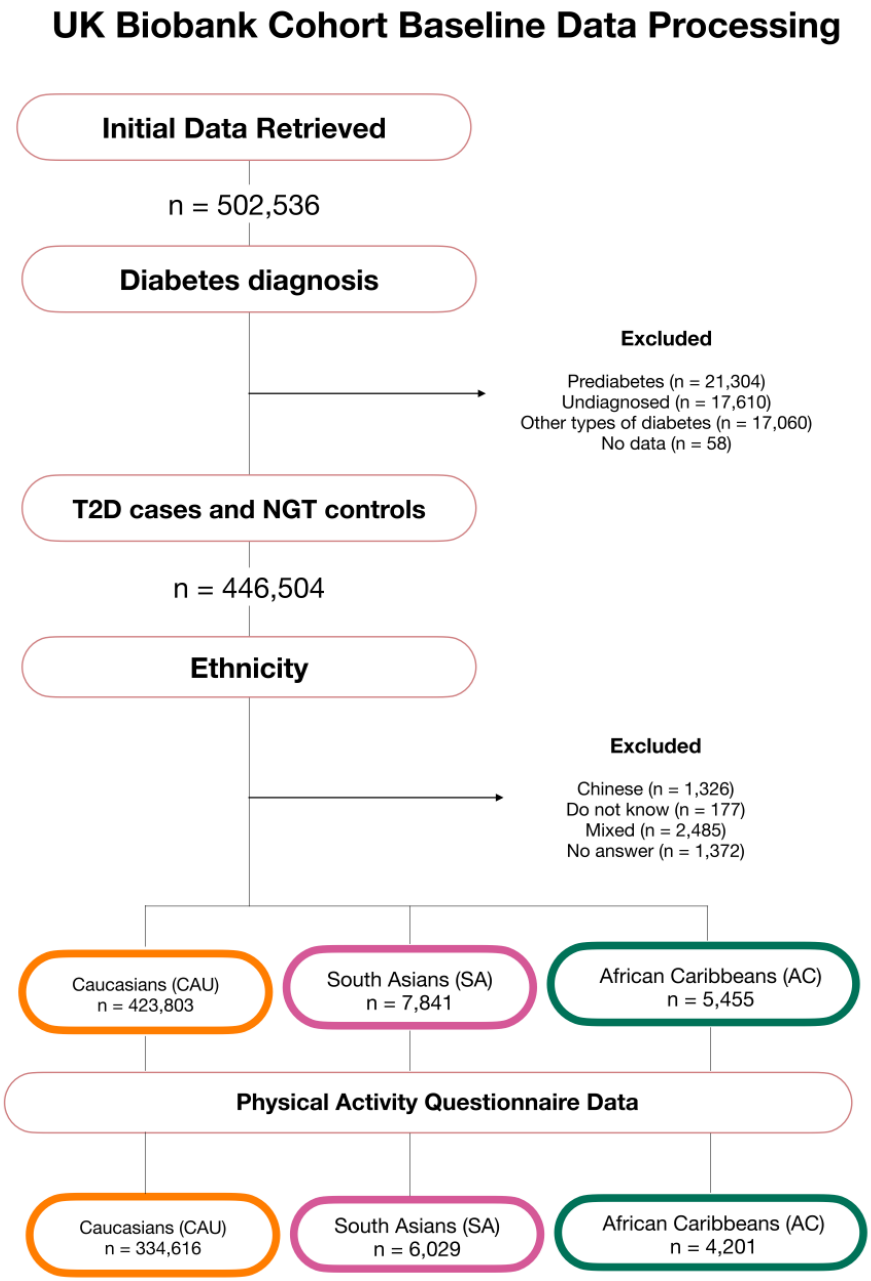
2.2.1 Study design

In order to perform trans-ethnic comparisons throughout my study, I used cross-sectional baseline data from individuals who participated in the UK Biobank (UKBB) (cohort) study. Access to the UKBB data was granted under application ID 23889. **Figure 2.3** describes processing of the UKBB data retrieval.

2.2.2 UK Biobank Cohort: background

The UKBB generic cohort was established with funding from the Wellcome Trust and Medical Research Council. It is a population-based observational cohort study designed to provide sufficient power to assess a wide range of chronic disease outcomes[276], including T2D. Detailed phenotyping of participants at baseline was anticipated to facilitate research of the mechanisms underlying exposure-outcome associations [277] in different study designs such as cohort, case-control and cross-sectional approaches. The UKBB includes over 500,000 volunteer participants, aged between 40-69 years at recruitment (2006–2010). Baseline assessments of participants included completion of lifestyle questionnaires and collection of biological samples (i.e., blood, saliva and urine). Participants were also clinically-assessed and given written informed consent to health follow-up via medical record linkage[277].

Figure 2.3. Flow diagram summarising data retrieval and processing of the UK Biobank baseline cohort.



Middle- and older-aged UK residents were recruited based on NHS primary care registrations. Following enrolment, participants attended an initial baseline assessment visit at one of 22 centres located throughout England, Scotland and Wales. Study protocols have been described in detail in previous publications[278]. In brief, UKBB delivered questionnaires via touch-screen devices to assess lifestyle factors that included dietary habits, physical activity behaviour via physical activity questionnaires[278], psychological wellbeing, cognitive functioning, social factors, personal and familial self-reported medical histories, and sociodemographic information. After completion, trained staff then clarified and expanded upon the given responses. Research nurses recorded physical and anthropometric measures, including blood pressure, height, and weight, using standardised equipment, and drew blood samples. Processing of samples followed standardised protocols mandated by the UKBB to minimise random and systematic error and produce adequately harmonised data across assessment centres. The UKBB catalogued the collected data in an online database – the Data Showcase (ukbiobank.ac.uk/data-showcase) – with detailed protocols of data collection and handling for each variable. The UKBB study received ethical approval from the UK National Research Ethics Committee Northwest.

2.2.3 Diagnoses and definitions

2.2.3.1 Glycaemic profile and T2D in UKBB

I categorised individuals into T2D and controls using the following data fields of UKBB: age diabetes diagnosed (Data field 2976), self-reported diabetes (Data field 20002), diabetes diagnosed by doctor (Data field 2443), started insulin within one year of diagnosis (Data field 2986), and Gestational diabetes only (Data field 4041). Subjects were defined to have T2D if they fit any of the following criteria: (1) self-reported having T2D, (2) self-reported having diabetes and age of diagnoses ≥ 30 years, or (3) did not

self-report any type of diabetes but were indicated to have doctor-diagnosed diabetes at ≥ 30 years. I excluded subjects if they; (1) self-reported having T1D, (2) self-reported having gestational diabetes, (3) self-reported diabetes and age of diabetes diagnosed was < 30 years, (4) self-reported having diabetes and started insulin within one year of diagnosis, or (5) did not report any type of diabetes but had age < 30 years when diabetes was diagnosed. Individuals with prediabetes were also excluded. Those are the ones that had HbA1c between 42.0 and 47.9 mmol/mol as defined previously [279]. All other individuals were defined as controls with normal glucose tolerance (NGT). After following these exclusion criteria, 446,504 individuals were included: 423,410 controls with NGT and 23,094 cases with T2D (**Figure 2.3**). HbA1c levels were initially not included when defining T2D in the UKBB. I performed a sensitivity analysis to check whether my initial definition of T2D that I described here, underestimates the actual number of cases in the UKBB cohort. I did that by checking if any of the individuals who were initially defined as having NGT have an HbA1c > 47.9 mmol/mol (Chapter 4, section 4.5.8).

2.2.3.2 Adiposity

Body mass index (BMI) was calculated as weight in kilograms divided by height in metres squared. Individuals were classified into four BMI categories; (1) Normal (BMI = 18.0-24.9 Kg/m^2); (2) Overweight (BMI = 25.0-29.9 Kg/m^2), (3) Class I/II Obese (BMI = 30.0-39.9 Kg/m^2); and (4) Class III Obese (BMI ≥ 40.0 Kg/m^2). Waist-to-hip ratio (WHR) was calculated as waist measurement (cm) divided by hip measurement (cm). Central adiposity based on WHR was defined as > 0.9 in males and > 0.8 in females. Lin et al.[280] equation (expanded version) was used to estimate visceral adipose tissue (VAT) volume from simple anthropometric parameters:

VAT volume (L) = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip +

37.74

Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

2.2.3.4 Ethnicity

Ethnicity definition was based on self-reported ethnicity (UKBB Data-Field 21000). Out of 446,504 individuals, data from three ethnic groups were included in the present study: (1) individuals of Caucasian descent (CAU: British, Irish, other White background) (n = 423,803), (2) individuals of South Asian descent (SA: Indian, Pakistani, Bangladeshi, other Asian) (n = 7,841), and (3) individuals of Afro-Caribbean descent (AC: Caribbean, African, other Black background) (n = 5,455). Individuals that belonged to any other ethnic group were excluded (n = 5,360) (**Figure 2.3**).

2.2.4 Screening

Here I describe processing of data available as part of my study application (application ID 23889).

2.2.4.1 Physical activity questionnaire

Physical activity in the UKBB baseline cohort was measured using the short International Physical Activity Questionnaire (IPAQ)[281]. IPAQ assesses physical activity undertaken across a set of domains including leisure time, domestic and gardening activities, work-related and transport-related activity. The short form asks about three specific types of activity undertaken in the last 7 days, these include (1) vigorous physical activities (refer to activities that take hard physical effort and make the participant breathe much harder than normal), (2) moderate activities (refer to activities that take moderate physical effort and make participant breathe somewhat harder than normal), and (3) transport/walking activities (includes at work, home, travel, or any form of recreation, sport, exercise). Frequency (measured in days per week) and duration (measured in time per day) are collected separately for each specific type of activity (vigorous, moderate, walking). The items are structured to provide separate

scores on walking, moderate-intensity, and vigorous-intensity activity, as well as a combined score to describe total activity levels. Computation of the total score requires the summation of the duration (in minutes) and frequency (in days) of walking, moderate-intensity and vigorous-intensity activity. A standardised measure of volume of activity is computed by its requirements defined in METs to yield a score in MET-minutes (as previously described in Chapter 1 section 1.4.2.3). MET-minute scores are equivalent to kilocalories for a 60-kilogram person. According to the Ainsworth Compendium MET score estimates for different types of activities[122], the following values are used for the analysis of IPAQ data: 3.3 METs for walking activities, 4.0 METs for moderate-intensity activities, and 8.0 METs for vigorous-intensity activities[282]. Individuals were then categorised into three IPAQ activity groups based on their total activity levels: (1) Low- (LPA), (2) Moderate- (MPA), or High- (HPA) physical activity. Physical inactivity prevalence was defined as the prevalence of LPA.

Data fields used to retrieve IPAQ-derived estimates in the UKBB included: (1) 22032 - IPAQ activity group, (2) 22038 - MET minutes per week for moderate activity (MET_mod), (3) 22039 - MET minutes per week for vigorous activity (MET_vig), (4) 22037 - MET minutes per week for walking (MET_walk), and (5) 22040 - Summed MET minutes per week for all activity (TPA). IPAQ-derived physical activity data were available for 334,616 Caucasians (CAU), 6,029 South Asians (SA), and 4,201 Afro-Caribbeans (AC).

2.2.3.2 Blood biochemical analyses

As part of my study application, blood biochemical analyses data retrieved from the UKBB included: plasma glucose, fasting plasma glucose, haemoglobin A1C (HbA1c), total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), and triglycerides. Majority of UKBB participants were non-fasting. Fasting blood (plasma) glucose data were available for 18,013 CAU, 308 SA, and 227 AC participants only. Insulin data was not available for the UKBB participants. Insulin data from Arab-

Emiratis (EM2 sample population, described in section 2.3) was compared to that reported for the LOLIPOP (London Life Sciences Prospective Population) study[283] among South Asians and Europeans living in the UK.

2.3 Objective 3: Establish a cross-sectional sample of Arab-Emiratis that is representative of the general population to investigate the contribution of obesity towards risk of T2D in the population (EM1)

2.3.1 Study design

I carried out a cross-sectional study to investigate the contribution of adiposity, as measured by BMI, towards the increased risk of T2D in Arab-Emiratis living in the UAE in comparison to individuals of CAU, SA, and AC descent, living in the UK (UKBB cohort, described in section 2.2). In order to assess how representative the EM1 sample of the general UAE population, I compared age structure, as well as age-specific obesity and T2D rates observed in EM1 to that reported by the Weqaya program, a population-wide cardiovascular program that included 50,138 Arab-Emiratis aged 18 years or older[284].

2.3.2 Participants

2.3.2.1 Recruitment centre

Imperial College London Diabetes Centre (ICLDC) is an outpatient facility based in Abu Dhabi, specialising in diabetes treatment. In 2015, 81,503 patients attended ICLDC, with 38% of them diagnosed with T2D, 18.2% with prediabetes, 2.6% with T1D, 1.0% with other types of diabetes, and 40.1% with no diabetes.

Patients attending ICLDC for general health screening makeup 7-10% of all patients, with over 90% of them attending for the first time. Patients who are attending ICLDC for the first time and reported (self-

reported) to be free from diabetes, are allocated a specific care-package, named “screening package”. A care-package includes several blood tests and anthropometric measures that trained nurses must complete for a patient at their clinic visit. Patients attending ICLDC for the first time and reported to have been diagnosed previously with diabetes, are allocated a different care-package called; “New diabetes package”. Lastly, patients who have been previously diagnosed with diabetes at ICLDC and are attending clinic regularly, are allocated a care package called; “Diabetes 3-month follow-up package”, where they are required to attend the clinic every 3 months for follow-up. This part of the study followed the ethical guidelines for retrospective studies, approved by the Medical Research Ethics Committee of ICLDC. Informed consent for use of clinical data in an anonymized form was obtained from patients at the time of enrolment.

2.3.2.2 Inclusion criteria

Medical records of patients attending ICLDC between January and October 2015 were reviewed. This included data from 72,350 patients in total. Patients were included in my analyses if they fit the following inclusion criteria: (1) self-reported ethnicity and ethnicity indicated in their health insurance is Arab-Emirati, (2) aged ≥ 18 years, and (3) attended ICLDC for general health screening (care-package = “screening package”). 7,556 patients were identified that complied this inclusion criteria. I named this sample population (n = 7,556): **EM1**. All analyses that involved the EM1 sample population were performed in November 2015. Data prior to October 2015 were not included in my analyses as individuals were invited and started attending ICLDC for general health screening by the end of 2014. Prior to that, the majority of attendees had diabetes or a related comorbidity, and thus were highly unlikely to represent the general population.

2.3.3 Screening

All patients attending ICLDC for general health screening are screened by trained nurses following ICLDC's established protocols and the American Diabetes Association (ADA) guidelines. Nurses at ICLDC are registered with HAAD (Health Authority of Abu Dhabi) and DOH (Department of Health) licence. They are also Advanced Cardiovascular Life Support (ACLS), Basic Life Support (BLS), Paediatric Advanced Life Support (PALS), and Wound Care certified. Anthropometric measures, including weight, height, systolic and diastolic blood pressure were recorded. Blood was withdrawn after an overnight fast, and biochemical analyses were carried out following standardised methods in the pathology laboratories at ICLDC. Pathology laboratories at ICLDC are managed by National Reference Laboratory (NRL), which is accredited by the College of American Pathologists (CAP) laboratories and the International Organisation for Standardisation (ISO).

Results extracted from ICLDC's medical records included fasting measurements of plasma glucose (FBG), glycated haemoglobin (HbA1c), triglycerides, total cholesterol, low-density lipoprotein (LDL), and high-density lipoprotein (HDL). Waist and hip measurements are not part of general screening at ICLDC, and thus, data on these measures were not available for EM1.

2.3.3.1 Anthropometric measures

2.3.3.1.1 Height

Height was measured using a stadiometer, mounted on a hard, flat surface. Patients were measured without shoes, standing with their back to the height rule, both feet together, and with head, buttocks and heels touching the wall. Participants were asked to look directly ahead, and the stadiometer bar moved down to touch the top of the head, pressing the hair flat. Height was recorded to the nearest 0.1cm.

2.3.3.1.2 Weight

Weight was recorded using digital scales. The scales were mounted on a hard, flat surface. The patient was weighed after an overnight fast, wearing light clothing only. Weight was recorded to the nearest 0.1 Kg.

2.3.3.1.3 Blood pressure

Blood pressure was measured using Dinamap 1846 XT automated blood pressure monitor. Blood pressure was measured in the right arm, with the subject in the supine position, after 10 minutes rest. The cuff was positioned level with the heart, with the centre of bladder over the brachial artery, and the lower edge of the cuff 2-3 cm above the antecubital fossa.

2.3.3.2 Blood sample collection and nurses training

Trained staff collected venous blood samples by venepuncture technique. Nurses at ICLDC are registered with HAAD (Health Authority of Abu Dhabi) and DOH (Department of Health) licence. They are also Advanced Cardiovascular Life Support (ACLS), Basic Life Support (BLS), Paediatric Advanced Life Support (PALS), and Wound Care certified.

Venepuncture was performed at the antecubital fossa, with a tourniquet applied gently around the upper arm. Once the vein had been cannulated, the tourniquet was released. Samples were collected directly into pre-labelled vacuum filled sample tubes, with additives where appropriate, and once filled the sample tubes were gently inverted 2-3 times.

2.3.3.3 Blood biochemical analyses, laboratory certification and quality control

As part of ICLDC's patients' clinical care, blood biochemical analyses included lipid profile, glucose and HbA1c. Blood sample tests were analysed in the pathology laboratories at ICLDC, using commercially

available assays summarised in **Table 2.3**. Pathology laboratories at ICLDC are managed by National Reference Laboratory (NRL), which is accredited by the College of American Pathologists (CAP) laboratories and the International Organisation for Standardisation (ISO). The laboratories' quality control (QC) utilises both internal and external quality programs. The internal quality program consists of 2 levels of controls (normal and pathological), programmed, and performed daily based on previously-approved QC frequency brackets. As for the external QC, pathology laboratories at ICLDC are enrolled to CAP, and as part of this program they test CAP surveys throughout the year and their results are compared to peer laboratories and the methods are referenced across all laboratories enrolled in CAP. Laboratory performances are reviewed in the CAP website. The laboratories also utilise peer group review systems such as Unity to ensure optimal performance across the UAE. They also send out samples to be processed at a secondary testing laboratory (inhouse, Cleveland Clinic Abu Dhabi) or a referral laboratory at Mussafah Civils Laboratory (ICAD); samples are packed, picked up daily by an authorised logistics company (Agility) and delivered to the referral lab.

2.3.4 Diagnoses and definitions

2.3.4.1 Glycaemic profile and T2D in EM1

T2D and prediabetes were diagnosed by ICLDC's endocrinologists following American Diabetes Association (ADA) guidelines. Adults (≥ 18 years) are evaluated for T2D if they are overweight (BMI ≥ 25 Kg/m^2) and have one or more of the following risk factors: first-degree relative with diabetes, women who delivered a baby weighing > 9 lb, diagnosis of hypertension > 140/90 mmHg, diagnosis of polycystic ovarian syndrome, history of gestational diabetes mellitus (GDM), or acanthosis nigricans. Individuals with FBG between 5.5-7.0 mmol/L, and HbA1c between 5.7-6.4% were defined to have prediabetes. Individuals with FBG of > 7.0 mmol/L and HbA1c of > 6.4% were defined to have T2D. Individuals with prediabetes (n = 273) were excluded from EM1.

2.3.4.2 Obesity

Body mass index (BMI) was calculated as weight in kilograms divided by height in metres squared. Individuals were classified into four BMI categories; (1) Normal (BMI = 18.0-24.9 Kg/m^2); (2) Overweight (BMI = 25.0-29.9 Kg/m^2), (3) Class I/II Obese (BMI = 30.0-39.9 Kg/m^2); and (4) Class III Obese (BMI ≥ 40.0 Kg/m^2).

2.4 Objective 4: Conduct a case-control study to collect a sample of Arab-Emirati men and women (EM2) with or without T2D to investigate the contribution of environmental and genetic risk factors for T2D in the UAE

2.4.1 Study design

I conducted a case-control study to investigate the contribution of environmental risk factors including obesity and physical inactivity, and genetic risk factors towards risk of T2D among Arab-Emiratis. In total, 1,485 individuals were recruited for this part of the study, with 72% of them recruited by myself and the rest by research recruiters. I named this sample population **EM2**. The study design and data processing are described for each objective of the study which required the use of this sample population in the analyses (**Figure 4.2** in Chapter 4, **Figure 5.1** in Chapter 5, and **Figure 6.4** in Chapter 6).

2.4.2 Participants

2.4.2.1 Recruitment centre

Participants were recruited from two ICLDC branches in Abu Dhabi and Al Ain. This part of the study was approved by the Medical Research Ethics Committee of ICLDC and was registered in clinical trials (ID: NCT04270097). Informed consent to participate in this part of the study was obtained from patients at the time of enrolment.

2.4.2.2 Inclusion criteria

Targeted patients had to fit the following inclusion criteria; (1) males or females, (2) reported ethnicity and ethnicity indicated in their health insurance is Arab-Emirati, (3) fasting (overnight fast), and (4) aged ≥ 18 years. Cases and controls were recruited from two separate source populations. The control source population comprised patients attending ICLDC for general health screening (care-package = "Screening package"). Participants with prediabetes or T2D were either newly-diagnosed patients or were existing prediabetes/T2D patients attending ICLDC for follow-up (Patients with prediabetes or T2D attend ICLDC for follow-up every 3 months, care-package = "Diabetes 3-month follow-up package").

2.4.2.3 Recruitment process

Patients with (cases) or without (controls) T2D were recruited from ICLDC between 2017 and 2020. I identified my target population through ICLDC's reporting system. If a patient fits the study's inclusion criteria, a research package specific to my study was added to their records. Targeted patients were approached in the assessment area by myself or other research recruiters. Prior to assessment and blood sample collection, patients were first welcomed and informed about the study and its significance. They were also handed an Arabic version of an information sheet that detailed what the study is about and what is required from individuals who agree to participate (English version attached in Appendix 1). If a patient agreed to participate, they were asked to sign an informed consent form (English version attached in Appendix 1). The nurse then proceeded with patient assessment and blood sample collection. During this time, the patient was interviewed by myself or the recruiter for approximately 10-15 minutes to complete a physical activity questionnaire (described in section 2.4.4.1). Blood samples collected for research purposes were then taken to ICLDC's Research laboratory for processing and storage (described in section 2.4.3.4). Signed consent forms were scanned and added to the participant's records.

2.4.3 Screening

All patients attending ICLDC are screened by trained nurses following ICLDC's established protocols and the American Diabetes Association (ADA) guidelines.

2.4.3.1 Anthropometric measures

Sections 2.3.3.1.1 – 2.3.3.1.3 above in this chapter describe how height, weight, and blood pressure measurements were taken and recorded for patients attending ICLDC.

Waist and hip measurements were taken with patients wearing light clothing only, and standing erect, with feet together. Waist was defined as the point midway between iliac crest and lowest rib, identified by palpation, and waist circumference measured as the minimum circumference at this level. Hip circumference was defined as the maximum circumference over the greater trochanters and buttocks. Measurements were made using a non-stretchable measuring tape, applied around the body without twists, and gently touching the body, without compressing soft tissue. All measurements were recorded in centimetres, to the nearest 0.1 cm.

2.4.3.2 Blood sample collection and nurses training

Described in section 2.3.3.2 in this chapter.

2.4.3.3 Blood biochemical analyses, laboratory certification and quality control

Described in section 2.3.3.3 in this chapter.

2.4.3.4 Blood for storage

In addition to blood samples that were normally taken as part of patient's clinical care, we collected: (1) 8-millilitre of blood in Ethylenediaminetetraacetic acid (EDTA) (AB Medical, Inc.) anticoagulant tubes for

DNA extraction, (2) 8-millilitre of blood in PAXgene RNA tubes (BD Biosciences), and (3) 10-millilitre of blood in serum separation tubes (SST) (AB Medical, Inc.). Sample handling was performed by a trained laboratory technician. EDTA samples were stored at -80°C freezers. PAXgene samples were maintained at room temperature (18-22°C) for 2 hours after blood collection, and then transferred to a freezer set at -20°C for 24 hours. Following that, tubes were transferred and stored at -80°C freezers. SST samples were first centrifuged at 2000g for 10 minutes at room temperature. Plasma and buffy coat layers were aliquoted into separated tubes and freeze at -80°C.

2.4.4 Physical activity

The following section describes physical activity data collection in Arab-Emiratis using a self-report questionnaire. Physical activity data was also collected using accelerometers in a subset of the EM2 sample population.

2.4.4.1 Global Physical Activity Questionnaire

The WHO developed the Global Physical Activity Questionnaire (GPAQ) for physical activity surveillance in 2002. Since then, the validity and reliability of this questionnaire has been tested repeatedly[285], and was proven to be adaptable to incorporate cultural differences[286]. It has been used in over 100 countries globally, mainly through the WHO STEPwise approach to NCD risk factor surveillance (STEPS) programme. GPAQ is designed to collect information on physical activity participation in a 'typical week' in three domains as well as sedentary behaviour, comprising 16 questions in total. It covers several components of physical activity, including intensity, duration (in time per day), and frequency (in days per week). The three domains in which physical activity is assessed are occupational/work physical activity, transport-related physical activity, and physical activity during discretionary/leisure time. In each domain, the duration and frequency of physical activity participation are recorded.

When calculating an individual's energy expenditure using GPAQ, METs (the metabolic equivalent of task, described previously Chapter 1 section 1.4.2.3) are used to express the intensity of physical activity: (1) moderate-intensity activities at work, transport, or recreation are assigned a value of 4.0 METs, (2) vigorous-intensity activities are assigned a value of 8.0 METs, and (3) inactivity is assigned a value of 1.0 MET. Total moderate and vigorous physical activity levels (MVPA) was computed as the sum of all MET-min/week from MVPA performed at work, transport, and recreation. Data collected were then filtered to remove any invalid or missing values by following the GPAQ analysis guide[287]. GPAQ scoring protocol was used to create the following indicators: total (TPA) and domain-specific physical activity (i.e., work, transport, leisure). According to GPAQ guidelines[287], individuals were categorised into having: (1) Low- (LPA), (2) Moderate- (MPA), or High- (HPA) physical activity (**Table 2.4, 2.5**). Physical inactivity prevalence was defined as the prevalence of LPA.

2.4.4.2 Physical activity accelerometers: ActiGraph monitors

I used ActiGraph devices, which are accelerometer-based activity monitors, to quantify physical activity in a subset of participants from the EM2 sample population. I initially used ActiGraph GT3X+ (ActiGraph, Pensacola, FL), a trial-axial accelerometer that measures and records time at varying accelerations, ranging from ± 6 G's. Each participant was given an activity monitor and was asked to position it on their hip (as a belt) for 7 days (including weekends) for at least 10 hours per day. After 3 months, 65 out of 190 participants agreed to wear the accelerometers. Valid data was defined as wear time of at least 3 days including at least 1 weekend day. Out of the 65 participants, valid data was obtained from 22 (33.8%).

To increase recruitment speed, and improve participant compliance, I switched to using ActiGraph GT9X monitors, which can be worn by participants on the wrist (as a watch). Activity monitor data described in this thesis is based on data collected using ActiGraph GT9X only. The ActiGraph GT9X is a small (3.5 x 3.5 x 1 cm), lightweight (14 grams), water-resistant device that measures accelerations from the subject's amplitude (g) and frequency (Hz) of movement in three axes (X axis: anterior-posterior, Y axis: vertical, Z axis: medial-lateral). I had 14 Actigraph GT9X monitors at the start of the study. In total, 143 participants were recruited for this part of the study.

Using ActiLife v6.13.4 software, I initialised the ActiGraph GT9X monitors at a sampling rate of 30 Hz. Participants were asked to position the monitor on their non-dominant wrist and wear it for 7 days (including weekends) for at least 10 hours per day. Recruiters called participants twice during the wearing period to ensure that participants were complying to wearing their monitors. After the 7-day period, participants returned the watches and the raw data collected was downloaded into ActiLife (Version 6.13.4). Wear-time validation tool in ActiLife was used to flag invalid data for exclusion from further analysis, which is based on the Troiano technique. After data were downloaded successfully into ActiLife, monitors were then re-initialised and were ready to be given to new participants.

Wear-time validation parameters computed included: (1) wear and non-wear length, (2) total dataset length, (3) calendar days with wear time, and (4) weekend days with wear time. One day was defined as wearing the GT9X monitor for at least 10 hours in a 16-hour period (excluding sleep time). Valid data for this study was defined as wear time of at least 3 days including at least 1 weekend day. Data scoring was also performed in ActiLife using the Freedson Combination (1998)[288] algorithm to calculate energy expenditure, Freedson Adult (1998) algorithm[288] to calculate METs, cut points and MVPA. The Freedson Combination algorithm uses the work-energy theorem to calculate caloric expenditure below 1951

counts and the Freedson '98 equation to calculate energy expenditure above 1952 counts. Equations are described in Chapter 5 section 5.4.2.1.2. Out of the 143 individuals recruited, 9 individuals failed to return the Actigraph monitors (6.3%). Valid data as defined by wear time of at least 3 days including at least 1 weekend day was available for 106 participants (74.1%).

2.4.5 Diagnosis and definitions

2.4.5.1 Glycaemic profile and T2D

Diagnosis of T2D and prediabetes at ICLDC were described above in section 2.3.4.1 in this chapter. Insulin resistance and β-cell dysfunction were estimated by the Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) and β-cell function (HOMA-B) which were defined using the following formulas[289] :

HOMA-IR = [Glucose (mmol/L) x Insulin (mU/L)]/22.5

HOMA-B = 20 * [Insulin (mU/L)/Glucose (mmol/L)] - 3.5

2.4.5.2 Adiposity

BMI calculation and classification were described in section 2.3.4.2 in this chapter. Waist-to-hip ratio (WHR) was calculated as waist measurement (cm) divided by hip measurement (cm). Central adiposity based on WHR was defined as > 0.9 in males and > 0.8 in females. Lin et al. equation (expanded version) was used to estimate visceral adipose tissue (VAT) volume from simple anthropometric parameters:

VAT volume (L) = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74

Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

2.5 Objective 5: Conducting a genome-wide association analysis to identify novel and known genetic loci influencing, and contributing to the increased risk of T2D in the UAE

In summary, material and methods described below are for a subpopulation from EM2. As described above (section 2.4.3), individuals were recruited by myself and EDTA blood samples were collected by ICLDC trained nurses. Blood samples were sent to a service provider for DNA extraction and genotyping. Intensity files were sent back, and all data processing and analyses were performed by myself from that point.

2.5.1 Study design and participants

I carried out a genome-wide association study (GWAS) on 1,200 Arab-Emirati men and women from the EM2 sample population. As described previously in section 2.4.2, the EM2 sample population was collected in a case-control study. GWAS was performed in 426 controls with normal glucose tolerance (NGT), 589 cases with T2D, and 185 with prediabetes.

2.5.2 Extraction of Genomic DNA

Genomic DNA extraction from the 1,200 EDTA samples was performed by a service provider, Agiomix Fz LLC (Dubai, UAE), using the NucleoMag® Blood 200 ul procedure (MACHEREY-NAGEL, Düren, Germany), following manufacturer's instruction[290]. DNA extraction and purity were measured using the NanDrop 2000c/2000 UV-Vis Spectrophotometer (NanoDrop, Thermo Fisher Scientific Inc.). In addition, the samples were quantified using the Qubit 2.0 fluorometer from Invitrogen with the Qubit dsDNA BR

Assay Kit (Invitrogen Life Technologies, USA) for accurate quantification. DNA samples were stored at -80°C.

2.5.3 Genome-Wide SNP Genotyping using the Infinium Global Screening Array -24 v3.0 BeadChip

Genome-wide genotyping of the 1,200 extracted DNA samples was performed by Agiomix Fz LLC (Dubai, UAE) using the Infinium Global Screening Array -24 v3.0 BeadChip. The Beadarray® technology forms the fundamental basis of the high throughput single nucleotide polymorphism (SNP) genotyping arrays developed by Illumina. This technology involves the assembly of hundreds of thousands of copies of 50-mer oligonucleotides that are complementary in their sequences immediately adjacent to target SNPs onto 3-micron silica beads. These beads are self-assembled in microwells of fibre optic bundles with uniform spacing of ~5.7 microns and each bead is designed to be specific for tagging a particular SNP.

The Infinium Global Screening Array -24 v3.0 BeadChip features 654,027 SNPs across the genome. It combines highly curated clinical research variants, and quality control (QC) markers for a broad range of clinical research and variant screening applications. Genome-wide SNP genotyping procedure includes the amplification, fragmentation, precipitation, and resuspension of DNA samples. Samples are then loaded to the BeadChip for hybridisation to the oligonucleotide probes, followed by a washing step to remove the non-specifically hybridised DNA. A two-colour single-base extension step is then carried out in which differentially labelled nucleotides are added to the oligonucleotide probes in a complementary manner to the nucleotides present in the hybridised DNA sample. The HiScan or iScan system then scans the BeadChip by laser to excite the fluorophore of the extended base, and high-resolution images of the fluorophores emitted are recorded. The output images and intensity files are imported into GenomeStudio® software 2.0 (Illumina) to extract the genotype information.

2.5.4 Genotype calling in GenomeStudio

2.5.4.1 Input files

Agiomix Fz LLC provided me with intensity output files (idat format) for each of the green and red channels per sample (2,400 idat files for 1,200 samples). I uploaded the files as a genotyping project in GenomeStudio® 2.0. Along with the idat files, I uploaded the cluster (egt) and the bead pool manifest (bpm) files, which were specific for the Infinium Global Screening Array -24 v3.0 BeadChip.

2.5.4.1 Sample and SNP metrics

In GenomeStudio, I computed sample and SNP metrics, which include sample call rate and SNP frequency. Sample call rate is the percentage of loci that a genotype was called for each sample. SNP frequency is the percentage of samples that a genotype was called for each SNP. The cluster algorithm used in GenomeStudio's genotyping module is called GenTrain. It is a measurement of SNP calling quality, ranging from 0 to 1, with higher value meaning better quality. Another quality metric computed by GenomeStudio for each genotype is the GenCall score, which also ranges from 0 to 1. GenCall scores generally decrease in value the further a sample is from the centre of the cluster to which the data point is associated. Genotypes with a GenCall score less than 0.15 were not assigned genotypes because they were considered too far from the cluster centroid to make reliable genotype calls. They were instead assigned a "no call" for that locus.

2.5.4.2 Output files

No sample or SNP filtering was performed in GenomeStudio. Final report was extracted as a tab-delimited file, with the following data fields: (1) Sample Group, (2) Sample ID, (3) SNP Name, (4) Allele 1-Forward, (5) Allele 2 - Forward, (6) B allele frequency, (7) LogRatio. Along with the Final report file, a map file with information for each SNP was exported.

2.5.5 Preparing PLINK files

From GenomeStudio output files, I prepared the long-format file sets to process in PLINK 2.0. These included the lgen, map, and fam tab-delimited files. The lgen file is a tab-delimited file with the following data fields: (1) Family ID, (2) Individual ID, (3) SNP ID (rs number), (4) Allele 1, (5) Allele 2. The map file is a tab-delimited file with the following data fields: (1) Chromosome number, (2) SNP ID (rs number), (3) Genetic distance, (4) Base pair position. The fam file is a tab-delimited file with the following data fields: (1) Family ID, (2) Individual ID, (3) Paternal ID, (4) Maternal ID, (5) Sex, (6) Phenotype. Paternal and maternal ID were missing and thus entered as "0" for all samples.

I then used the lfile and recode options in PLINK to read the lgen, map, and fam files, and convert them into .ped and .map files for further processing.

2.5.6 GWAS quality control

I performed GWAS quality control in PLINK 2.0. **Figure 2.4** is a flow diagram demonstrating the processing of intensity files in GenomeStudio and quality control steps undertaken to produce a cleaned dataset for further analyses (imputation and association analyses). This section describes the methodology only and results describing success rates are described and discussed in Chapter 6 (section 6.5.1.1).

2.5.6.1 Sample quality

2.5.6.1.1 Sample genotyping efficiency and call rate

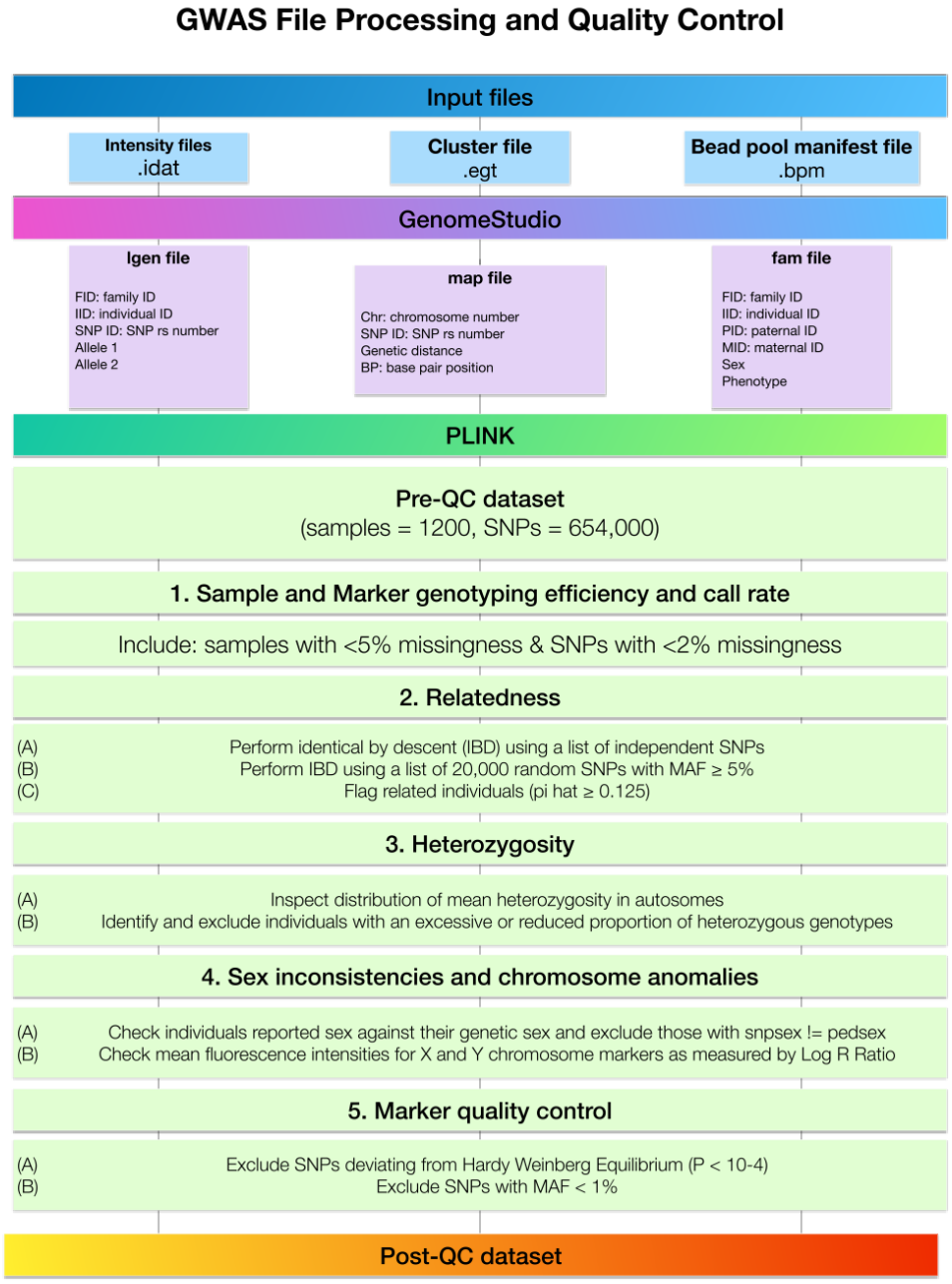
To identify samples with poor DNA quality and/or possible aberrant genotype calling, I used the missing option in PLINK. I eliminated samples with low genotyping efficiency or call rate (those with no genotype called on > 5% of all SNPs) from further analysis using the mind option.

2.5.6.1.2 Sample relatedness

Relatedness indicates how strongly a pair of individuals is genetically related. To avoid biased estimations of standard errors of SNP effect sizes, I adopted appropriate correction methodologies to eliminate or account for relatedness in my sample. Hidden relatedness among individuals or duplicate samples were examined using dense marker data obtained in GWAS through pruning to compute pairwise kinship estimates between every individual in the study using the genome option in PLINK. Pruning is a method used to select a subset of markers that are in approximate linkage equilibrium and was performed in PLINK using the indep-pairwise option (indep-pairwise 50 5 0.1).

Using this set of pruned SNPs, PLINK was used to calculate the proportion of loci where two individuals share zero, one, or two alleles identical by descent (IBD). Individuals with evidence of relatedness were flagged (Pi-HAT ≥ 0.125). To investigate the sensitivity of IBD analysis, I selected 20,000 SNPs at random with a minor allele frequency (MAF) of at least 5%, repeated my IBD analysis, and flagged the individuals who show evidence of relatedness (Pi-HAT ≥ 0.125).

Figure 2.4. Arab-Emirati GWAS genotype calling and quality control. Flow diagram demonstrating the processing of intensity files in GenomeStudio and quality control steps undertaken in PLINK to produce a cleaned dataset for imputation and association analysis.



2.5.6.1.3 Heterozygosity

Another quality control step involves checking for samples' heterozygosity rate, which is the proportion of heterozygous genotypes for each sample. High levels of heterozygosity within an individual might be an indication of low sample quality whereas low levels of heterozygosity may be due to inbreeding. I used the het option in plink to calculate heterozygosity using the set of pruned SNPs (as described in section 2.5.6.1.2). Individuals with excessive or reduced proportion of heterozygotes were excluded (+/-3SD of the mean).

2.5.6.1.4 Sex inconsistencies and chromosomal Anomalies

In order to check for potential sample identity problems that typically result from sample handling errors, I checked the reported sex of all individuals against that predicted by their genetic data using the check-sex option in PLINK. PLINK estimates the inbreeding coefficient over all X-linked markers (FX) and classifies the sample as male if FX < 0.2, and female if FX > 0.8. A value of FX between 0.2 and 0.8, or an apparent sex that disagrees with the putative sex, is flagged as a "problem". If discrepancies were found, reported sex files are reviewed and if problem is not resolved, then sample is excluded from analysis. As a secondary screen, mean fluorescence intensities in X and Y chromosome markers as measured by Logratio (LRR), a normalised measure of overall signal strength, were used to check for sex and sex-chromosome anomalies. A plot of mean X and Y LRR values showed distinct major clusters of points, representing the genetically male (with high LRR for Y, low for X) and female (with high LRR for X, and low for Y).

2.5.6.2 Marker Quality

2.5.6.2.1 Marker genotyping efficiency and call rate

Marker genotyping efficiency (the proportion of samples with a genotype call for each marker) is a good indicator of marker quality. SNP assays that failed on a large number of samples are poor assays and are likely to result in spurious data. SNPs with a call rate <98% were removed from further analysis using the geno option.

2.5.6.2.2 Minor allele frequency

Minor allele frequency (MAF) for each SNP was reported using the freq option in PLINK, and SNPs with MAF lower than 1% were removed using the maf option.

2.5.6.2.3 Hardy-Weinberg Equilibrium

Under Hardy-Weinberg (HWE) assumptions, allele and genotype frequencies can be estimated from one generation to the next. Violation of the HWE law indicates that genotype frequencies are significantly different from expectations. Departure from this equilibrium can be indicative of potential genotyping errors, population stratification, or even actual association to the trait under study. HWE was assessed using the hardy option in PLINK. Chi-square tests for HWE are rejected at P-values less than 10^{-4}.

2.5.7 Genotype imputation

Genotype imputation is the process of inferring missing genotypes in silico using haplotype information from reference samples with genotypes from denser genotyping arrays or sequencing. In this study, I imputed my dataset to the Haplotype Reference Consortium (HRC) panel. The HRC reference panel combines data from 20 different cohort studies[167]. Most of these studies have low-coverage whole-genome sequencing data (4-8x coverage), which consist of samples with predominantly European ancestry. However, the 1000 Genomes Project Phase 3 cohort (one of the cohort studies included in the

HRC panel), includes individuals of diverse ancestry. The HRC panel consists of 64,976 haplotypes at 39,235,157 SNPs with evidence of having a minor allele count ≥ 5. I described pre-imputation checking and post-imputation methodology in Chapter 6 (section 6.4.4). Briefly, I used the pre-imputation checking program developed by the McCarthy group (University Oxford, Oxford)[291], to compare PLINK markers (bim) and frequency (freq) files to the HRC panel. After pre-imputation checking, I uploaded updated files into the Michigan Imputation Server for imputation to the HRC panel[334]. Output files were then cleaned prior to association analyses (Chapter 6, section 6.4.4.2).

2.6 Statistical Methods and Approaches

Detailed statistical methods for all objectives are described in their respective chapters (Chapter 3-6). Statistical analyses were carried out according to a prespecified plan drawn up before data collection and the results were available (Research proposal as presented at the Early Stage Review of my PhD is attached in Appendix 2). All analyses were stratified by ethnicity, and diabetes-status (NGT, prediabetes, T2D). Briefly, continuous data are expressed as mean ± standard deviation (SD) and median and interquartile range (IQR). Categorical data are expressed as percentages and 95% confidence intervals. Data normality was tested by performing the Shapiro Normality test which directed the selection of statistical tests and potential transformations (e.g., log transformation). Wilcoxon signed-rank test and X^2-test were used to compare values of continuous and categorical variables respectively between groups. Relationships with continuous variables were examined using correlation coefficients and linear regression. A logarithmic transformation was used for positively skewed variables prior to regression analyses. Logistic regression was used to examine the relationship between risk factors and T2D, with adjustments for potential confounding variables. Linear regression was used to examine the relationship between risk factors and HOMA-IR and HOMA-B. To calculate the magnitude of T2D risk attributable to a risk factor in EM1, EM2 and UKBB, population attributable risk (PAR%) was calculated using standard

methodology. T2D risk attributable [population attributable risk (PAR)] to obesity or lack of physical activity in the EM2 sample population (case-control study design) was estimated using the following equation:

$$PAR\ (\%) = Pe * [(OR-1)/OR] * 100$$

Pe = proportion of cases exposed

To calculate PAR for UKBB and EM1, I used the following equation:

$$PAR = [Pe(RR-1)*100/Pe(RR-1)+1]$$

Pe = proportion of population exposed

RR = Relative risk (estimated as the OR)

Statistical significance was inferred at a P value of < 0.05. P values were corrected for multiple testing as necessary. For genetic studies, a correction for multiple tests was made, and a $P < 10^{-5}$ was used to infer statistical significance. Given the limited power that my GWAS had with a sample size of < 1,000 (after quality control), a less stringent P value cut-off was used, compared to standard for GWAS ($5x10^{-8}$), to increase discovery of potential SNP-phenotype associations. A Wald test was used to calculate power of the physical activity study. Power calculations for the genetic study were performed in Quanto Version 17.1.0.

Sensitivity analyses performed for each part of this study are presented in the last section of the results in each chapter. These included analyses that evaluate the effect of certain confounding factors, mediators, or the possibility of selection bias.

Table 2.1. List of countries included in NCD-RisC analyses to estimate diabetes and obesity prevalence rates, by super-region and region.			
Super-region	**Region**	**Number of countries**	**Countries**
Sub-Saharan Africa	Central Africa	6	Angola, Central African Republic, Congo, DR Congo, Equatorial Guinea, Gabon
	East Africa	17	Burundi, Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Rwanda, Seychelles, Somalia, Sudan (former), Tanzania, Uganda, Zambia
	Southern Africa	6	Botswana, Lesotho, Namibia, South Africa, Swaziland, Zimbabwe
	West Africa	19	Benin, Burkina Faso, Cabo Verde, Cameroon, Chad, Cote d'Ivoire, Gambia, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Sao Tome and Principe, Senegal, Sierra Leone, Togo
Central Asia, Middle East and North Africa	Central Asia	9	Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Mongolia, Tajikistan, Turkmenistan, Uzbekistan
	Middle East and North Africa	19	Algeria, Bahrain, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Occupied Palestinian Territory, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, Tunisia, Turkey, United Arab Emirates, Yemen
South Asia	South Asia	6	Afghanistan, Bangladesh, Bhutan, India, Nepal, Pakistan
East and SouthEast Asia	East Asia	4	China, China (Hong Kong SAR), North Korea, Taiwan
	Southeast Asia	12	Brunei Darussalam, Cambodia, Indonesia, Lao PDR, Malaysia, Maldives, Myanmar, Philippines, Sri Lanka, Thailand, Timor-Leste, Viet Nam
Oceania	Polynesia and Micronesia	13	American Samoa, Cook Islands, French Polynesia, Kiribati, Marshall Islands, Micronesia (Federated States of), Nauru, Niue, Palau, Samoa, Tokelau, Tonga, Tuvalu
	Melanesia	4	Fiji, Papua New Guinea, Solomon Islands, Vanuatu
High-income Asia Pacific	High-income Asia Pacific	3	Japan, Singapore, South Korea

Latin America and Caribbean	Andean Latin America	3	Bolivia, Ecuador, Peru
	Caribbean	18	Antigua and Barbuda, Bahamas, Barbados, Belize, Bermuda, Cuba, Dominica, Dominican Republic, Grenada, Guyana, Haiti, Jamaica, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago
	Central Latin America	9	Colombia, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Venezuela
	Southern Latin America	5	Argentina, Brazil, Chile, Paraguay, Uruguay
High-income Western countries	High-income English-speaking countries	6	Australia, Canada, Ireland, New Zealand, United Kingdom, United States of America
	North Western Europe	12	Austria, Belgium, Denmark, Finland, Germany, Greenland, Iceland, Luxembourg, Netherlands, Norway, Sweden, Switzerland
	South Western Europe	9	Andorra, Cyprus, France, Greece, Israel, Italy, Malta, Portugal, Spain
Central and Eastern Europe	Central Europe	13	Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Hungary, Macedonia (TFYR), Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia
	Eastern Europe	7	Belarus, Estonia, Latvia, Lithuania, Moldova, Russian Federation, Ukraine

Table 2.2. List of countries included in the analysis performed by Guthold et al.[105], by super-region.

Super-region	Number of countries	Countries
Central Asia, Middle East and North Africa	23	Algeria, Armenia, Egypt, Georgia, Iran (Islamic Republic of), Iraq, Jordan, Kazakhstan, Kuwait, Kyrgyzstan, Lebanon, Libya, Mongolia, Morocco, Oman, Qatar, Saudi Arabia, State of Palestine, Tajikistan, Tunisia, Turkey, United Arab Emirates, Uzbekistan
Central and Eastern Europe	17	Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Republic of Moldova, Romania, Russian Federation, Serbia, Slovakia, Slovenia, Ukraine
East and SouthEast Asia	13	Brunei Darussalam, Cambodia, China, Indonesia, Lao People's Democratic Republic, Malaysia, Maldives, Myanmar, Philippines, Sri Lanka, Thailand, Timor-Leste, Viet Nam
High-income Asia Pacific	3	Japan, Republic of Korea, Singapore

High-income Western countries	24	Andorra, Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States of America
Latin America and Caribbean	25	Argentina, Bahamas, Barbados, Bermuda, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Grenada, Guatemala, Jamaica, Mexico, Paraguay, Saint Kitts and Nevis, Saint Lucia, Suriname, Trinidad and Tobago, Uruguay, Venezuela (Bolivarian Republic of)
Oceania	17	American Samoa, Cook Islands, Fiji, French Polynesia, Kiribati, Marshall Islands, Micronesia (Federated States of), Nauru, Niue, Palau, Papua New Guinea, Samoa, Solomon Islands, Tokelau, Tonga, Tuvalu, Vanuatu
South Asia	5	Bangladesh, Bhutan, India, Nepal, Pakistan
Sub-Saharan Africa	41	Benin, Botswana, Burkina Faso, Cabo Verde, Cameroon, Central African Republic, Chad, Comoros, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, São Tomé and Principe, Senegal, Seychelles, Sierra Leone, South Africa, Swaziland, Togo, Uganda, United Republic of Tanzania, Zambia, Zimbabwe

Table 2.3. ICLDC Laboratory procedures for biochemical analysis.

	Method	**Analyser**
Glycaemic profile		
FBG	Enzymatic reference method with hexokinase	Cobas 8000
Insulin	Electrochemiluminescence Immunoassay (ECLIA)	Cobas C6000
HbA1c	High Performance Liquid Chromatography	D-100
Lipid profile		
Total cholesterol	Enzymatic Colorimetric	Cobas 8000

HDL	Homogeneous Enzymatic Colorimetric Test	Cobas C701/702
LDL	Homogeneous Enzymatic Colorimetric Test	Cobas C8000
Triglycerides	Homogeneous Enzymatic Colorimetric Test	Cobas C8000
Liver and Kidney function		
AST	Acc. to IFCC w/o pyridoxal phosphate activation	Cobas 8000
ALP	Colorimetric Assay	Cobas 8000
Albumin	Colorimetric Assay	Cobas C701/702
Bilirubin	Colorimetric Assay	Cobas C701/702
Creatinine	Kinetic Colorimetric Assay	Cobas C701/702
Potassium	Ion-Selective Electrode (ISE)	Cobas 8000
Urea	Kinetic Test with Urease and Glutamate Dehydrogenase	Cobas C701/702
Blood count		
Red blood cell count	Radio-frequency/Direct Current Detection, Hydrodynamic Focusing, Flow Cytometry & SLS-Haemoglobin	Sysmex XN-2000
White blood cell count	Radio-frequency/Direct Current Detection, Hydrodynamic Focusing, Flow Cytometry & SLS-Haemoglobin	
Full blood count	Radio-frequency/Direct Current Detection, Hydrodynamic Focusing, Flow Cytometry & SLS-Haemoglobin	

Abbreviations; ICLDC, Imperial College London Diabetes Centre; FBG, fasting blood (plasma) glucose; HbA1c, haemoglobin A1c; LDL, low-density lipoprotein; HDL, high-density lipoprotein; AST, aspartate transaminase; ALP, alkaline phosphatase

Table 2.4. Global Physical Activity Questionnaire guidelines to categorise distribution of activity participation in the three different domains: work, leisure, and transport.	
Activity domain	**Activity in ≥ 1 day in a typical week**
Work/household	**Moderate-intensity:** activity that causes small increases in breathing or heart rate for ≥10 minutes continuously
	Vigorous-intensity: activity that causes large increases in breathing or heart rate for ≥10 minutes continuously
Leisure	**Moderate-intensity:** sports or fitness activity that causes small increases in breathing or heart rate for ≥10 minutes continuously
	Vigorous-intensity: sports or fitness activity that causes large increases in breathing or heart rate for ≥10 minutes continuously
Transport	Walk or use bicycle to go to and from places for ≥10 minutes continuously

Table 2.5. Global Physical Activity Questionnaire guidelines to categorise total physical activity levels.	
Total physical activity (TPA) = Sum of total MET minutes of activity computed for work, transport, and leisure settings.	
Level of total physical activity	**Cut off criteria**
High (HPA)	(1) Vigorous activity on ≥3 days, achieving 1500 MET-minute/week
	OR
	(2) Walking, moderate, or vigorous activity on ≥7 days achieving ≥3000 MET-minute/week
Moderate (MPA)	(1) Vigorous activity on ≥3 days for ≥20 minutes
	OR
	(2) Moderate activity on ≥5 days for ≥30 minutes
	OR
	(3) Any combination of walking, moderate, or vigorous activities on ≥5 days, achieving ≥600 MET-minutes/week
Low (LPA)	Not meeting any of the above criteria

Chapter Three: Secondary analyses of global datasets: do obesity and physical activity explain variation in diabetes risk across populations?

In this chapter, I describe the use of publicly-available datasets on diabetes, obesity and physical inactivity prevalence rates and trends, as estimated by the NCD-RisC and the WHO. I present and discuss the relationship of obesity and physical inactivity with diabetes at a global, regional, and country level. Together these results help justify the hypothesis of my study and highlight the importance of conducting it. The work presented in this chapter forms part of my publication (Published article attached in Appendix 3)[292]. Supplementary figures and tables can be accessed through Zenodo using this link: https://zenodo.org/record/7825396#.ZDfrqC0Rro4.

3.1 OVERVIEW

I used estimates from the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) and the World Health Organisation (WHO) on rates and trends of diabetes (1975-2014), obesity (1975-2016), and insufficient physical activity (IPA) (2016) in 200 countries and 9 regions. I investigated region-specific correlation of obesity and IPA with diabetes, and their joint effects on diabetes risk. My results highlighted the rise of obesity and diabetes rates at a global level, albeit to varied extent between

regions. Similarly, the association of IPA with diabetes was not homogeneous across populations. The overall joint effects of obesity and IPA highlighted the larger impact of obesity on diabetes risk across populations. These data indicate that different populations around the world, who may be experiencing similar environmental shifts, are apparently differentially susceptible to diabetes risk. Data comparisons from the UAE and the UK demonstrate a representative example of this overall observation.

3.2 AIMS

To use publicly-available estimates by the NCD-RisC and the WHO to unravel the relative contribution of obesity and physical inactivity towards diabetes risk across different populations around the globe.

3.2.1 Objectives

1. To determine region-specific changes in diabetes and obesity rates in the last 34 years (1980-2014) and assess differences in these trends between regions.
2. To investigate region-specific correlation of obesity and insufficient physical activity with diabetes and evaluate the strength of this correlation across regions.
3. To assess the joint effects of obesity and insufficient physical activity on diabetes risk across different countries and regions.
4. To utilise country-specific data to assess the contribution of obesity and insufficient physical activity towards diabetes risk in the population of the UAE, in comparison to European populations (UK).

3.3 INTRODUCTION

T2D rates vary significantly between countries, geographic regions, and ethnic groups[3]. Understanding differences in T2D risk between ethnic groups has been challenging. Although the differential distribution of established behavioural and environmental exposures amongst geographic regions and ethnic populations can account for part of this variation, it remains largely unexplained[293]. As discussed previously (in Chapter 1, section 1.5.2), part of this ethnic disparity in T2D is thought to be due to differences in glycaemic control mechanisms, or preferential distribution of fat storage. For example, Indian Asians are more insulin resistant[207], and exhibit greater levels of visceral adipose tissue (VAT) and ectopic fat[216], and larger adipocytes[212], at similar levels of overall adiposity compared to Europeans. Whereas, T2D pathogenesis among African Americans appears to be largely driven by β-cell dysfunction, rather than insulin resistance[216], and is characterised with favourable body fat distribution qualities, as compared to other ethnic groups[294].

Several trans-ethnic studies were directed at elucidation of the impact of ethnicity on the association of age, BMI, and physical inactivity with T2D. Overall, their findings confirmed large ethnic, regional and sex-dependent differences in the age-adjusted prevalence of T2D, but also showed that ethnicity-specific variations were not fully explained by differences in BMI between ethnic groups[293]. As with BMI, numerous cross-sectional studies have demonstrated that differences in T2D rates between some ethnic groups cannot be explained by differences in the levels of physical activity levels that their populations undertake[295, 296]. Despite these efforts, there is still an under-representation of various ethnic populations in existing studies, particularly those in the AP and the Middle East. The UAE has continuously been estimated to be one of the top 20 countries with the highest rates of diabetes worldwide[3, 205]. Yet, studies investigating the factors contributing to this increased risk remain limited.

Networks and organisations like the NCD-RisC and the WHO offer a good opportunity to fill this research gap. Together they provide worldwide rigorous and timely data on risk factors for non-communicable diseases, including diabetes, obesity and physical inactivity. In 2016 and 2017, NCD-RisC provided some of the lengthiest and most complete estimates of worldwide adult diabetes rates, and trends in mean BMI and obesity rates for all countries in the world with the longest observation period[205, 260]. NCD-RisC diabetes estimates presented in this chapter are based on all diabetes cases worldwide, including people with T2D, although, T2D accounts for majority of all cases (75-85%). In addition to NCD-RisC data, in 2018, a group from the WHO[105] provided the some of the most complete description of global, regional, and country levels of insufficient physical activity, and presented regional and global trends over time.

Here, I performed a secondary-analysis that compiles publicly-available estimates by the NCD-RisC and the WHO on diabetes, obesity, together with insufficient physical activity rates and trends in adults. With this analysis, I aimed to unravel the relative contribution of obesity and physical inactivity to diabetes risk across different populations worldwide.

3.4 MATERIALS AND METHODS

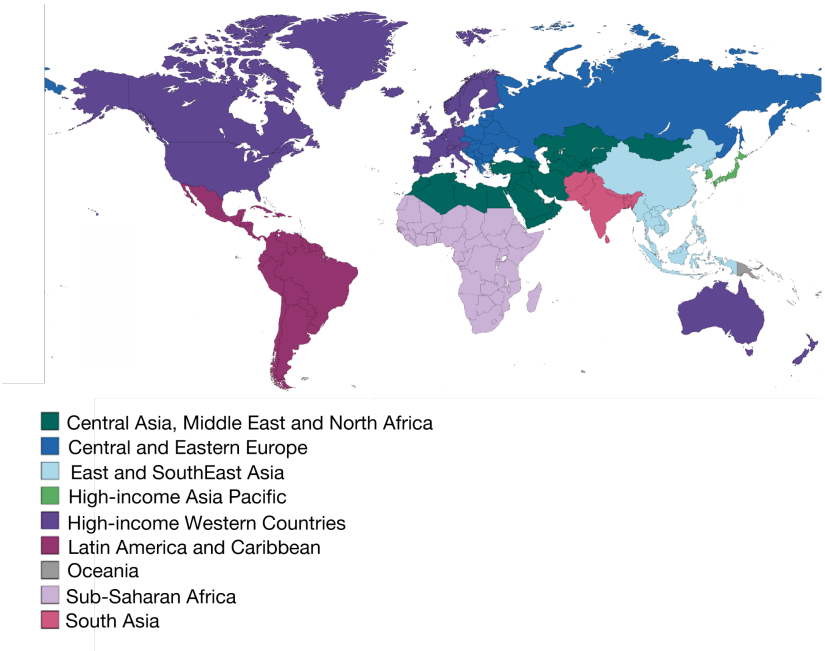
3.4.1 Data sources, data retrieval and processing

Data sources specification and data retrieval process were fully described in Chapter 2 section 2.1. Briefly, I reanalysed publicly-available estimates of diabetes, obesity, insufficient physical activity rates by the NCD-RisC and the WHO in 9 super-regions (**Figure 3.1**). Super-regions are referred to as "regions" throughout this chapter, and they include Central Asia, Middle East and North Africa, Central and

Eastern Europe, East and Southeast Asia, High income Asia Pacific, High income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa.

Estimates of diabetes, obesity, and insufficient physical activity rates were available for 200 (diabetes and obesity) and 168 (insufficient physical activity) countries, respectively. Region-, country-, and sex-specific estimates of, (1) diabetes prevalence rates from 1980 to 2014, and (2) obesity prevalence rates from 1975 to 2016, were downloaded from www.ncdrisc.org. Region-, country-, and sex-specific estimates of insufficient physical activity prevalence rates in 2016 were retrieved from a recent publication by a group from the WHO (35). All estimates were derived for males and females aged ≥ 18 years.

Figure 3.1. World map demonstrating the 9 regions as defined by the Non-Communicable Disease Risk Collaboration group (NCD-RisC). The NCD-RisC categorises 200 countries into 9 regions, including, (1) Central and Eastern Europe, (2) Central Asia Middle East and North Africa, (3) East and SouthEast Asia, (4) High-income Asia Pacific, (5) High-income Western countries, (6) Latin America and Caribbean, (7) Oceania, (8) South Asia, and (9) Sub-Saharan Africa.



3.4.2 Definitions

The NCD-RisC and the WHO organised countries into 9 super-regions (**Chapter 2, Table 2.1, 2.2**), mostly based on geography and national income. Diabetes was defined as fasting plasma glucose of 7.0 mmol/L or higher, history of diagnosis with diabetes, or use of insulin or oral hypoglycaemic drugs[205]. Obesity was defined as BMI ≥ 30 Kg/m^2. Insufficient physical activity in adults was defined as not meeting the WHO recommendations on physical activity for health—i.e., at least 150 min of moderate-intensity, or 75 min of vigorous-intensity physical activity per week, or any equivalent combination of the two. Prevalence of insufficient physical activity was reported by age and sex.

3.4.2 Statistical methods

Region- and sex-specific differences in diabetes prevalence rates and trends were assessed using scatter plots to compare the distribution of diabetes prevalence rates in males against that in females across 200 countries. This was done separately for 1980 and 2014, and for the change in diabetes prevalence throughout the 34 years. These analyses were repeated to assess region- and sex-specific differences in obesity prevalence rates in 1975 and 2016, and the change in obesity prevalence throughout the 41 years (1975-2016) studied.

Change in diabetes prevalence = Diabetes prevalence in 2014 - Diabetes prevalence in 1980

Change in obesity prevalence = Obesity prevalence in 2016 - Obesity prevalence in 1975

Region- and sex-specific correlation analyses of obesity with diabetes were based on obesity and diabetes prevalence rates in 200 countries in years 2016 and 2014, respectively. Region- and sex-specific correlation analyses of insufficient physical activity with diabetes were based on insufficient physical activity and diabetes prevalence rates in 163 countries in year 2016 and 2014, respectively. Diabetes and

obesity prevalence rates are presented as percentages with upper and lower limits. Insufficient physical activity rates are presented as percentages only (upper and lower limits were not available). Shapiro-Wilk tests were performed to test for the normality of diabetes, obesity, and insufficient physical activity prevalence rates distribution (**Supplementary Table 3.1**). Given that rates of obesity, insufficient physical activity and/or diabetes were not normally distributed in some regions, correlation was tested using both Pearson and Spearman tests. Significance was set at $p < 0.05$. Country- and sex-specific joint effects of obesity and insufficient physical activity prevalence rates with diabetes were based on estimates in 163 countries in years 2014 (diabetes) and 2016 (obesity and insufficient physical activity). Estimates of obesity, insufficient physical activity, and diabetes rates, were each split into tertiles. For each of obesity and insufficient physical activity rates, countries that fell in the first, second, and third tertiles, were coded as low, moderate, and high, respectively. Diabetes risk was defined based on the tertile split; countries that fell in the first tertile (with the lowest rates of diabetes) were defined to have low diabetes risk, and those that fell in the second and third tertiles, were defined to have moderate and high risk, respectively. All analyses were performed in Prism (Version 8).

3.5 RESULTS

Figures 3.2 and 3.3 are world maps demonstrating rates of diabetes (2014), obesity (2016), and insufficient physical activity (2016) in the 9 regions studied, colour-coded according to their global ranking, as estimated by the NCD-RisC[205, 273] and the WHO[105], in males and females, respectively.

3.5.1 Region- and sex-specific trends in diabetes rates from 1980 to 2014

Overall, in the 34 years studied (1980-2014), diabetes rates increased in all 9 regions. Based on my secondary-analysis of NCD-RisC data, in 1980, prevalence rates of diabetes were very similar across all regions, but as years passed, this regional variation increased. In 2014, the age-standardised diabetes

prevalence was highest in Oceania among both males [Prevalence (upper limit-lower limit): 15.8% (9.1-24.8)] and females [15.7% (9.3-24.5)] (**Figure 3.4**). In addition, Oceania along with Central Asia, Middle East and North Africa region, were estimated to have the highest rates of diabetes across all years, particularly after 1985 (**Figure 3.4**).

Figure 3.2. World maps demonstrating region-specific rates of (A) diabetes, (B) obesity, and (C) Insufficient physical activity in males. Age-standardised diabetes and obesity prevalence rates as estimated by Non-Communicable Disease Risk Collaboration (NCD-RisC) group, in 2014 and 2016, respectively. Insufficient physical activity (IPA) rates were estimated by the World Health Organisation (WHO) in 2016. Based on NCD-RisC group definition, estimates were reported for 9 regions including, Central and Eastern Europe, Central Asia Middle East, and North Africa, East and SouthEast Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. The colour fill of each region reflects the rank of prevalence rates with the darkest and the lightest indicating the highest and the lowest rates at a global level, respectively.

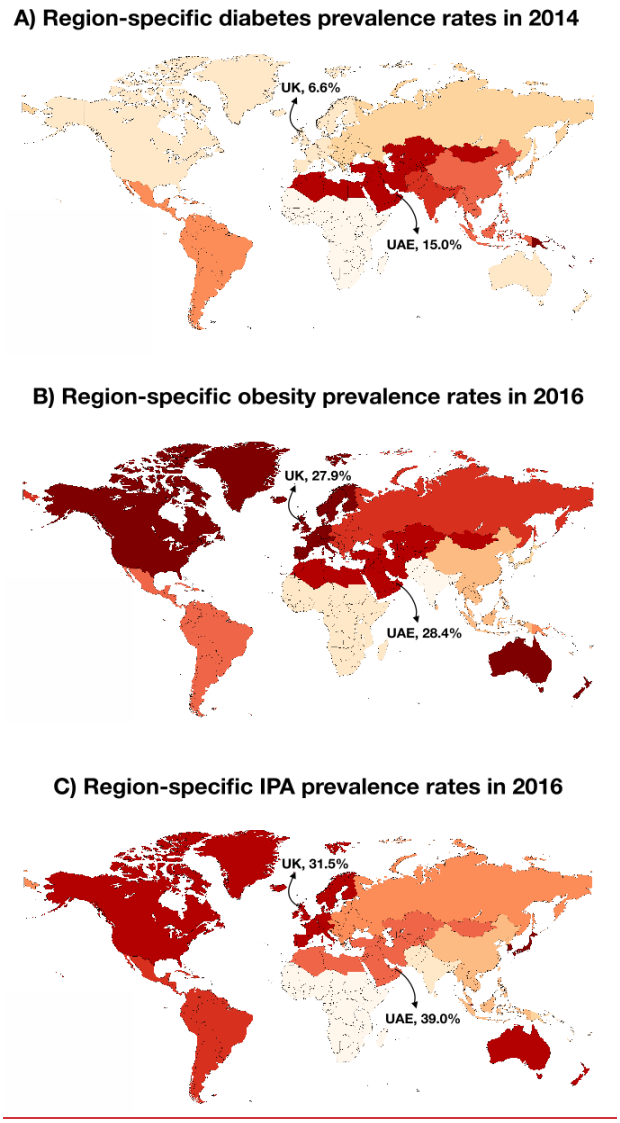
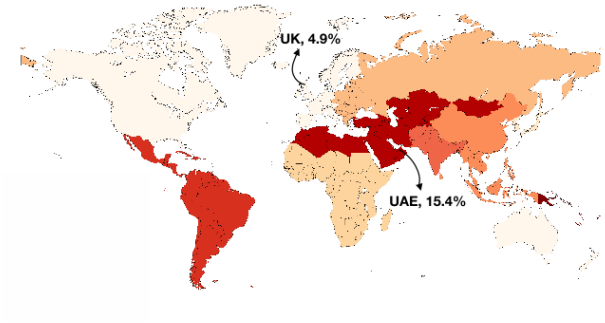
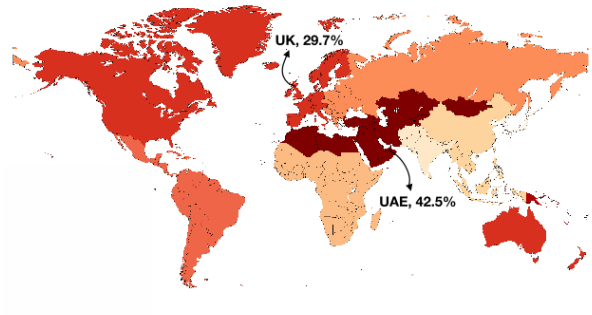


Figure 3.3. World maps demonstrating region-specific rates of (A) diabetes, (B) obesity, and (C) Insufficient physical activity in females. Age-standardised diabetes and obesity prevalence rates as estimated by Non-Communicable Disease Risk Collaboration (NCD-RisC) group, in 2014 and 2016, respectively. Insufficient physical activity (IPA) rates were estimated by the World Health Organisation (WHO) in 2016. Based on NCD-RisC group definition, estimates were reported for 9 regions including, Central and Eastern Europe, Central Asia Middle East, and North Africa, East and SouthEast Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. The colour fill of each region reflects the rank of prevalence rates with the darkest and the lightest indicating the highest and the lowest rates at a global level, respectively.

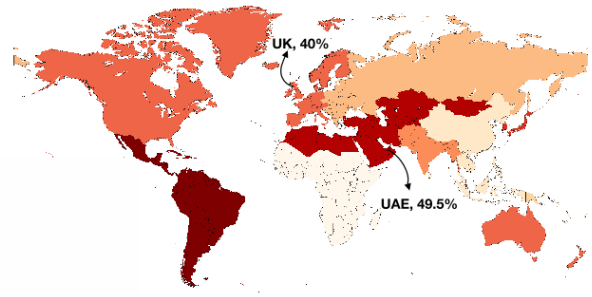
A) Region-specific diabetes prevalence rates in 2014



B) Region-specific obesity prevalence rates in 2016



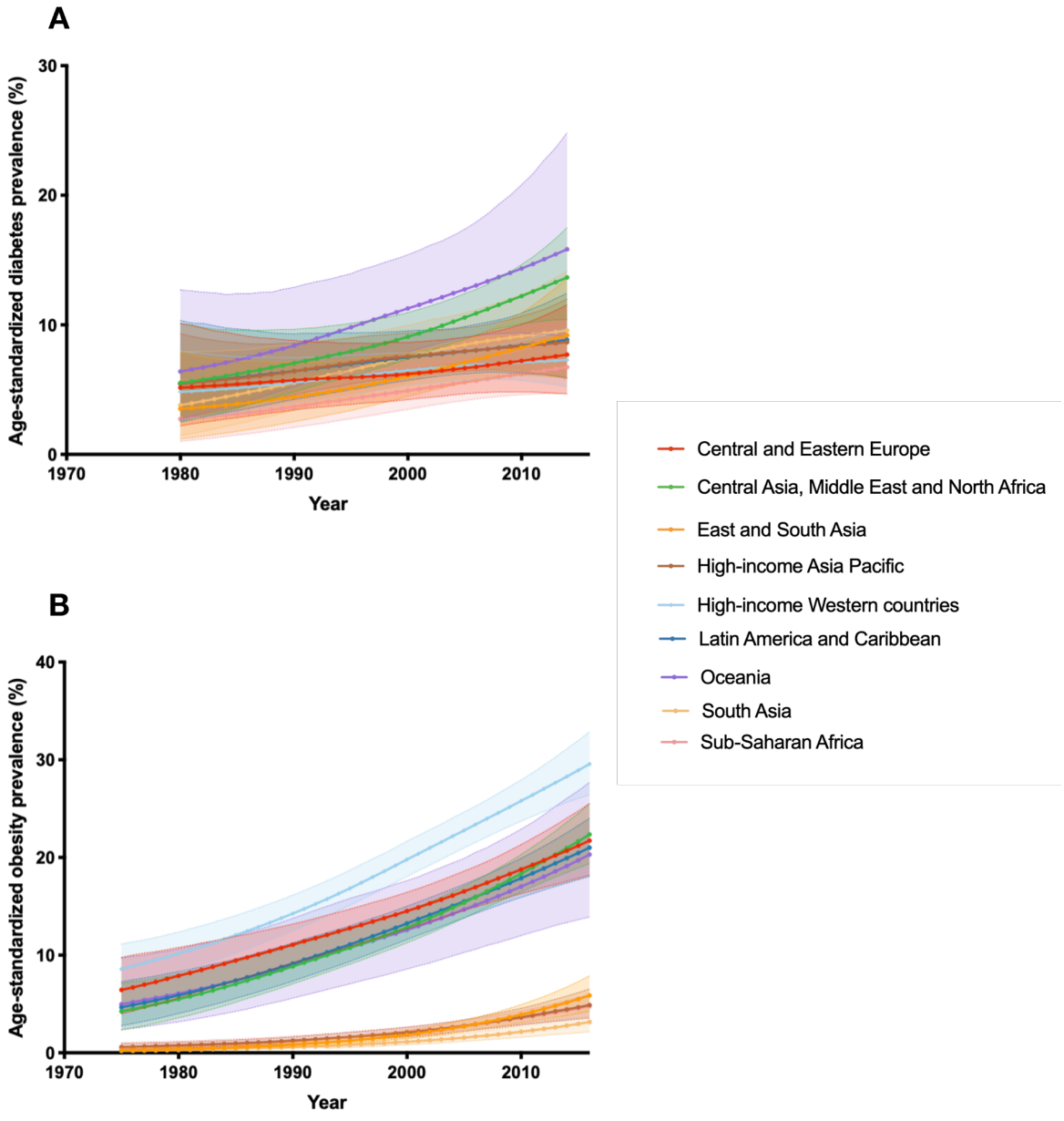
C) Region-specific IPA prevalence rates in 2016



In 2014, diabetes rates in these regions reached a prevalence (upper limit-lower limit) of 13.7% (10.4-17.5) and 15.8% (9.1-24.8) in males and 14.4% (11.2-18.3) and 15.7% (9.3-24.5) in females, respectively (**Figure 3.4**). On the other hand, the Sub-Saharan Africa region had the lowest rates of diabetes between 1980 and 1995 in both sexes. Between 1995 and 2014, however, Sub-Saharan Africa continued to have the lowest rates of diabetes among males only. In both 1980 and 2014, 4 out of the top 5 countries with the highest rates of diabetes in the world were in Oceania. These included Nauru [in 2014: 30.1% (17.9-43.9) in males and 28.4% (17.0-41.9) in females], American Samoa [in 2014: 30.8% (19.1-43.7) in males and 32.9% (21.3-46.6) in females], Palau [in 2014: 24.8% (15.0-36.1) in males and 21.6% (12.8-32.4) in females], and the Cook Islands [in 2014: 28.3% (17.6-40.4) in males and 26.7% (16.8-38.2) in females].

Sex-specific differences in diabetes rates increased with the years; in 1980, overall diabetes rates were almost identical between sexes (**Figure 3.6 A**). By 2014, diabetes rates were higher around the world than in the 1980s, rising more rapidly in some countries/regions/sex than others. Overall, increases in diabetes prevalence were higher in males (mean increase of 5.4%) compared to females (mean increase of 4.2%) **(Figure 3.6 C**). In both males and females, during the 34 years studied (1980-2014), the largest increase in diabetes prevalence rates were observed in Oceania (males: 12.9% and females: 12.1%) and Central Asia, Middle East and North Africa (males: 7.6% and females: 6.9%), whereas the High-Income Western countries region have undergone the least increase in diabetes rates during this time period (males: 1.8% and females: 0.01%). There were sex differences between populations too: rates were higher in males, in High-income Western countries, and higher in females among countries in Latin America and the Caribbean **(Figure 3.6 B**).

Figure 3.4. Region-specific trends of (A) diabetes (1980-2014) and (B) obesity (1975-2016) in males. Age-standardised diabetes and obesity prevalence rates as estimated by Non-Communicable Disease Risk Collaboration (NCD-RisC) group, for each year from 1980 to 2014, and 1975 to 2016, respectively. NCD-RisC group reported diabetes and obesity prevalence estimates in 200 countries, which were categorised into 9 regions, including, Central and Eastern Europe, Central Asia Middle East, and North Africa, East and SouthEast Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]



3.5.2 Region- and sex-specific trends in obesity rates from 1975 to 2016

Like diabetes, obesity rates also showed an increasing trend across all regions in the 41 years studied (1975-2016). In males, between 1975 and 2016, obesity rates were highest in the high-income Western countries region, and lowest in South Asia, particularly post-2000 (**Figure 3.5 C**). In females, obesity rates were highest in Central and Eastern Europe in the first 15 years, but this increasing trend slowed after 1990 (**Figure 3.5 D**). In females, between 1990 and 2016, obesity rates were highest in the Central Asia Middle East and North Africa region, reaching a prevalence rate of 35.2% (31.9-38.7) in 2016. While in males, the highest obesity rates during this period were observed in the High-income Western Countries region, reaching a prevalence of 27.8% (27.5-28.1) in 2016 **(Figure 3.5 D**). Overall, females had higher rates of obesity compared to males across all regions both in 1975 and 2016 (**Figure 3.7 A, B**). In both males and females, during the 41 years (1975-2016), the largest increase in obesity prevalence rates was observed in Oceania (males: 26.2% and females: 24.8%), whereas the smallest increases were observed in South Asia (males: 3.5% and females: 6.7%) and High-income Asia Pacific regions (males: 4.4% and females: 2.8%) (**Figure 3.7 C**). The top 5 countries with the highest rates of obesity in 2016 were all in the Oceania region, including Nauru [in 2016: 59.9% (51.8-67.7) in males and 64.8% (57.0-72.1) in females], American Samoa [in 2016: 58.8% (50.3-66.6) in males and 65.3% (57.6-72.3) in females], and the Cook Islands [in 2016: 54.0% (45.5-61.9) in males and 60.9% (52.8-68.4) in females]. Those five countries/territories (Nauru, American Samoa, Cook islands, Palau, and Marshall islands) were also among the top 10 countries with the highest obesity rates in 1975.

Figure 3.5. Region-specific trends of (A) diabetes (1980-2014) and (B) obesity (1975-2016) in females. Age-standardised diabetes and obesity prevalence rates as estimated by Non-Communicable Disease Risk Collaboration (NCD-RisC) group, for each year from 1980 to 2014, and 1975 to 2016, respectively. NCD-RisC group reported diabetes and obesity prevalence estimates in 200 countries, which were categorised into 9 regions, including, Central and Eastern Europe, Central Asia Middle East, and North Africa, East and SouthEast Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]

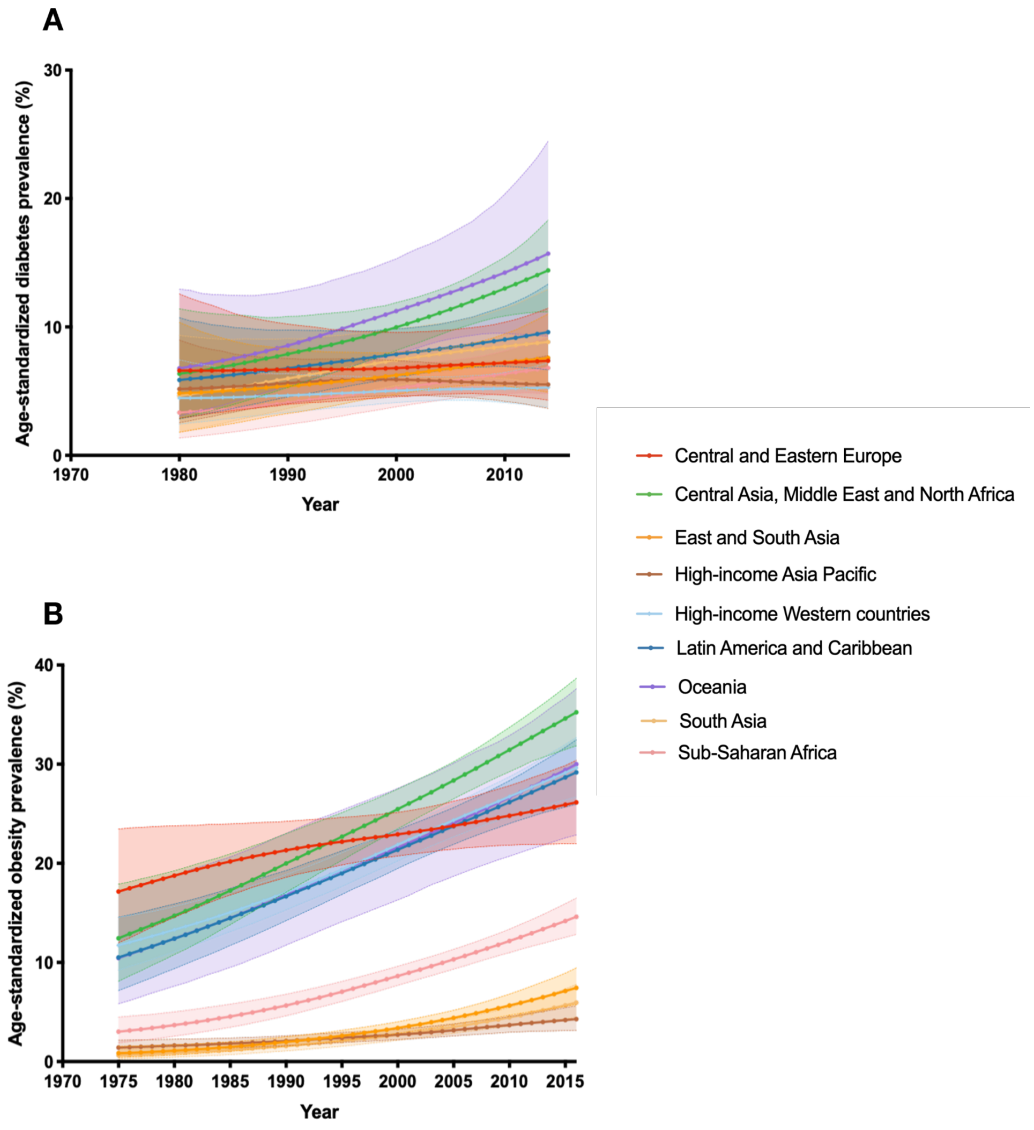


Figure 3.6. Comparison of age-standardised diabetes prevalence in (A) 1980, (B) 2014, and (C) change in diabetes prevalence between 1980 and 2014 by sex. Data presented is based on Non-Communicable Disease Risk Collaboration (NCD-RisC) group estimates of diabetes prevalence rates in 1980 and 2014, across 200 countries. Scattergraphs A and B represent the age-standardised diabetes prevalence rates in 200 countries, in males (x-axis) against females (y-axis), in 1980 and 2014, respectively. Scattergraph C represents the change in diabetes rates during the 34 years (1980-2014) in 200 countries, in males (x-axis) against females (y-axis). The 200 countries were categorised into 9 regions and colour coded in the figure, which include: Central and Eastern Europe, Central Asia Middle East, and North Africa, East and SouthEast Asia, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]

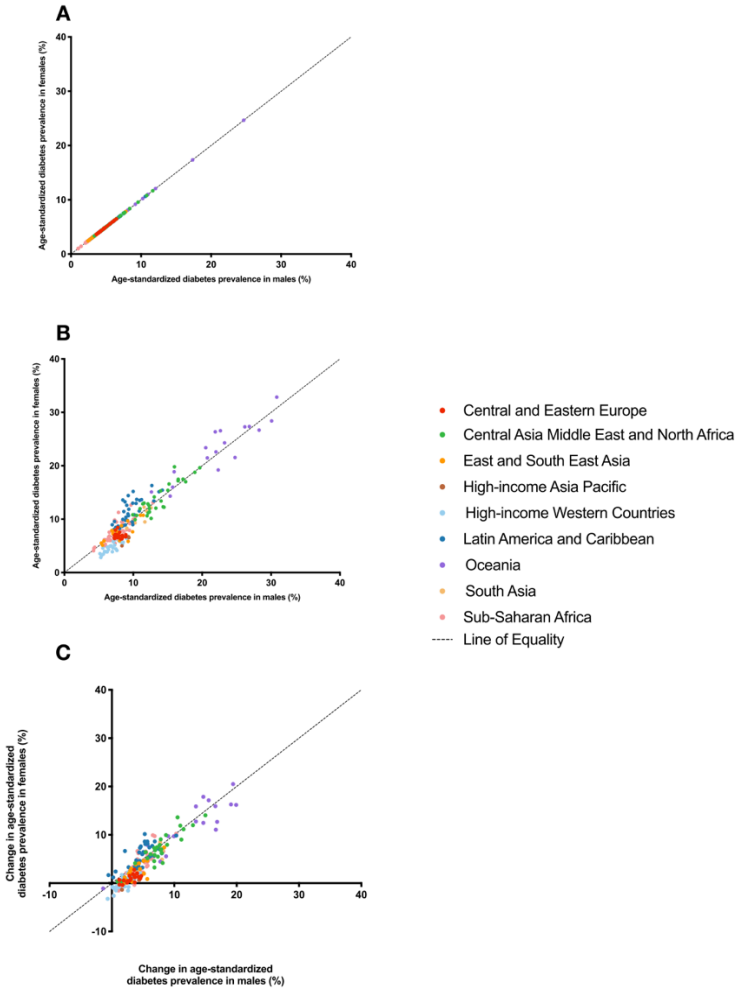
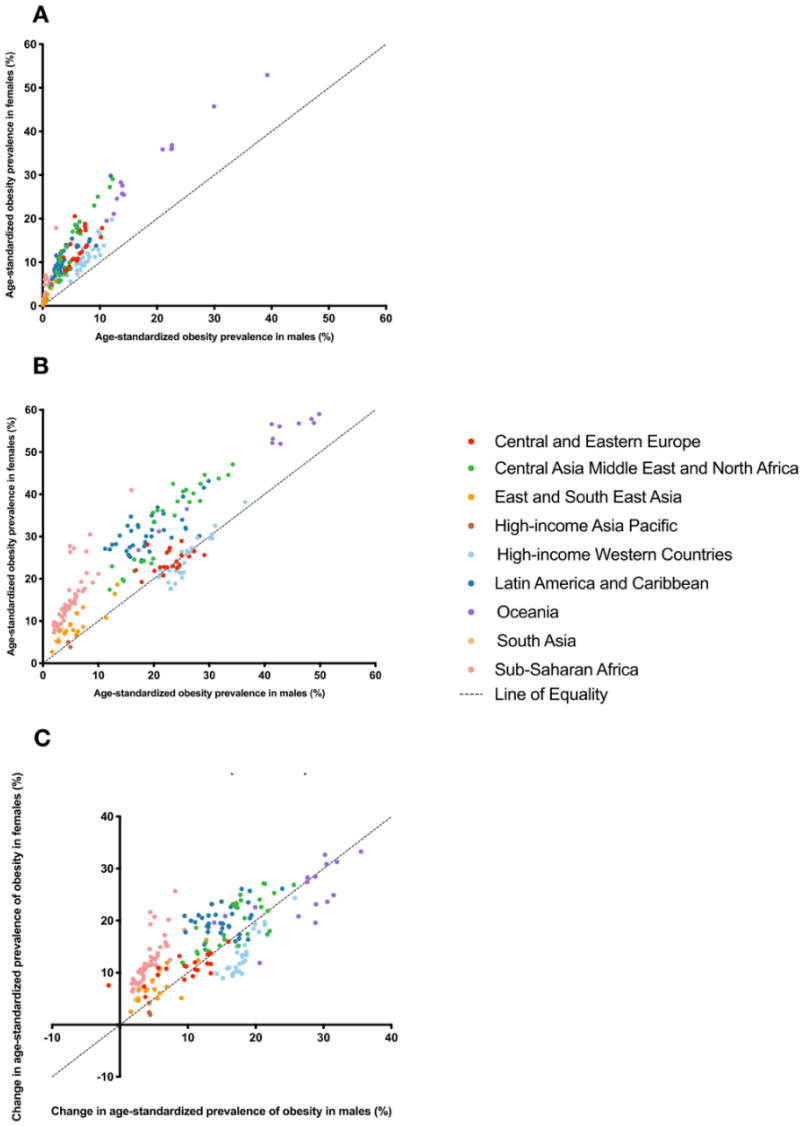


Figure 3.7 Comparison of age-standardised obesity prevalence rates in (A) 1975 and (B) 2016, and (C) change in obesity prevalence between 1975 and 2016 by sex. Data presented is based on Non-Communicable Disease Risk Collaboration (NCD-RisC) group estimates of diabetes prevalence rates in 1975 and 2016, across 200 countries. Scattergraphs A and B represent the age-standardised obesity prevalence rates in 200 countries, in males (x-axis) against females (y-axis), in 1975 and 2016, respectively. Scattergraph C represents the change in obesity prevalence rates during the 41 years (1975-2016) in 200 countries, in males (x-axis) against females (y-axis). The 200 countries were categorised into 9 regions and colour coded in the figure, which include: Central and Eastern Europe, Central Asia Middle East, and North Africa, East and SouthEast Asia, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]



3.5.3 Region- and sex-specific estimates of insufficient physical activity rates in 2016

In 2016, according to estimates by the WHO, the top 3 countries with the highest age-adjusted insufficient physical activity rates were (1) High-income Asia Pacific (34.1% in males and 37.8% in females), (2) Latin America and Caribbean regions (29.2% in males and 40.9% in females), and (3) High-income Western Countries (29.8% in males and 36.8% in females) (**Figure 3.2 and 3.3 C**). On the other hand, the Sub-Saharan Africa region was estimated to have the lowest rates of insufficient physical activity in 2016 of 18.0% in males and 25.6% in females **(Figure 3.2 and 3.3 C**). Three out of the top 5 countries with the highest rates of insufficient physical activity rates were in the Central Asia, Middle East and North Africa region, including Kuwait (61.3% in males and 74.6% in females) (highest in the world), Saudi Arabia (44.9% in males and 53.1% in females), and Iraq (39.5% in males and 64.6% in females). On the other hand, 3 out of the 5 countries with the lowest rates of insufficient physical activity rates were in the Sub-Saharan Africa region, including Uganda (5.2% in males and 5.8% in females) lowest in the world), Mozambique (5.0% in males and 6.0% in females), and Lesotho (6.5% in males and 6.3% in females).

3.5.4 Contribution of obesity towards differences in diabetes risk across populations

To assess region-specific correlation of obesity with diabetes, I used NCD-RisC estimates of diabetes (2014) and obesity (2016) prevalence rates for 200 countries and 9 regions. In males, obesity rates were significantly correlated with diabetes rates in 4 out of the 9 regions, with the strongest correlation in Central Asia, Middle East and North Africa ($r = 0.83$, $p < 0.001$), followed by Oceania ($r = 0.80$, $p < 0.001$), East and SouthEast Asia ($r = 0.74$, $p = 0.002$), and Sub-Saharan Africa ($r = 0.67$, $p < 0.001$) (**Table 3.1**).

In females, a significant correlation was observed in 6 regions. As in males, the strongest correlation was observed in Central Asia Middle East and North Africa ($r = 0.85$, $p < 0.001$) (**Table 3.1**). On the other hand, unlike males, the correlation of obesity with diabetes in High-income Western countries region was stronger and significant in females ($r = 0.80$, $p < 0.001$) (**Table 3.1**).

In line with these observations, most countries with obesity prevalence rates lower than the worldwide mean (10.8% in males, 14.9% in females), also had lower diabetes prevalence rates compared to the worldwide mean (males, 9%; females, 7.9%), and vice versa. This was observed in males and females (**Figure 3.8**). Despite this well-established association, several countries, particularly in Latin America and Caribbean and Central and Eastern Europe regions yield surprising data: they have some of the highest rates of obesity, but still have some of the lowest rates of diabetes.

When comparing diabetes and obesity rates trends between 1980 and 2016 (**Figure 3.9**), the increasing obesity trends in some regions, like Central Asia Middle East and North Africa reflect the increasing diabetes trend in the region. In other regions, however, including Central Eastern Europe and High-income Western countries regions, diabetes rates have been consistent during that time period, despite the increasing rates of obesity in their populations.

3.5.5 Contribution of insufficient physical activity towards explaining differences in diabetes risk across populations

To assess the correlation of insufficient physical activity with diabetes across regions and countries, I used estimates from 2016 of insufficient physical activity rates in 163 countries and 9 regions, as reported by the WHO. Prevalence rates of insufficient physical activity were higher in females compared to males across all regions (**Figure 3.10**).

Figure 3.8. Regional correlation of obesity with diabetes in (A) males and (B) females. Data presented is based on Non-Communicable Disease Risk Collaboration (NCD-RisC) group estimates of diabetes and obesity prevalence rates, in 2014 and 2016, respectively. Scatter Graphs A and B represent the age-standardised prevalence rates of diabetes in 2014 (x-axis) against obesity in 2016 (y-axis), across 200 countries, in males and females, respectively. For each graph, the vertical dotted line represents the worldwide age-standardised prevalence rates of diabetes in 2014, and the horizontal line represents the worldwide age-standardised prevalence rate of obesity in 2016. The 200 countries were categorised into 9 regions and colour coded in the figure, which include: Central and Eastern Europe, Central Asia Middle East and North Africa, East and SouthEast Asia, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]

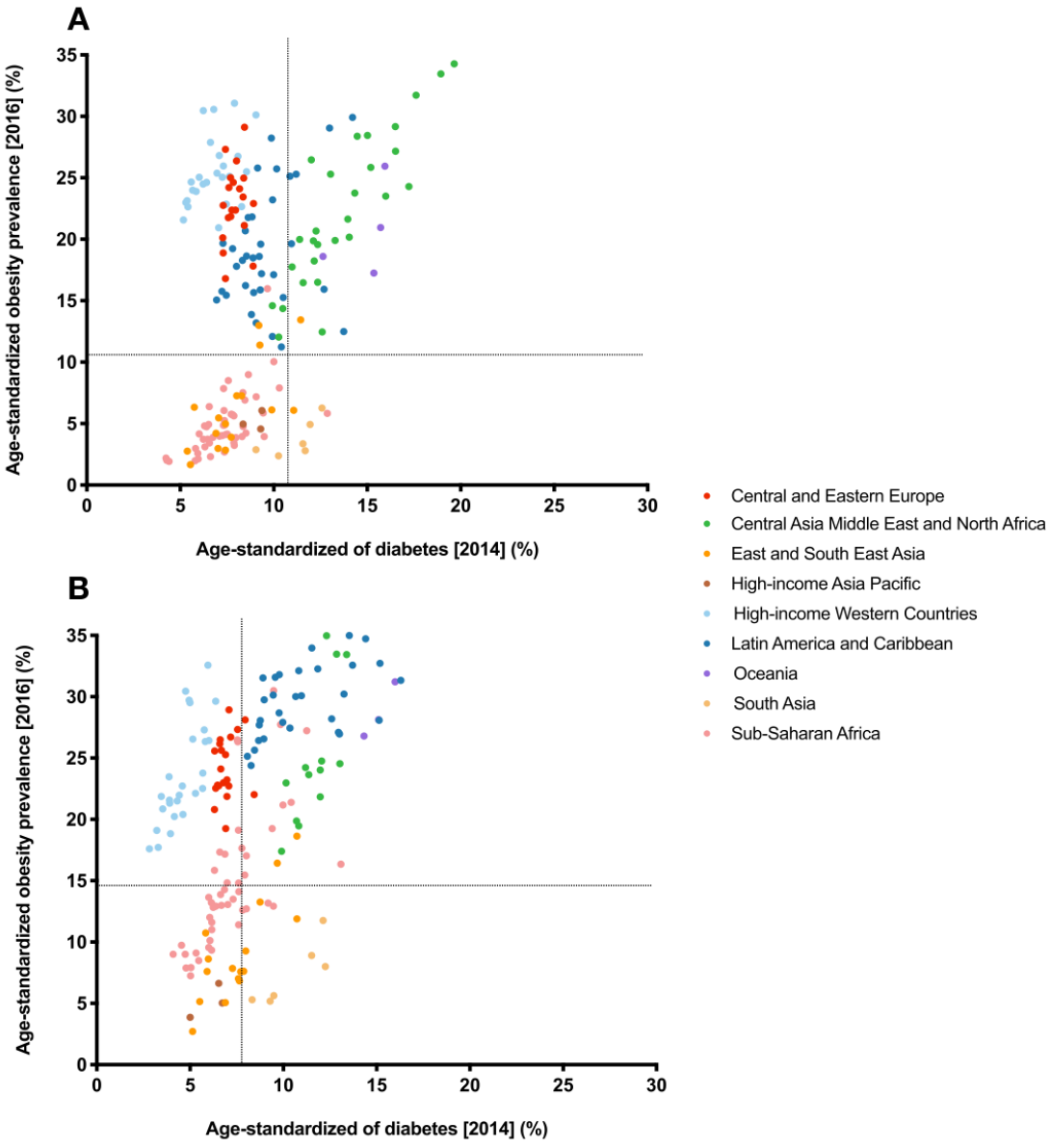


Figure 3.9. Obesity and diabetes trends in Central and Eastern Europe, Central Asia Middle East and North Africa, and High-Income Western countries regions, in (A) males and (B) females. Data presented is based on Non-Communicable Disease Risk Collaboration (NCD-RisC) group estimates of diabetes (1975-2014) and obesity (1980-2016) across 3 regions. Figure adapted from Alkaf et al.[277]

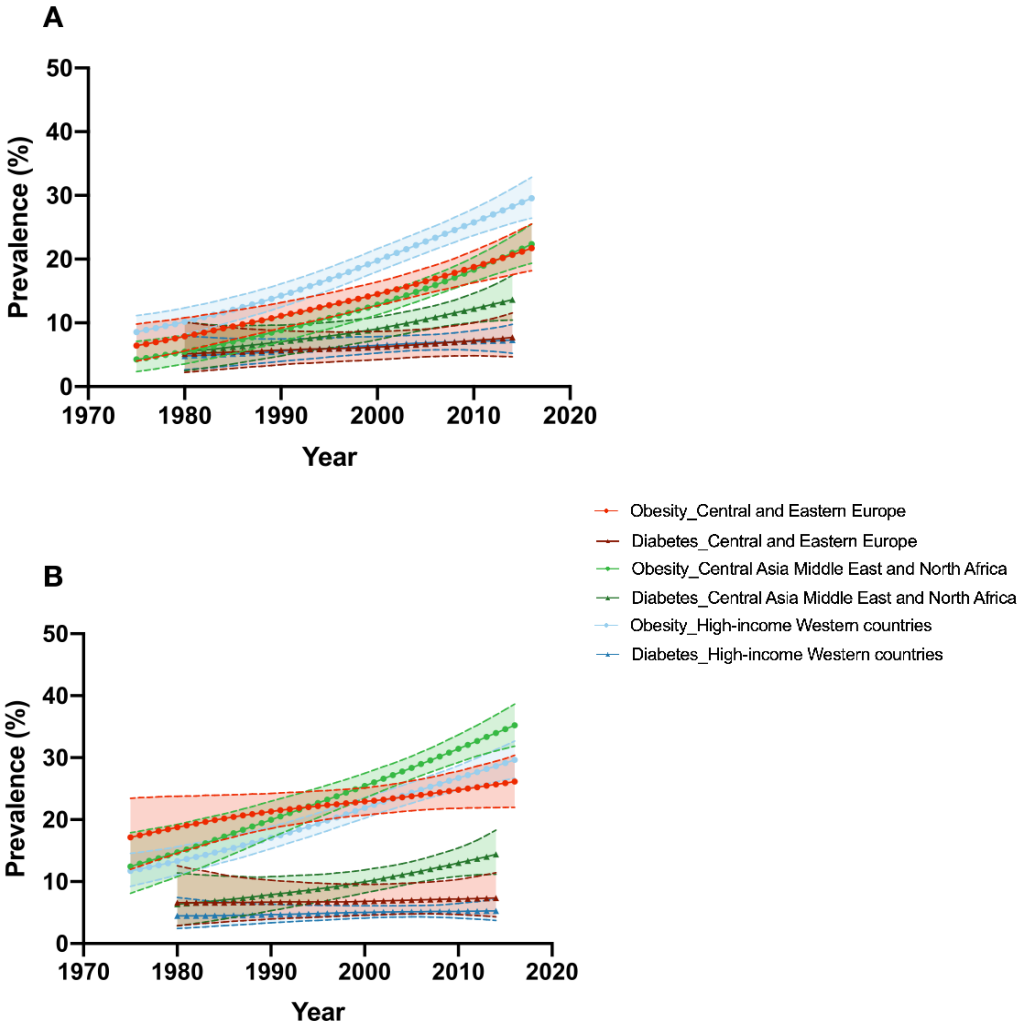
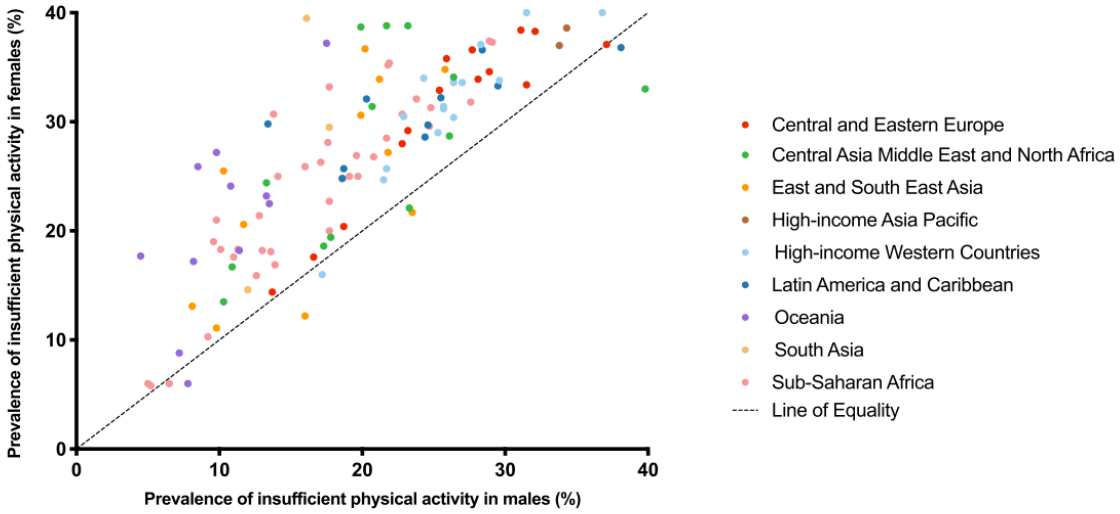


Figure 3.10. Comparison of age standardised insufficient physical activity prevalence in 2016 by sex. Data presented here is based on estimates of insufficient physical activity rates in 2016, from a pooled analysis of 358 population-based surveys, performed by Guthold and colleagues[105]. The scattergraph represents the age-standardised insufficient physical activity prevalence rates in 168 countries, in males (x-axis) against females (y-axis). All countries were categorised into 9 regions, based NCD-RisC group region definition and colour coded in the figure. The 9 regions include: Central and Eastern Europe, Central Asia Middle East and North Africa, East and SouthEast Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]



In males, insufficient physical activity rates were significantly correlated with diabetes rates in 2 out of the 9 regions, with the strongest correlation in Central Asia, Middle East and North Africa ($r = 0.66$, $p < 0.001$) (**Table 3.1**). Whereas in females, the correlation of insufficient physical activity rates with diabetes rates was significant in 3 out of the 9 regions, with the strongest correlation in Central and Eastern Europe ($r = 0.72$, $p = 0.002$) and Central Asia Middle East and North Africa ($r = 0.65$, $p < 0.001$) (**Table 3.1**).

Figure 3.11 demonstrates the correlation of insufficient physical activity with diabetes across all 9 regions in males (**Figure 3.11 A**) and females (**Figure 3.11 B**). Most countries with insufficient physical activity rates lower than the worldwide mean (males, 23.4%; females, 31.7%) also had lower diabetes rates compared to the worldwide mean (males, 9%; females, 7.9%), and vice versa. This was observed in both males and females (**Figure 3.11**). Despite that, several countries, particularly in High-income Western countries and Central and Eastern Europe regions, have some of the highest rates of insufficient physical activity, but still have low rates of diabetes, when compared to others. On the other hand, countries in Oceania in particular, have some of the lowest rates of insufficient physical activity, despite having some of the highest rates of diabetes in the world.

3.5.6 Joint effects of obesity and physical inactivity levels with diabetes

In order to assess country and sex-specific joint effects of obesity and physical inactivity on diabetes, I analysed the distribution of obesity and insufficient physical activity rates in 163 countries in relation to diabetes risk (**Figure 3.12**). Overall, males and females across all countries have varying risk of diabetes (low, moderate, or high), irrespective of differences in their insufficient physical activity levels. Also, countries with obesity rates above 30% have high diabetes risk, irrespective of their insufficient physical activity rates. However, there is no such clear cut-off for insufficient physical activity (**Figure 3.12**).

Despite this overall picture, there are some countries that are outliers of this observation, and there seems to be a difference between the sexes. The majority of males with high obesity rates have a high risk of diabetes; however, there are males in some countries with high diabetes risk despite having low or moderate rates of obesity. That is not the case in females, where most countries with high obesity rates also have high risk of diabetes (**Figure 3.12**). In addition, males but not females, in some countries of South Asia, including Bhutan, Bangladesh and Nepal, have high diabetes risk despite having low rates

of both obesity and insufficient physical activity (**Table 3.2**). In contrast to this scenario, males and females in some countries particularly in the High-income Western region, including the United Kingdom, have a relatively low risk of diabetes despite having high rates of both obesity and insufficient physical activity (**Table 3.2**).

Figure 3.11. Regional correlation of insufficient physical activity with diabetes in (A) males and (B) females. Data presented is based on diabetes prevalence rates estimates by Non-Communicable Disease Risk Collaboration (NCD-RisC) group in 2014, and insufficient physical activity prevalence rates estimates in 2016, from a pooled analyses of 358 population-based surveys, performed by Guthold and colleagues[105]. Scatter Graphs A and B represent the age-standardised prevalence rates of diabetes in 2014 (x-axis) against insufficient physical activity 2016 (y-axis), across 163 countries, in men and women, respectively. For each graph, the vertical dotted line represents the worldwide age-standardised prevalence rate diabetes in 2014, and the horizontal line represents the worldwide age-standardised prevalence rate of insufficient physical activity in 2016. Based on NCD-RisC group definition, the 163 countries were categorised into 9 regions and colour coded in the figure, which include: Central and Eastern Europe, Central Asia Middle East and North Africa, East and SouthEast Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa. Figure adapted from Alkaf et al.[277]

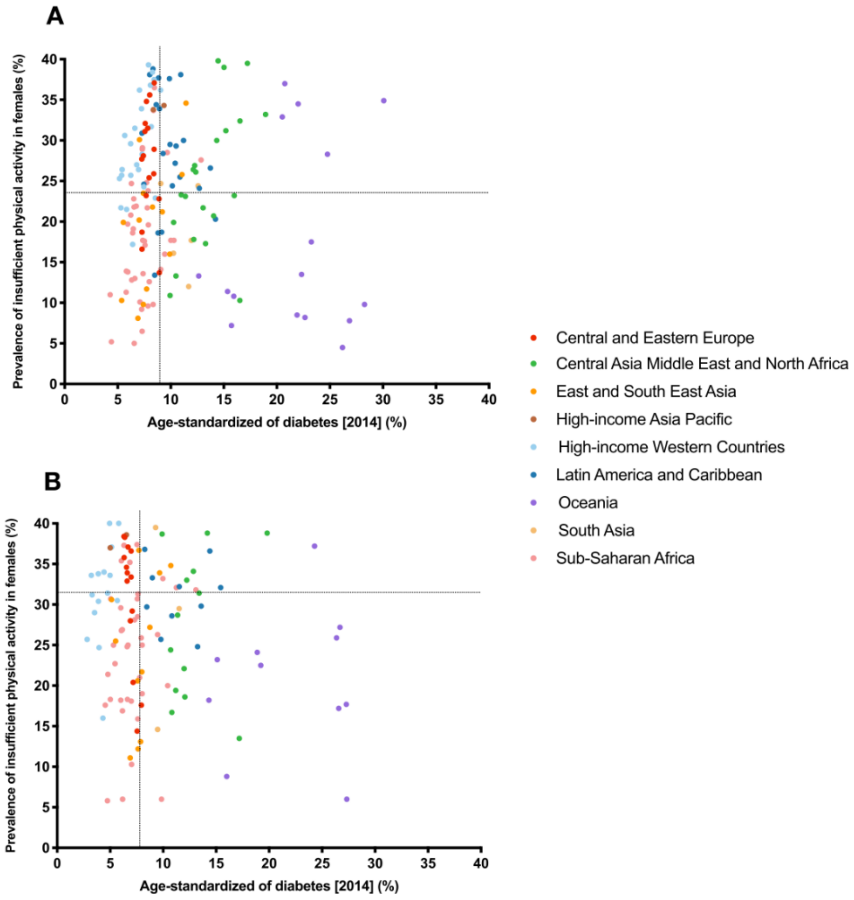
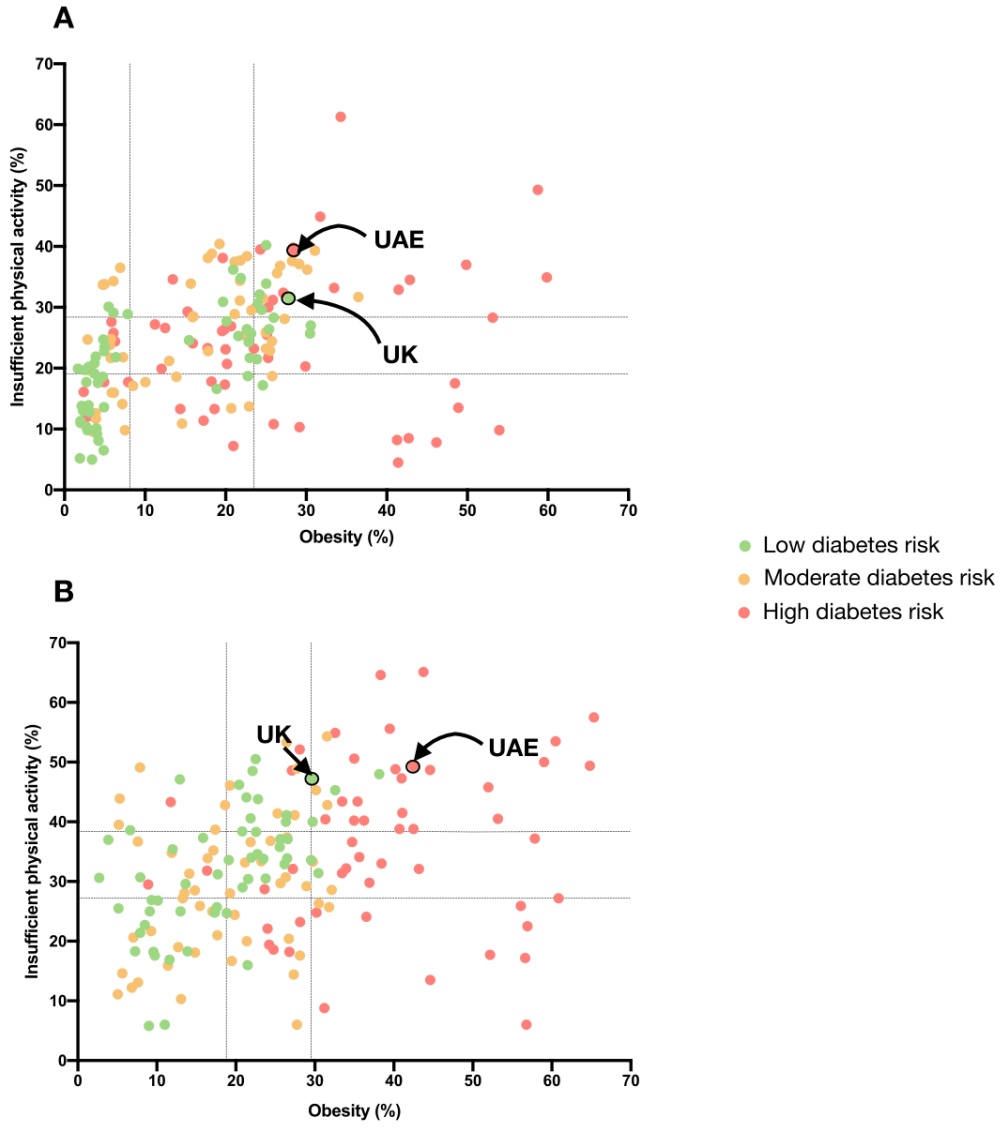


Figure 3.12. The joint relationship of obesity and insufficient physical activity (IPA) with diabetes (DM) risk in (A) males and (B) females. Diabetes and obesity rates were estimated by the Non-Communicable Disease Risk Collaboration (NCD-RisC) group in 2014 and 2016, respectively. IPA rates were estimated by the World Health Organisation (WHO) in 2016. Scatter Graphs represent distribution of age-standardised prevalence rates of obesity (x-axis) against IPA rates in 163 countries. Estimates of obesity, IPA, and diabetes were grouped into three quantiles. For diabetes rates, countries that fell in the first quantile (with the lowest rates of diabetes) were coded as Low risk (green), and those that fell in the second and third quantiles, were coded as Moderate (yellow) and High (red) risk, respectively. Dotted vertical and horizontal lines represent the first and second quantiles of obesity and IPA rates, respectively.



3.5.7 The UAE versus the UK: contribution of obesity and insufficient physical activity towards diabetes risk

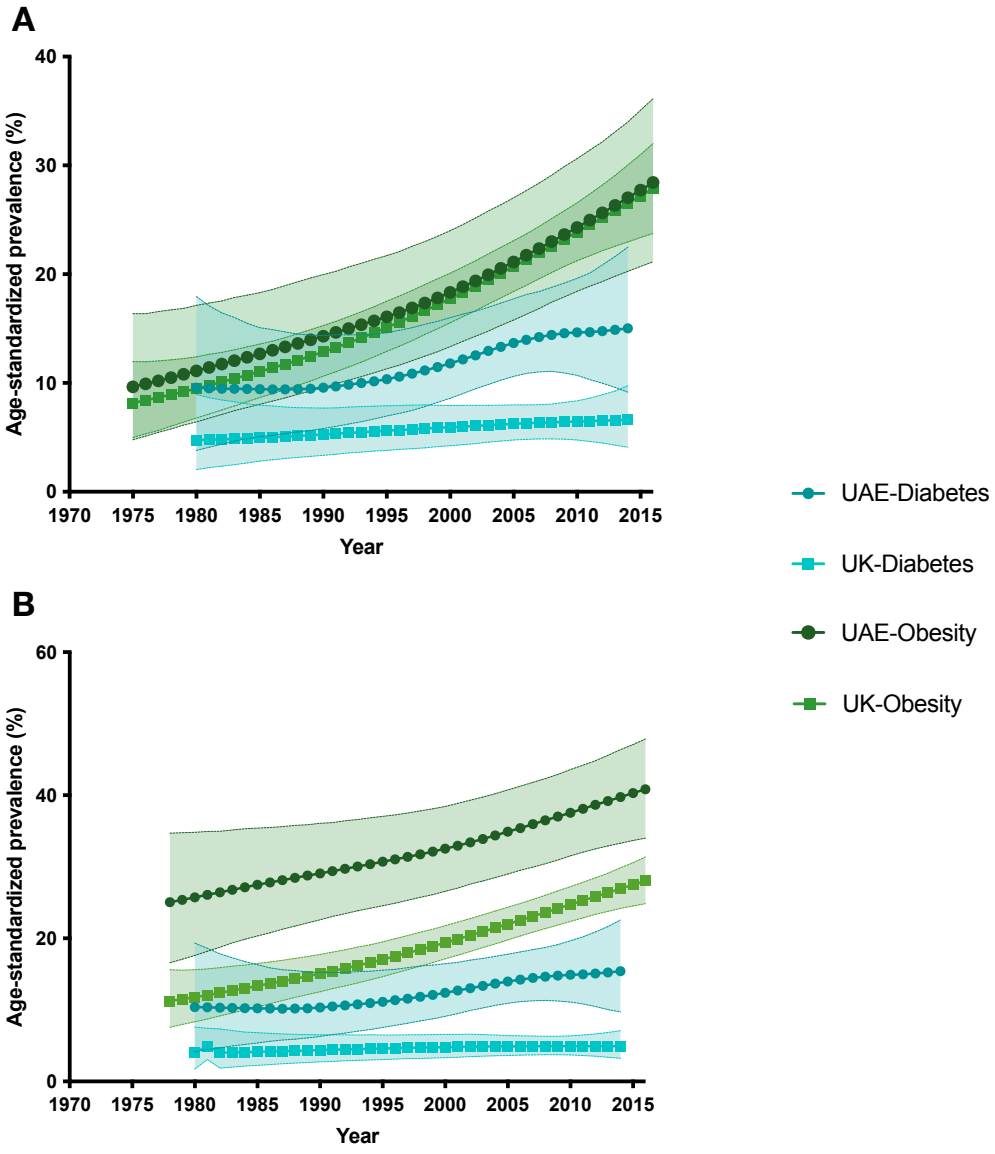
According to the NCD-RisC group, the UAE was estimated to have a diabetes prevalence of 15.0% (9.2-22.5) in males and 15.4% (9.7-22.6) in females in 2014[205]. This ranks the UAE as the 25th country with the highest rates of diabetes worldwide. The UAE was also ranked among the top 20 and 25 countries according to their obesity prevalence rates in males [28.4% (21.2-36.1)] and females [42.5% (35.0-50.2)], respectively, at a global level. In addition, according to the WHO estimates in 2016, insufficient physical activity rates in the UAE were 39% and 49.5% in males and females, respectively. This ranks the UAE as the 9th and the 15th country with some of the highest rates of insufficient physical activity rates worldwide, in males and females, respectively.

In comparison to the UAE, males in the United Kingdom (UK) have very similar rates of obesity [UAE vs. UK; 28.4% (21.2-36.1) vs. 27.9% (23.7-32.0)], but lower rates of insufficient physical activity (UAE vs. UK; 39.0% vs. 31.5%), and remarkably lower diabetes rates [UAE vs. UK; 15.0% (9.2-22.5) vs. 6.6% (4.1-9.7)]. Over the 34 years studied (1980-2014), the percentage increase in diabetes rates among males in the UAE was 56.9% compared to 39.1% in the UK (**Figure 3.13 A**). Whereas the percentage increase in obesity prevalence rates among males in the UAE was 155%; lower than that of the UK (194%) (**Figure 3.13 A**).

Observations were different in females. In 2016, females in the UAE have higher rates of insufficient physical activity rates compared to those in the UK (UAE vs. UK; 49.5% vs. 40.0%) and have approximately 1.5-fold higher obesity rates [UAE vs. UK; 42.5% (35.0-50.2) vs. 29.7% (25.6-34.0)] (**Figure 3.13 B**). In addition, females in the UAE had higher rates of both obesity and diabetes across all years

(1980-2014) compared to those in the UK. Between 1980 and 2014, females in the UAE have undergone a 48.1% and 54.9% increase in their diabetes and obesity rates, respectively. Whereas diabetes and obesity rates in females in the UK increased by 22.5% and 128%, respectively (**Figure 3.13 B**).

Figure 3.13. Obesity and diabetes trends in the UAE and the UK in (A) males and (B) females. Data presented is based on Non-Communicable Disease Risk Collaboration (NCD-RisC) group estimates of diabetes (1975-2014) and obesity (1980-2016).



3.6 DISCUSSION

The worldwide rising rates of diabetes, particularly in Westernised populations, have been associated with environmental changes related to obesity. Excessive caloric intake and reduced physical activity are the main features of a high-risk environment for diabetes development, and the relationship of these environmental stressors with disease risk was sometimes considered to be straightforward. In this chapter, I present and discuss data that conflicts with this supposition. Obesity rates have been rapidly rising since 1980 in every part of the world, but this was not reflected in a proportional increase in diabetes rates in several regions, including Europe and High-income Western Countries. These data suggest that certain populations, such as those in the Middle East (including the UAE), may be exposed to similar environmental stressors to some Westernised/European populations, yet have much higher rates of diabetes.

The findings presented here must be interpreted with caution. Obesity rates reported here are based on the universal BMI cut-off point of ≥ 30 Kg/m^2, and that could have resulted in underestimating obesity levels in some regions, and unreliable comparisons with other regions. BMI is a simple and non-invasive method for assessing obesity and excessive fat stores within a single population but becomes less reliable when comparing adiposity measures between populations[62, 297]. One of the largest studies conducted to investigate the impact of ethnicity on the association of age and BMI with T2D was the Diabetes Epidemiology Collaborative analysis of Diagnostic criteria in Europe - Diabetes Epidemiology Collaborative analysis of Diagnostic criteria in Asia (DECODE-DECODA) study[293]. Nakagami et al. analysed studies/surveys representing different ethnic groups, including European, Maltese, Indian, Chinese, and Japanese populations. They found that the variation in T2D prevalence across ethnic groups cannot be explained by differences in BMI alone, and that ethnicity modified the effect of BMI on age-adjusted prevalence of T2D[293].

These findings support the need for region- and ethnicity-specific definitions of obesity. The National Institute for Health and Care Excellence (NICE) recommendations for T2D prevention includes using lower BMI thresholds for South Asian, Chinese, black African, and Afro-Caribbean populations (23 Kg/m^2 to indicate increased risk and 27.5 Kg/m^2 to indicate high risk)[298]. These raise awareness of the need for lifestyle interventions at a lower BMI threshold for specific ethnic populations. Such guidelines for populations in the Gulf region, such as the UAE, are not yet in place, and need to be considered.

The heterogeneity in the relationship between obesity (as measured by BMI) and diabetes observed here could also be explained by differences in the contribution of reduced insulin secretion and reduced insulin sensitivity to the development of diabetes between populations[207, 299]. Differential patterns in body fat distribution between populations should also be acknowledged. Two populations could have identical distributions of BMI, but still show large differences with respect to their intra-abdominal fat accumulation[207, 300, 301]. South Asian populations, for example, are more insulin-resistant and more prone to abdominal obesity and low muscle mass than other ethnic groups[207, 300, 301], explaining their increased risk of developing diabetes, despite the absence of significant obesity as assessed by BMI >30. This is in accord with the data presented here. Obesity and physical inactivity rates in South Asia did not correlate with diabetes rates in the region overall: countries including Bhutan, Bangladesh and Nepal were estimated to have some of the highest diabetes rates in the world yet have relatively low obesity and physical activity rates.

Sex differences were observed across all my analyses. Previous studies, largely based on Europeans, showed that diabetes is more frequently diagnosed at a lower age and BMI in males, despite the fact that obesity is more common in females[302]. This trend was also observed here, not only in European

populations (Central and Eastern Europe and High-income Western countries regions), but across all regions. Diversities in culture, lifestyle, and socioeconomic status are some of the factors that may impact differences between males and females in predisposition, development, and clinical presentation of diabetes[302]. Besides obesity, other risk factors including genetic effects and epigenetic mechanisms, nutritional factors and sedentary behaviours affect disease risk and complications differently in both sexes[302]. Indeed, in this study, females had higher insufficient physical activity rates compared to males, despite having similar rates of diabetes.

The secondary analyses presented here were based on two of the most well-established risk factors for T2D and their contribution to differences in disease risk across populations. It is important to note, however, that the development of T2D involves a complex interplay of adverse lifestyle and genetic exposures. This means that other factors may be confounding these findings, particularly those that are also associated with cultural and faith/religious groups: fat distribution (as discussed above), patterns of physical activity, smoking, alcohol consumption, and nutrition including balance of macronutrients, as well as micronutrient deficiencies), for example. The WHO produces yearly estimates (2007-2018) of cigarette and tobacco smoking prevalence rates in more than 190 countries. Secondary analyses that account for the effect of factors like smoking, diet and nutrition, socioeconomic status on findings presented here are urgently required.

Despite extensive efforts by the NCD-RisC in deriving worldwide estimates, some countries had none or limited data sources, including those in Sub-Saharan Africa, the Caribbean, Central Asia, and Central Europe[205]. Therefore, estimates for these countries relied mostly or entirely on the statistical model. In addition, diabetes rates reported here include those with T1D, however as the majority of diabetes

cases in adults are type 2[303], the observed rates and rising trends in diabetes prevalence presented here are likely due to T2D.

It should also be noted that when NCD-RisC and WHO derive country estimates, they do not consider ethnicity breakdown within the country. Thus, estimates from countries like the UAE, which only comprises ~15% Arab-Emiratis, may not be representative of the Arab-Emiratis. That could potentially mean when studying Arab-Emiratis separately, we would observe different results to those presented in this secondary-analysis. More details on this are in Chapter 4 (the next chapter).

In conclusion, populations like those of the UAE, can be exposed to similar environmental stressors as other populations (e.g., UK), but still have higher rates of diabetes. One reason behind this increased risk may be an increased genetic/epigenetic susceptibility in this population. Whether or not there is supporting evidence for this will be further discussed in the coming chapters.

Table 3.1. Pearson and Spearman correlation tests for region-specific correlation of obesity and insufficient physical activity (IPA) with diabetes. Based on NCD-RisC group definition, estimates were reported for 9 regions including, Central and Eastern Europe, Central Asia Middle East and North Africa, East and South East Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa.

	Pearson				**Spearman**			
Males	**Obesity**		**IPA**		**Obesity**		**IPA**	
Region	**r**	***P*-value †**	**r**	***P*-value †**	**r**	***P*-value †**	**r**	***P*-value †**
Central and Eastern Europe	0.136	0.567	0.007	0.765	0.236	0.316	0.082	0.763
Central Asia Middle East and North Africa	0.864	<0.001	0.486	<0.001	0.829	<0.001	0.661	<0.001
East and South East Asia	0.700	0.003	0.281	0.062	0.735	0.002	0.550	0.055
High-income Asia Pacific	0.314	0.797	Too few pairs (only 2 countries)		0.500	0.999	Too few pairs (only 2 countries)	
High-income Western Countries	0.366	0.060	0.243	0.014	0.373	0.055	0.469	0.021
Latin America and Caribbean	0.317	0.064	0.077	0.200	0.208	0.231	-0.303	0.160
Oceania	0.910	<0.001	0.134	0.015	0.802	<0.001	0.145	0.579
South Asia	0.710	0.114	0.042	0.740	0.714	0.136	-0.100	0.950
Sub-Saharan Africa	0.600	<0.001	0.137	0.019	0.666	<0.001	0.286	0.073
Females	**Pearson**				**Spearman**			
	Obesity		**IPA**		**Obesity**		**IPA**	
Region	**r**	***P*-value †**	**r**	**P-value**	**r**	***P*-value †**	**r**	***P*-value †**
Central and Eastern Europe	0.211	0.372	0.671	<0.001	0.287	0.220	-0.729	0.002
Central Asia Middle East and North Africa	0.818	<0.001	0.396	0.001	0.848	<0.001	0.653	<0.001
East and South East Asia	0.78	<0.00	0.10	0.281	0.70	0.003	0.335	0.263

	7	1	5		3			
High-income Asia Pacific	0.758	0.452	Too few pairs (only 2 countries)		0.500	0.999	Too few pairs (only 2 countries)	
High-income Western Countries	0.770	<0.001	0.348	0.002	0.800	<0.001	0.600	0.002
Latin America and Caribbean	0.589	<0.001	0.003	0.816	0.664	<0.001	0.034	0.879
Oceania	0.837	<0.001	0.114	0.184	0.679	0.004	0.203	0.432
South Asia	0.859	0.029	0.000	0.960	0.771	0.103	-0.300	0.683
Sub-Saharan Africa	0.741	<0.001	0.106	0.041	0.779	<0.001	0.284	0.076

† P-value not corrected for multiple testing.
IPA, Insufficient physical activity

Table 3.2. Countries that fall as outliers of the positive association of obesity and insufficient physical activity (IPA) with diabetes risk in males and females. Age-standardised prevalence rates of obesity, IPA, and diabetes rates in 163 countries were each split into tertiles. For each of obesity and insufficient physical activity rates, countries that fell in the first, second, and third tertiles, were colour-coded as Low (green), Moderate (yellow), and High (red), respectively. For diabetes rates, countries that fell in the first tertile (with the lowest rates of diabetes) were colour-coded as Low risk (green), and those that fell in the second and third tertiles, were colour-coded as Moderate (yellow) and High risk (red), respectively. Code for interpreting the table: Bhutan in South Asia has low obesity (green) and low IPA (green), but high diabetes risk (red).

Males				
Country	**Region**	**Obesity**	**IPA**	**Diabetes**
Bhutan	South Asia	Low	Low	High
Bangladesh	South Asia	Low	Low	High
Nepal	South Asia	Low	Low	High
Seychelles	Sub-Saharan Africa	Low	Low	High
Cabo Verde	Sub-Saharan Africa	Low	Low	Moderate
Comoros	Sub-Saharan Africa	Low	Low	Moderate
Gambia	Sub-Saharan Africa	Low	Low	Moderate
China	East and SouthEast Asia	Low	Low	Moderate
Lao PDR	East and SouthEast Asia	Low	Low	Moderate
Germany	High-income Western Countries	High	High	Low
Greece	High-income Western Countries	High	High	Low
Belgium	High-income Western Countries	High	High	Low
Norway	High-income Western Countries	High	High	Low
Ireland	High-income Western Countries	High	High	Low
United Kingdom	High-income Western Countries	High	High	Low
Romania	Central and Eastern Europe	High	High	Low
Females				
Country	**Region**	**Obesity**	**IPA**	**Diabetes**
Nepal	South Asia	Low	Low	Moderate
Togo	Sub-Saharan Africa	Low	Low	Moderate
Cambodia	East and SouthEast Asia	Low	Low	Moderate
China	East and SouthEast Asia	Low	Low	Moderate

Myanmar	East and SouthEast Asia	Low	Low	Moderate
Malta	High-income Western Countries	High	High	Low
United Kingdom	High-income Western Countries	High	High	Low
United States	High-income Western Countries	High	High	Low
New Zealand	High-income Western Countries	High	High	Low

IPA, Insufficient physical activity

Chapter Four: The contribution of adiposity towards increased risk of T2D among Arab-Emiratis

In the previous chapter, results highlighted the high rates of T2D, obesity and insufficient physical activity in populations of the Gulf region, including the UAE. It is therefore likely that a large proportion of disease risk is attributable to the increased exposure to adiposity and lack of physical activity, but data supporting this notion is limited. In this chapter, I assess the contribution of adiposity (as estimated by BMI, waist-to-hip ratio, and derived estimates of visceral adipose tissue volume) towards differences in magnitude of T2D risk observed between populations using data I collected in Arab-Emiratis (study population in the UAE), comparing that to data from various other ethnic groups including Europeans, South Asians, and Afro-Caribbeans from the UKBB data set. Supplementary figures and tables can be accessed through Zenodo using this link: https://zenodo.org/record/7825396#.ZDfrqC0Rro4.

4.1 OVERVIEW

I utilised two Arab-Emirati populations: EM1 included 7,283 Arab-Emirati adults, a representative sample of the UAE population. EM2 included 1,448 Arab-Emirati adults recruited in a case-control study

(459 with normal glucose tolerance (NGT), 371 with prediabetes, 618 with T2D). Cross-sectional baseline data on 423,803 Caucasians (CAU), 7,841 South Asians (SA), and 5,455 Afro-Caribbeans (AC) were retrieved from the UK Biobank (UKBB). I investigated age-specific T2D and obesity prevalence rates in Arab-Emiratis (EM1) as compared to CAU, SA and AC individuals. I tested the contribution of BMI to the differences in risk in Arab-Emiratis, SA, and AC individuals as compared to CAU. Lastly, I tested the strength of association of other adiposity measures with T2D across ethnic groups, and with insulin resistance and β-cell function in Arab-Emiratis (EM2). I found a striking 12 to 14-fold excess risk of T2D in Arab-Emiratis, compared to CAU individuals that is not attributable to differences in BMI. Unlike BMI, my findings indicate that central adiposity measures may contribute more to differences in T2D risk between populations. This highlights the importance of more precise characterisation of unfavourable fat distribution in the Arab-Emirati population that could potentially better explain their increased disease risk.

4.2 AIMS

To determine the burden of T2D and related glycaemic disturbances amongst Arab-Emiratis and investigate the contribution of adiposity measures to differences in risk of T2D between Arab-Emiratis and other ethnic groups.

4.2.1 Objectives

1. To establish a cross-sectional sample of Arab-Emiratis that is representative of the general population to investigate the contribution of obesity towards increased risk of T2D among Arab-Emiratis (EM1), including:

a. To assess how well the EM1 sample represents the characteristics of the general UAE population in terms of age structure, age-specific T2D and obesity rates.

b. To determine age-specific prevalence rates of T2D amongst Arab-Emirati men and women compared to individuals of CAU, SA, and AC descent from the UKBB.

c. Evaluate the relationship of ethnicity with risk of T2D and investigate the contribution of BMI as a measure of adiposity to increased risk of T2D amongst Arab-Emiratis compared to CAU individuals.

2. To collect a sample of Arab-Emirati men and women (**EM2**) with or without T2D in a case-control study to enable the investigation of:

a. The relationship of central adiposity as measured by waist circumference, waist-to-hip ratio, and anthropometry-derived visceral adipose tissue volume with T2D risk among Arab-Emiratis in comparison to other ethnic groups in the UKBB.

b. The association of adiposity measures with insulin resistance and β-cell function as measured by the homeostatic model assessment among Arab-Emiratis.

Figure 4.1 is a flowchart depicting the cohorts and sample populations investigated in this part of my study to achieve the described specific aims.

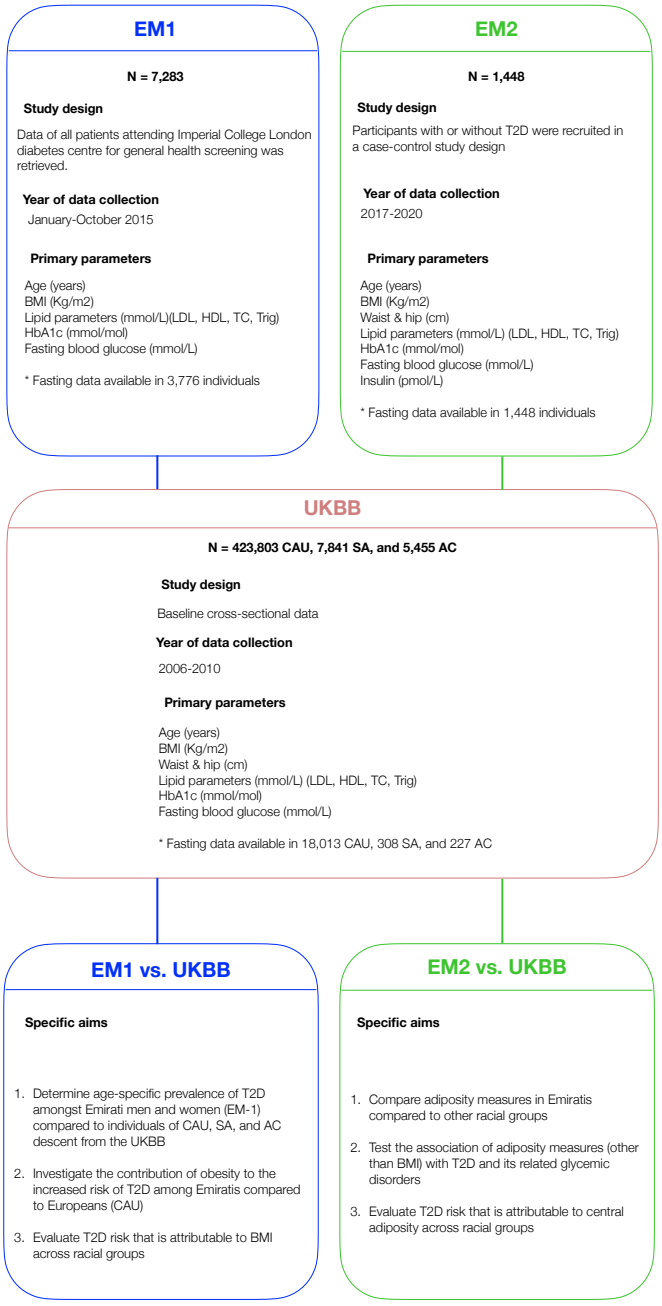
4.3 INTRODUCTION

T2D is the fifth leading cause of death at a global level, and a major contributor to the development of coronary heart disease, stroke, peripheral vascular disease, end-stage renal failure and lower limb amputation[304, 305]. The rise in obesity prevalence rates in ageing populations is considered one of the main driving factors[306] . The strong link between age and BMI and the prevalence and incidence of T2D has been consistently reported in cross-sectional and prospective studies[79-81].

As highlighted in Chapter 3, the prevalence of T2D varies significantly between geographical regions and

ethnic populations[205], however, this variation has only been partially attributed to differential distribution of established risk factors across regions. According to my secondary analyses of global datasets (Chapter 3), populations in the Gulf region including the UAE, have a strong correlation between obesity levels and prevalence of T2D. However, given the established limitations of BMI at capturing the complex distribution of fat, I hypothesise that other simple and clinically applicable indices, such as waist and waist-to-hip ratio, may better explain differences in T2D risk between populations.

Figure 4.1 Flowchart describing the sample populations investigated in this part of the study to achieve the specific aims. EM1 included 7,283 Arab-Emirati men and women from the general population of the UAE. EM2 included 1,448 Arab-Emirati men and women collected in a case-control study. UK biobank (UKBB) baseline data from 423,803 Caucasians (CAU), 7,841 South Asians (SA), and 5,455 Afro-Caribbeans (AC) were retrieved.



Compared to European populations, such as the UK, the UAE has over 2-fold higher rates of T2D. Despite that striking fact, little is known about the epidemiology and public health burden of T2D and its related metabolic disturbances amongst Arab-Emirati men and women. I, therefore, carried out a population-based study to determine the prevalence of T2D in Arab-Emiratis living in the UAE, compared to individuals of CAU, SA, and AC descent living in the UK. In addition, I investigated the contribution of adiposity measures towards the risk T2D and its related glycaemic disturbances in Arab-Emiratis.

4.4 MATERIALS AND METHODS

4.4.1 Study design and participant characterisation

I investigated 8,731 Arab-Emiratis from EM1 (n = 7,283) and EM2 (n = 1,448) cohorts and 437,099 individuals from the UKBB cohort. Study design, participant recruitment, screening and diagnosis were described in detail in Chapter 2 (UKBB; section 2.2, EM1; section 2.3, EM2; section 2.4). Here, I describe them again briefly.

4.4.1.1 EM1 and EM2

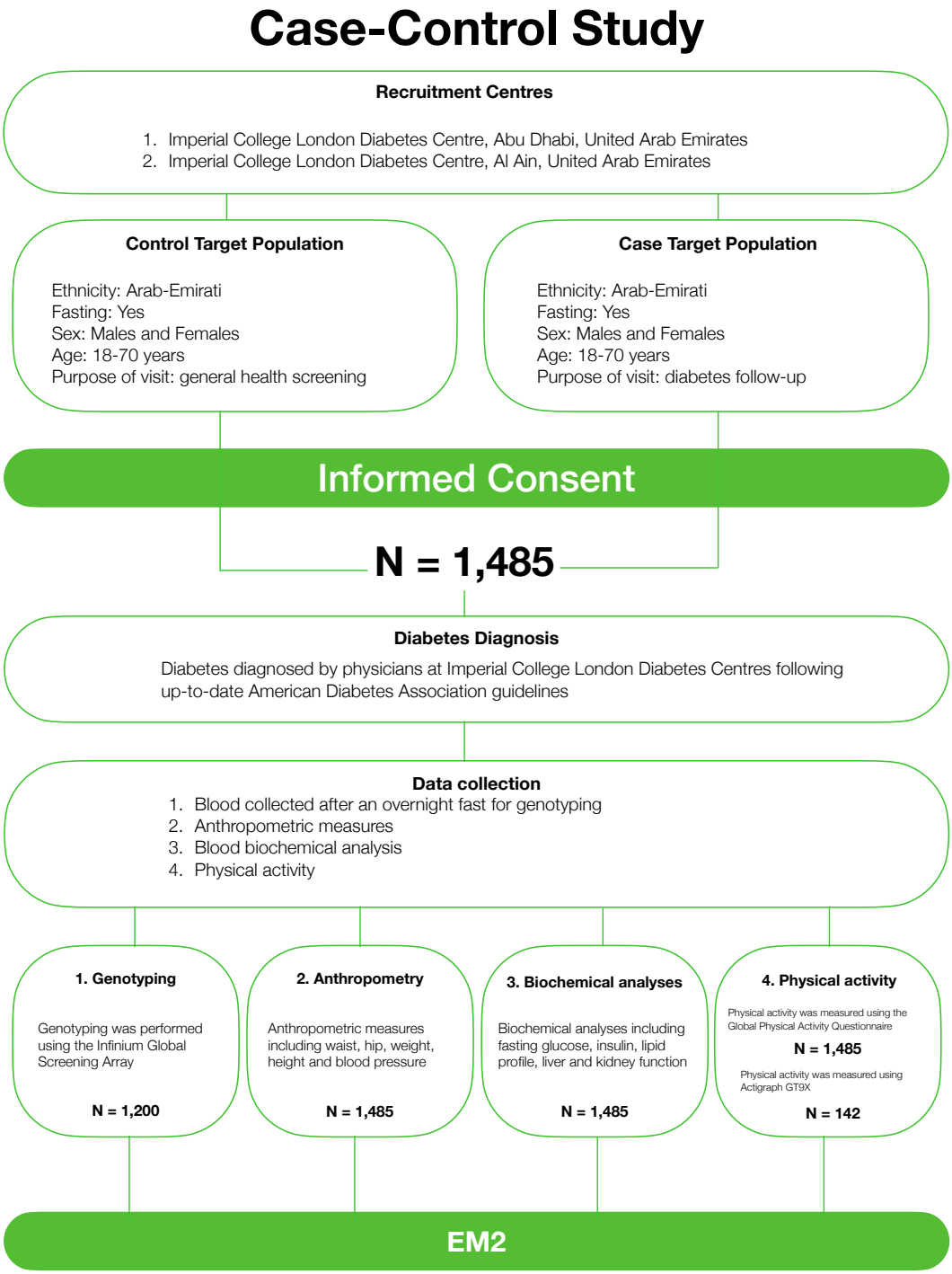
The EM1 sample population included 7,283 Arab-Emirati males and females, aged ≥ 18 years, who attended ICLDC between January and October 2015, for general health screening. A more detailed description of what proportion of patients attend ICLDC for general health screening or diabetes follow-up can be found in Chapter 2 section 2.3.2.1. In order to assess how representative the EM1 sample of the general UAE population, I compared age structure, as well as age-specific obesity and T2D rates observed here in EM1 to that reported by the Weqaya program, a population-wide cardiovascular

program including 50,138 Arab-Emiratis aged 18 years or older[284].

The EM2 sample population included Arab-Emirati men and women recruited in a case-control study design from ICLDC between 2017 and 2020 (**Figure 4.1 and 4.2**). Arab-Emirati participants were screened by trained nurses following ICLDC's established protocols and the American Diabetes Association guidelines. A more detailed description of participants screening can be found in Chapter 2 section 2.3.3.

Results were extracted from ICLDC's system records, which included fasting plasma glucose (FBG), glycated haemoglobin (HbA1c), triglycerides (TG), total cholesterol, low-density lipoprotein cholesterol (LDL), and high-density lipoprotein cholesterol (HDL). Additional anthropometric measures and biochemical analyses that were collected and assessed only in EM2 (not available for EM1) included: waist and hip measurements, fasting insulin, liver and kidney function tests [aspartate transaminase (AST), alanine transaminase (ALT), bilirubin, creatinine, and urea]. Moreover, fasting data was available for all EM2 participants (n = 1,448), but only in 3,776 individuals in EM1.

Figure 4.2 Flowchart describing the study design and data collection for the EM2 sample population. 1,485 Arab-Emirati men and women were recruited in a case-control study between 2017 and 2020. Physical activity data collected is presented in Chapter 5 of this thesis. Genotyping data are presented in Chapter 6 of this thesis.



4.4.1.2 UKBB

Data retrieved from the UKBB included: plasma glucose, HbA1c, total cholesterol, HDL, LDL, triglycerides, aspartate aminotransferase (AST), alkaline phosphatase (ALP), albumin, bilirubin, creatinine, and urea. The majority of UKBB participants were not fasting; fasting data were available for 18,013 CAU, 308 SA, and 227 AC participants. Insulin data were not available for the UKBB cohort.

4.4.2 Statistical analyses

R studio (Version 1.4.1106) was used for statistical analyses. Figures presented were plotted in R studio or Graphpad Prism (Version 9.2.0). Results are expressed as mean (standard deviation) and median (interquartile range) for continuous variables and as proportions and/or percentages for categorical variables with 95% confidence intervals as necessary. Data normality was tested by Shapiro Normality test or Anderson-Darling normality test (for larger than 5,000 sample sizes). **Supplementary figures 4.1-4.7** show the distribution of all parameters tested and **Supplementary Table 4.1** presents the results of normality tests performed. **Supplementary Table 4.2** summarises data completeness for EM1, EM2 and UKBB. The Wilcoxon rank-sum test was used to compare continuous parameters between sexes and diabetes status groups. The Chi-squared test was used to compare categorical parameters between sexes and diabetes status groups. The association of BMI, VAT volume, waist circumference, and WHR with T2D were tested by logistic regression in Arab-Emiratis (EM2) and CAU, SA, and AC individuals separately. Odds ratios (ORs) are presented for one unit change in each variable. Heterogeneity of effect between ethnic groups was tested by comparison of effect sizes, as measured by odds ratios. Differences between odds ratios across ethnic groups were considered significant ($p < 0.05$) if overlap between confidence intervals was less than 50%[307]. Linear regression to test for the association of BMI, waist, WHR, and anthropometry-derived VAT with HOMA-IR and HOMA-B was performed in EM2 only

(insulin data only available for EM2). HOMA-IR and HOMA-B were both log-transformed for the linear regression analyses. Standardised β coefficients were computed using the following equation[308] :

Standardised coefficient = β * (SD of predictor/SD of outcome)

Change in outcome by 1 SD change in predictor was then computed using the following equation:

Change in outcome = [exp(standardised coefficient) - 1] * SD of outcome

Multinomial logistic regression (MLR) was used to test for the relationship of ethnicity with T2D risk, and to determine the contribution of BMI to increased T2D risk among Arab-Emiratis (EM1), SA and AC individuals, compared to CAU individuals (reference category). In addition, MLR was used to compute ORs for T2D risk in Arab-Emirati (EM1), SA, and AC individuals at every level of general obesity; (1) Normal (BMI = 18.0-24.9 Kg/m^2) (2) Overweight (BMI = 25.0-29.9 Kg/m^2), (3) Class I/II obese (BMI = 30.0-39.9 Kg/m^2); and (4) Class III obese (BMI ≥ 40.0 Kg/m^2), with CAU individuals as the reference category. Unadjusted and adjusted ORs with 95% confidence intervals (CIs) were computed. All statistical tests were two-sided. P values were adjusted for multiple testing where necessary. 95% CIs for proportions (prevalence rates presented) were calculated as follows:

se = sqrt(proportion*(1-proportion)/total number of individuals

Lower limit = proportion – (se*1.96)

Upper limit = proportion + (se*1.96)

To calculate the magnitude of T2D risk attributed to BMI (BMI ≥ 30 Kg/m^2) and WHR (WHR > 0.9 for males and WHR > 0.8 for females) [population attributable risk (PAR)] amongst the EM2 sample population (case-control study) was computed using the following equation:

PAR (%) = Pe * [(OR-1)/OR] * 100

Pe = proportion of cases exposed

To calculate PAR for UKBB and EM1, I used the following equation:

PAR = [Pe(RR-1)*100/Pe(RR-1)+1]

Pe = proportion of population exposed

RR = Relative risk (estimated as the OR)

4.5 RESULTS

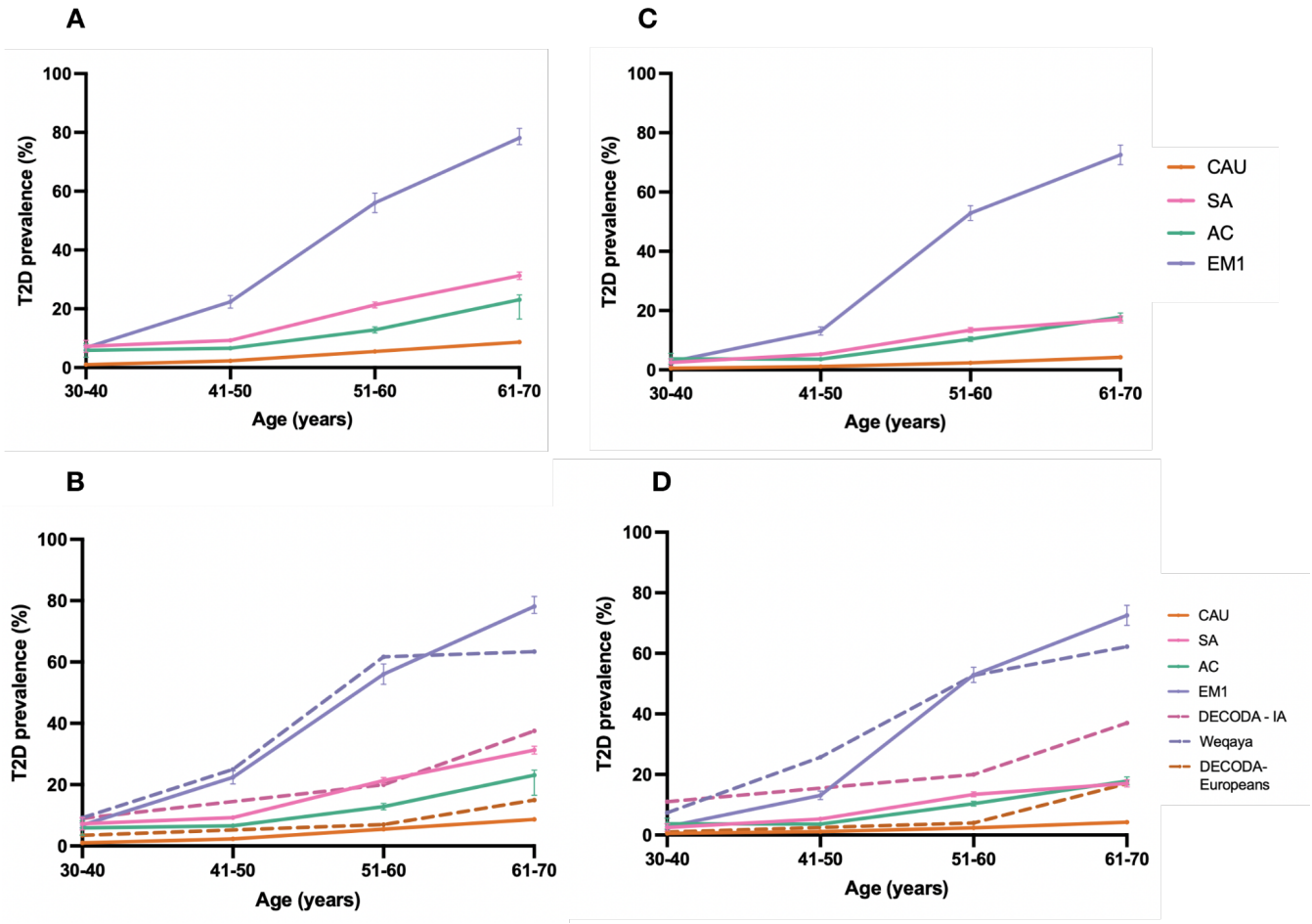
4.5.1 EM1 sample representativeness of the UAE population

7,283 Arab-Emirati individuals (2,704 males and 4,579 females) constituted the EM1 sample population; median (IQR) age was 33.0 (26.0-42.0) years, and median (IQR) BMI was 27.6 (24.1-32.2) Kg/m^2. EM1 demography was representative of the young general Arab-Emirati population (**Table 4.1**), with similar diabetes (14.7%) and obesity (36.2%) levels to those estimated by the IDF and NCD-RisC for the UAE in 2016/2017 [3, 205]. In addition, the EM1 sample had similar age-structure to the Weqaya program sample (n = 50,134 Arab-Emirati adults)[284], with a mean (SD) age of 37.1 (14.6) versus 37.6 (14.5) years in males and 35.1 (13.5) and 37.0 (14.2) years in females, respectively. Age-specific T2D prevalence rates in the EM1 sample were also similar to those reported by the Weqaya program (**Figure 4.3 C and D**)[284]. Lastly, age-specific obesity prevalence rates in EM1 were similar to that reported by the Weqaya program[284]. Obesity prevalence rates in males aged 30-40 years were 36.8% in EM1 (**Supplementary table 4.4**) compared to 36.2% in Weqaya[269]. In addition, obesity rates in males aged 51-60 years were 34.6% in EM1 (**Supplementary table 4.4**) compared to 37.5% in Weqaya[269]. Similarly, obesity rates in females aged 30-40 years and 51-60 years were 38.5% and 62.7% in EM1 (**Supplementary table 4.4**) and 39.0%

and 64.9% in Weqaya[284], respectively.

Furthermore, 3% of the studied population had heart disease and 1.6% had fatty liver disease. 1.9% were on blood pressure medications, 7% were on obesity medications, and 16.8% were on lipid-lowering medications.

Figure 4.3 Age- and ethnicity-specific prevalence rates of type 2 diabetes (T2D) in Arab-Emirati males (A, B) and females (C, D) from the EM1 sample population, compared to Caucasians (CAU), South Asians (SA) and Afro-Caribbeans from the UK Biobank (UKBB). The EM1 sample population included 2,704 males and 4,579 females who attended Imperial College London in 2015 for general health screening. The UKBB baseline data included 423,803 CAU, 7,841 SA, and 5,455 AC individuals. T2D rates were compared to estimates from the Weqaya program on age-specific T2D prevalence in 50,138 Arab-Emiratis [284], and 14,940 Europeans and 3,291 Indian Asians (IA) from the DECODA study[293]. Error bars present 95% confidence intervals.



4.5.3 Basic characteristics of the EM2 sample population

I recruited 1,485 Arab-Emirati individuals in a case-control study design. Individuals with other types of

diabetes (Type 1 diabetes, gestational, MODY) were excluded (n = 37). Participants included 693 males and 755 females, whom I categorised into 3 groups: (1) NGT (n = 459), (2) prediabetes (n = 371), and (3) T2D (n = 618). Clinical characteristics of the EM2 participants are shown in **Table 4.2** for males and **Table 4.3** for females. The mean duration of T2D was 3.6 years, mean age of T2D onset was 34.5 years, and 66.5% of individuals with T2D were on insulin. **Figure 4.2**. flowchart describes the study design and data collection for EM2.

4.5.4 Comparison of T2D prevalence rates between Arab-Emiratis (EM1) and other ethnic groups

Data from EM1 (7,283 Arab-Emiratis) and UKBB (423,803 CAU, 7,841 SA, and 5,455 AC) individuals were used to compare age-specific T2D prevalence rates across ethnic groups. Individuals included either had NGT or T2D. Arab-Emirati males and females (EM1) were significantly younger than all other ethnic groups with significantly higher fasting plasma glucose (FBG) [mean (SD) FBG (mmol/L): 5.94 (2.24) in Arab-Emiratis, 5.07 (1.08) in CAU, 5.65 (2.04) in SA, and 4.98 (1.34) in AC, $p < 0.001$] (**Table 4.4**). Similar differences in FBG were observed when individuals with NGT were separately compared. Arab-Emiratis had significantly higher BMI compared to their CAU and SA counterparts only; AC males and females had significantly higher BMI than Arab-Emiratis [mean (SD) BMI (Kg/m^2): 28.5 (6.3) in Arab-Emiratis, 27.3 (4.7) in CAU, 26.9 (4.4) in SA, and 29.3 (5.4) in AC, $p < 0.001$] (**Table 4.4**).

The overall T2D prevalence rate in EM1 (14.7%, 95%CI: 14.6-14.8) was over 3-fold higher than that observed among CAU individuals (4.7%, 95%CI: 4.6-4.8) ($p < 0.001$). In addition, T2D rates were highest in Arab-Emiratis (EM1) across all ages studied ($p <0.001$), reaching a prevalence of 81.7% (95%CI: 81.6-81.8) in males and 79.5% (95%CI: 79.4-79.6) in females aged 61-70 years (**Supplementary Table 4.3**, **Figure 4.3**). The SA and AC groups had similar rates of T2D, reaching a prevalence of ~34-39% (males) and 24-30% (females) at 61-70 years. T2D prevalence was lowest in CAU males and females across all

ages studied, with the highest rates of 9.8% (95%CI: 9.7-10.0) and 4.9% (95%CI: 4.8-5.0) at 61-70 years, respectively (**Supplementary Table 4.3**, **Figure 4.3**).

4.5.5 Insulin, insulin resistance, and β-cell function in the EM2 sample population

As expected, fasting insulin levels were significantly higher in males and females with prediabetes or T2D, compared to individuals with NGT [mean (SD) insulin (pmol/L): males; 94.7 (65.3) with NGT 119.3 (60.6) with prediabetes, 141.9 (99.6) with T2D; females: 85.2 (51.0) with NGT, 121.7 (75.5) with prediabetes, 127.6 (89.0) with T2D] (**Tables 4.2 and 4.3**). Also, as expected, insulin resistance, as assessed by HOMA-IR, was significantly higher in individuals with T2D or prediabetes compared to those with NGT (**Tables 4.2 and 4.3, and Figures 4.4 and 4.5**). β-cell function as measured by HOMA-B was only significantly lower in individuals with T2D, but not those with prediabetes (**Tables 4.2 and 4.3**) (**Figures 4.4 and 4.5**).

4.5.6 Adiposity and glycaemic measures in EM2 sample population

Obesity rates differed between sexes and across diabetes status groups (NGT, prediabetes and T2D). In males, obesity rates were 31.8% in NGT, 44.8% in prediabetes, and 57.8% in T2D. Obesity rates in females were higher than those in males across all groups ($p < 0.001$): 46.2% in NGT, 64.5% in prediabetes, and 71.3% in T2D. Overall, all adiposity measures that I investigated here were significantly higher in individuals with prediabetes or T2D compared to those with NGT. That included higher BMI, waist circumference, WHR, and higher anthropometry-derived VAT volume (**Tables 4.2 and 4.3**) (**Figures 4.6 and 4.7**).

Differences in lipid measures between diabetes status groups differed across sexes. Compared to males with NGT, those with T2D had significantly lower HDL, LDL, and total cholesterol, but had significantly higher triglyceride levels (lipid-lowering medications data were not considered here) (**Table 4.2**). On the

other hand, females with NGT had significantly higher HDL and lower triglyceride levels compared to those with T2D (**Table 4.3**). No significant differences were observed for LDL and total cholesterol for females across diabetes status groups.

Figure 4.4 Box plots demonstrating differences in (A) fasting blood glucose (FBG), (B) haemoglobin A1c (HbA1c), (C) fasting insulin, (D) Homeostatic model assessment of insulin resistance (HOMA-IR), and (E) Homeostatic model assessment of β-cell function (HOMA-B) in males with normal glucose tolerance (NGT), from the EM2 sample population, compared to those with prediabetes and type 2 diabetes (T2D). The EM2 sample population was collected in a case-control study design and included 693 males: 171 with NGT, 349 with T2D, and 173 with prediabetes.

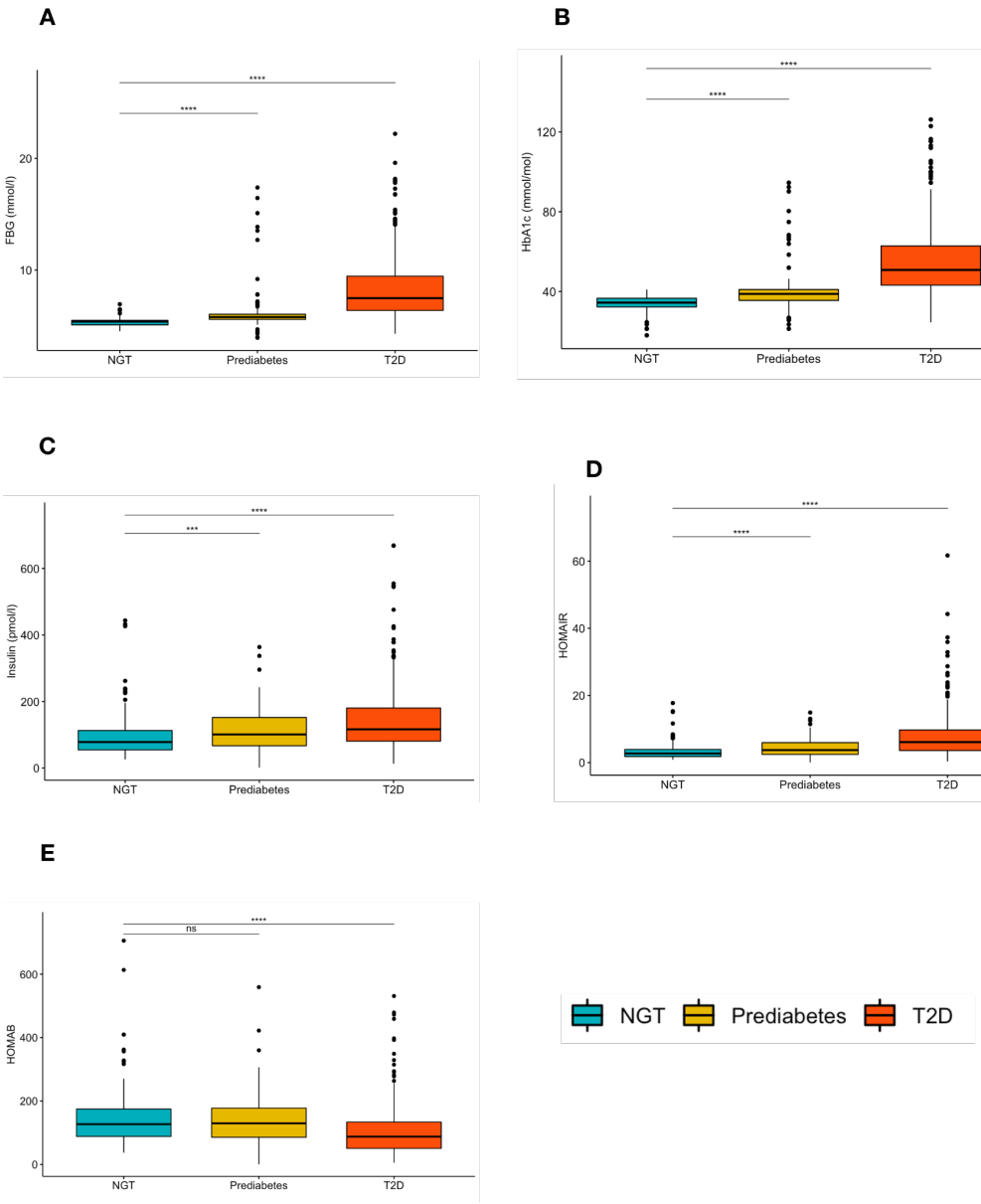
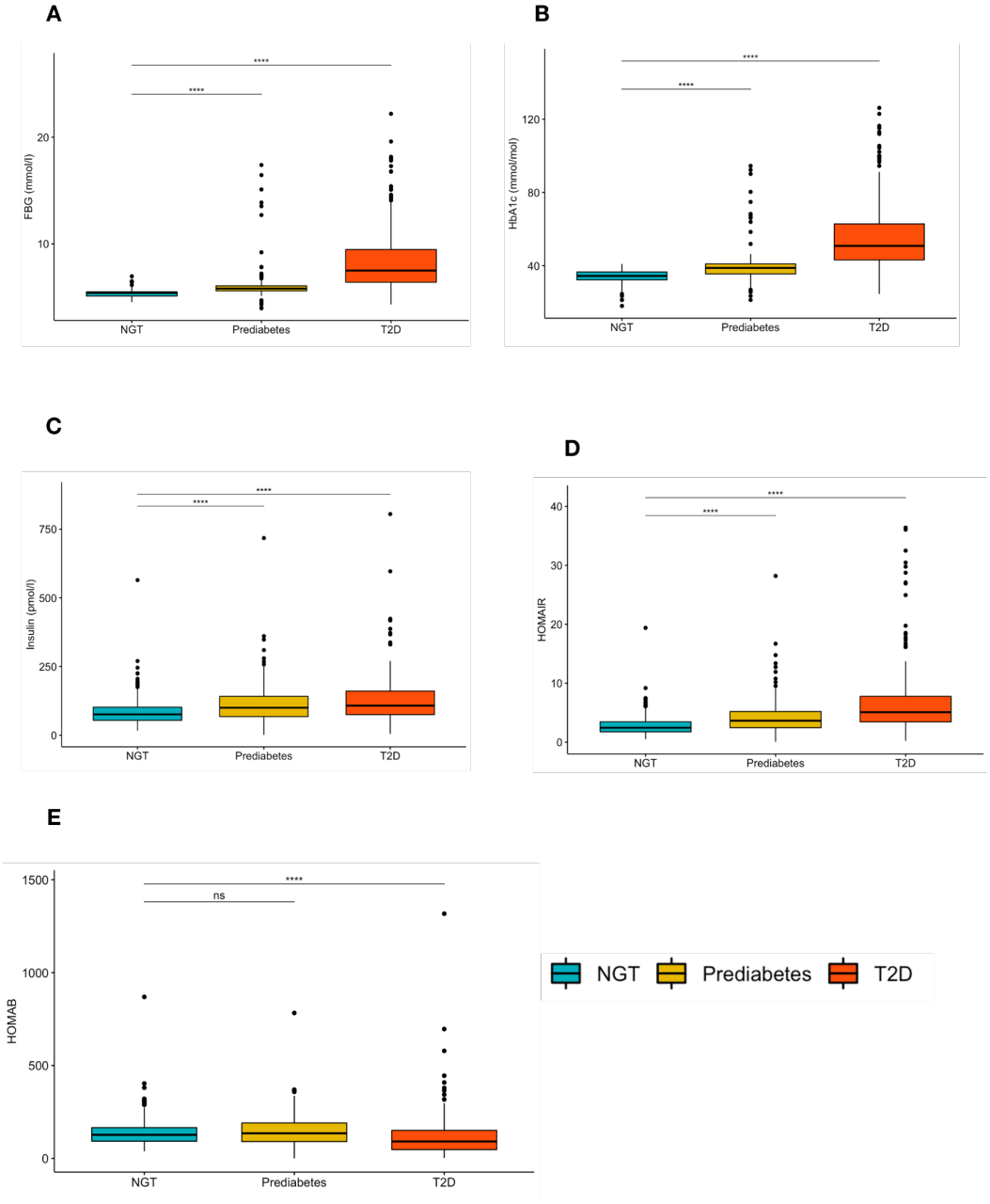


Figure 4.5. Box plots demonstrating differences in (A) fasting blood glucose (FBG), (B) haemoglobin A1c (HbA1c), (C) fasting insulin, (D) Homeostatic model assessment of insulin resistance (HOMA-IR), and (E) Homeostatic model assessment of β-cell function (HOMA-B) in females with normal glucose tolerance (NGT), from the EM2 sample population, compared to those with prediabetes and type 2 diabetes (T2D). The EM2 sample population was collected in a case-control study design and included 755 females: 288 with NGT, 269 with T2D, and 198 with prediabetes.



4.5.7 Contribution of adiposity to increased T2D risk and glycaemic disturbances in Arab-Emiratis compared to other ethnic groups

I present this section of results split into the different adiposity measures investigated. Waist and WHR measures were only available for EM2 and the UKBB (not EM1), and, thus, anthropometry-derived VAT volume was estimated in EM2 and UKBB only. As mentioned previously, insulin data were only available for EM2, and so the relationship of adiposity measures with insulin resistance and β-cell function was investigated in EM2 only.

4.5.7.1 Body-mass index

EM1 and UKBB data were used to perform trans-ethnic comparisons of overall and age-specific obesity prevalence rates. Overall obesity rates in males (BMI ≥ 30 Kg/m2) were significantly higher in Arab-Emiratis compared to all other ethnic groups [prevalence rates (95%CI): 33.4% (33.1-33.6) in Arab-Emiratis, 28.8% (27.5-30.0) in AC ($p < 0.001$), 17.6% in SA (16.7-18.5) ($p<0.001$), and 24.5% (24.3-24.7) in CAU ($p<0.001$)] (**Supplementary Table 4.4, Figure 4.8**). In addition, age-specific obesity rates were higher in Arab-Emirati males than their CAU and SA counterparts at all ages studied. Conversely, AC and Arab-Emirati males aged 51-70 years had similar obesity prevalence rates (no significant differences, $p > 0.05$).

On the other hand, overall obesity rates among Arab-Emirati females were significantly higher than CAU and SA females, but lower compared to their AC counterparts [prevalence rates (95%CI): 37.8% (37.6-38.0) in Arab-Emiratis, 44.6% (43.5-45.7) in AC ($p < 0.001$), 22.8% in SA (21.8-23.7) ($p<0.001$), and 22.1% (21.9-22.3) in CAU ($p<0.001$)] **(Supplementary Table 4.4, Figure 4.8**). Age-specific obesity rates in Arab-Emirati females, however, were significantly higher than all other ethnic groups at all ages, reaching a

prevalence of 62.1% (61.9-62.2) among those aged 61-70 years, compared to 23.0% (22.9-23.2), 23.1% (23.0-23.2), and 50.6% (50.5-50.8) among their CAU, SA, and AC counterparts, respectively.

Figure 4.6 Box plots demonstrating differences in (A) body mass index (BMI), (B) waist circumference (WC), (C) waist-to-hip ratio (WHR), and (D) anthropometry-derived visceral adipose tissue (VAT) in males with normal glucose tolerance (NGT), from the EM2 sample population, compared to those with prediabetes and type 2 diabetes (T2D). The EM2 sample population was collected in a case-control study design and included 693 males: 171with NGT, 349 with T2D, and 173 with prediabetes. VAT was estimated using the following formula:
VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

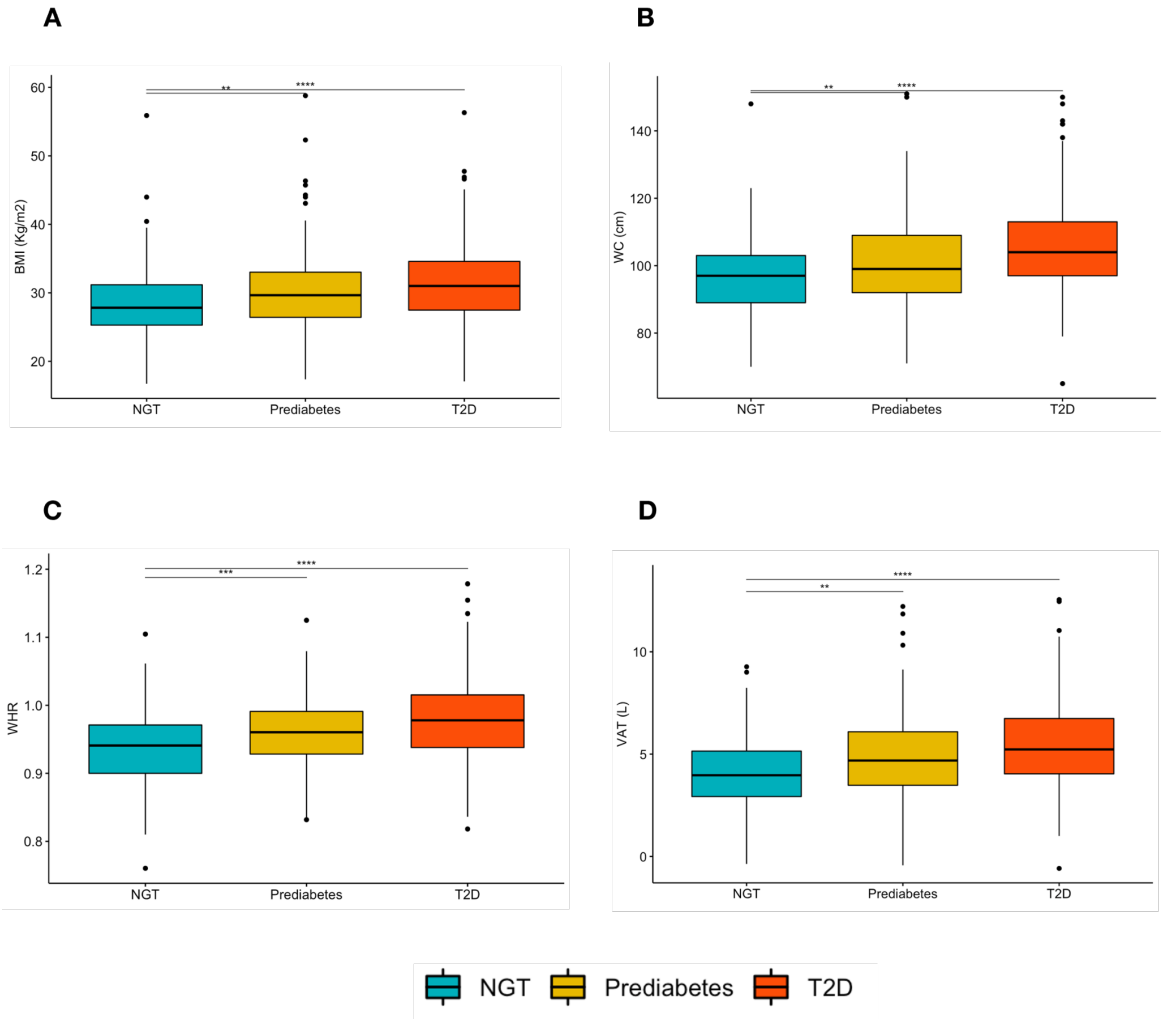


Figure 4.7 Box plots demonstrating differences in (A) body mass index (BMI), (B) waist circumference (WC), (C) waist-to-hip ratio (WHR), and (D) anthropometry-derived visceral adipose tissue (VAT) in females with normal glucose tolerance (NGT), from the EM2 sample population, compared to those with prediabetes and type 2 diabetes (T2D). The EM2 sample population was collected in a case-control study design and included 755 females: 288 with NGT, 269 with T2D, and 198 with prediabetes. VAT was estimated using the following formula: VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

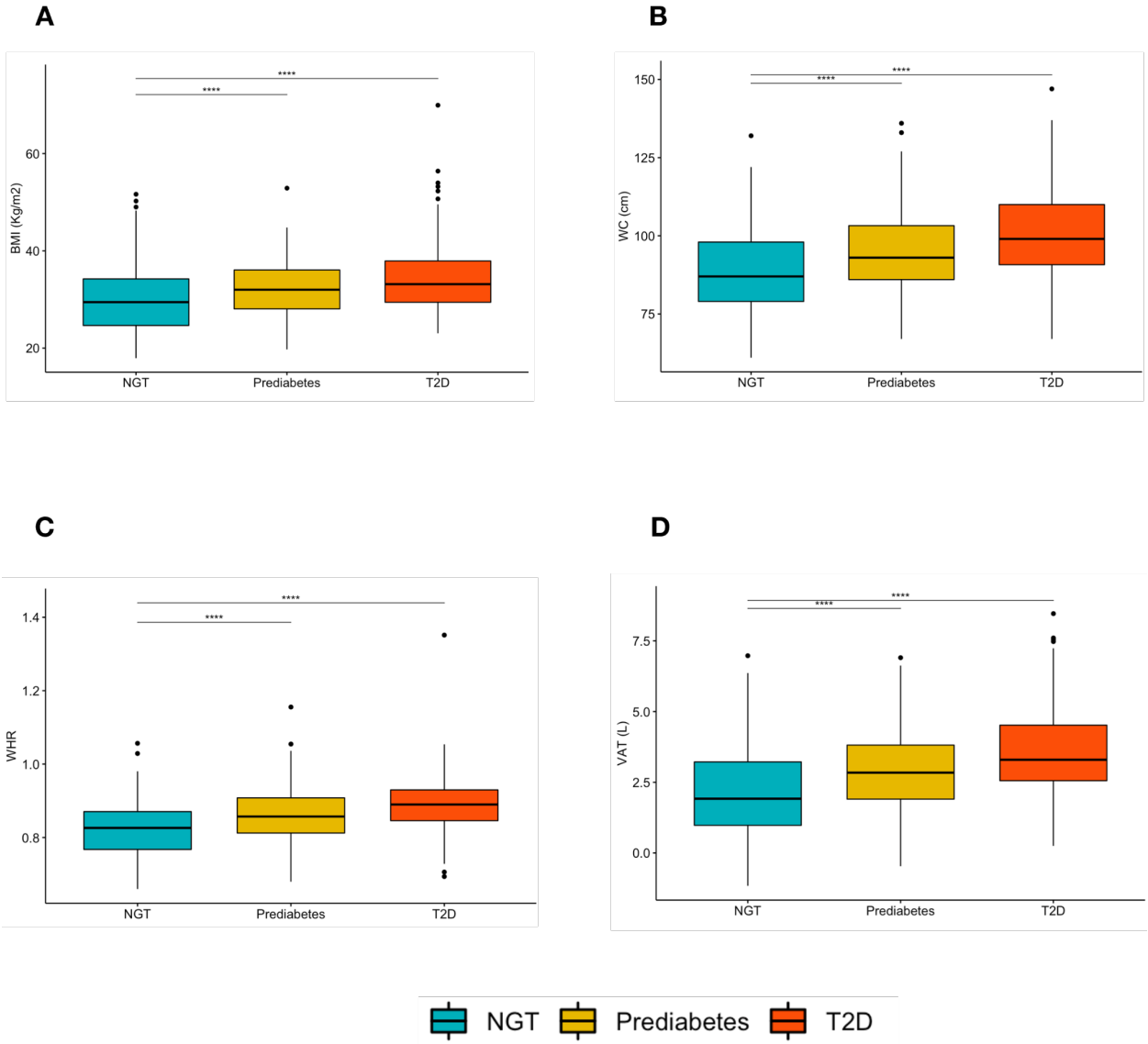
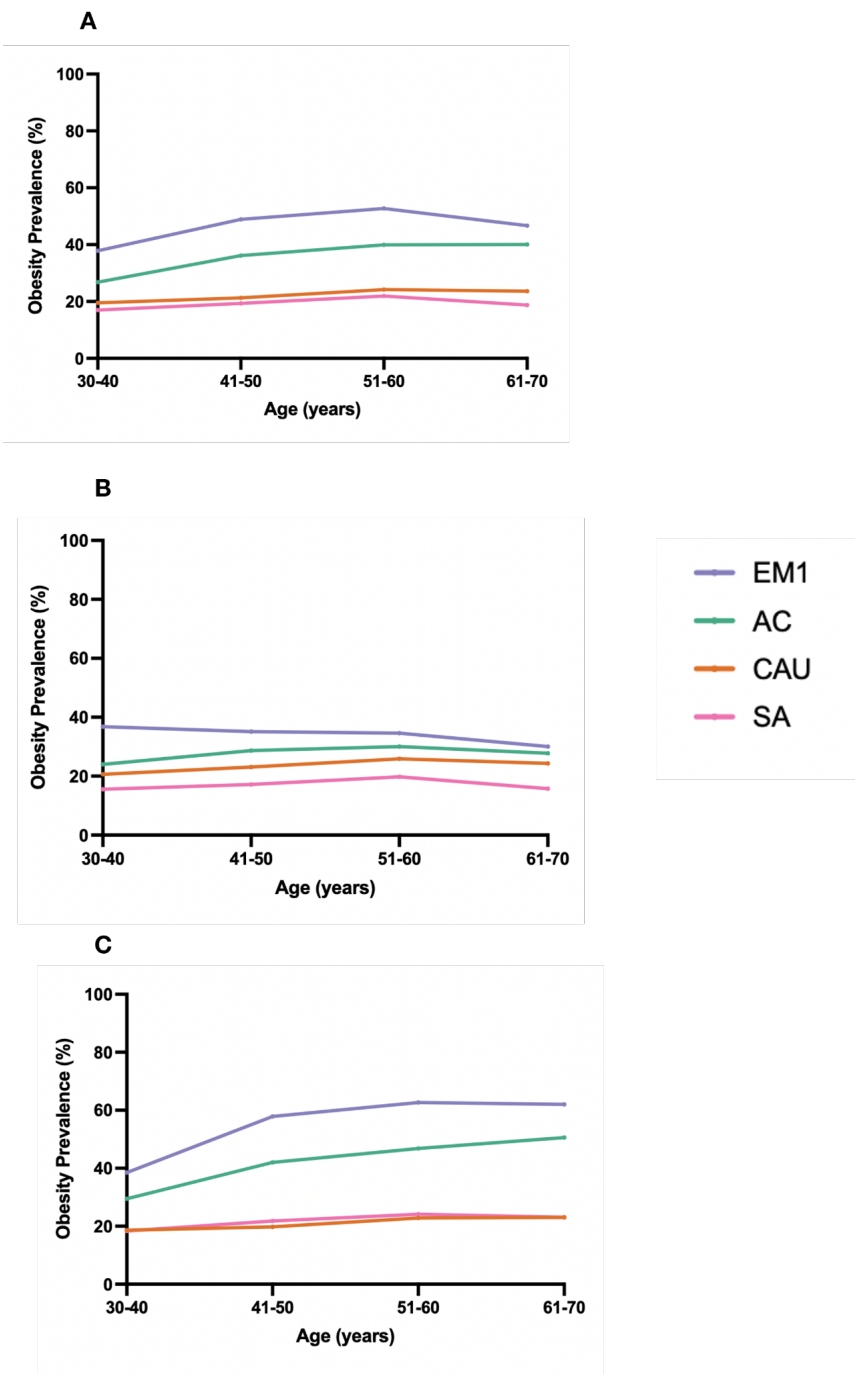


Figure 4.8. Line graphs demonstrating age-specific obesity prevalence rates in (A) all individuals, (B) males, and (C) females from the EM1 sample population compared to the UKBB cohort. The EM1 sample population included 2,704 Arab-Emirati males and 4,579 Arab-Emirati females who attended Imperial College London in 2015 for general health screening. The UKBB baseline data included 423,803 Caucasian (CAU), 7,841 South Asian (SA), and 5,455 Afro-Caribbean (AC) individuals.

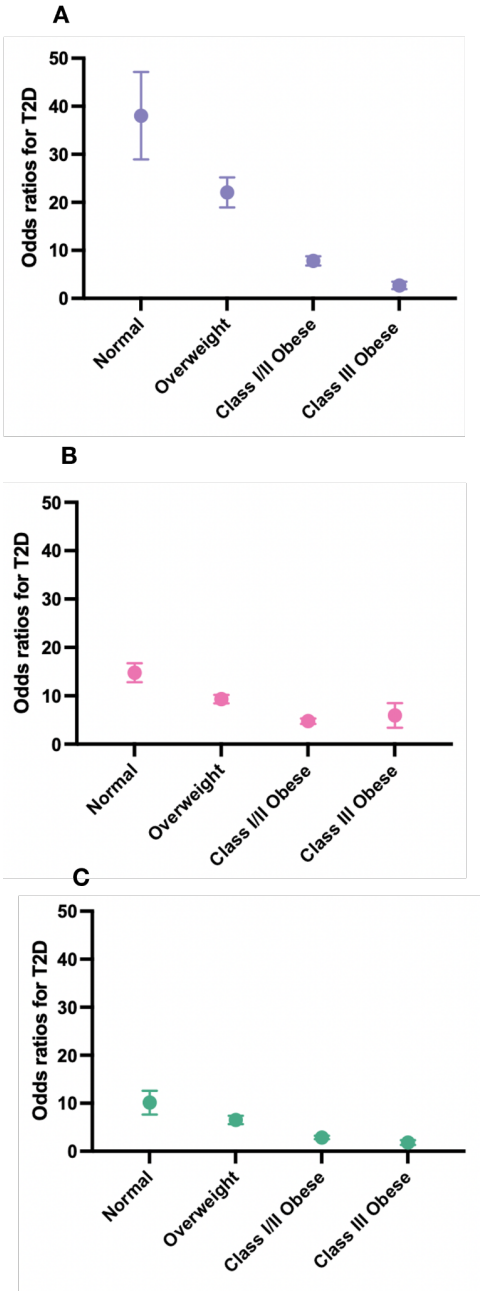


In order to test for the relationship of ethnicity with T2D risk, a multinomial regression analysis was performed with CAU as the reference category. Compared to CAU, age-adjusted ORs for T2D were 4.3-, 5.7-, and 11.9-fold higher in SA, AC, and Arab-Emirati (EM1) males, respectively (**Table 4.5**). In order to investigate the contribution of BMI towards ethnic differences in T2D risk, I adjusted for BMI. Ethnicity-specific ORs for T2D, relative to CAU (reference category), remained elevated even after adjustment for BMI (**Table 4.5**). Results were similar in females, with higher ORs for T2D, relative to reference category (CAU), particularly among Arab-Emirati females (**Table 4.5**).

Moreover, ORs for T2D were computed for Arab-Emirati (EM1), SA, and AC individuals at all levels of general obesity (normal weight, overweight, class I/II obese, and class III obese), relative to CAU individuals (reference category). ORs were higher among EM1, SA and AC individuals at all levels of general obesity, with most marked differences in EM1 and normal weight individuals overall (**Figure 4.9 and Table 4.6**). The population risk of T2D attributable to obesity (BMI ≥ 30 Kg/m2) was 33.4% and 32.4% in Arab-Emirati (EM1) males and females, respectively. This was lower than that computed for CAU males (44.3%) and females (54.7%) (**Table 4.7**).

The relationship of BMI with HOMA-IR and HOMA-B was only investigated in Arab-Emiratis (EM2). After adjustment for age, BMI was significantly associated with both HOMA-IR and HOMA-B in males and females (**Table 4.8**). This association persisted even after the adjustment for T2D status (**Table 4.8**). An increase in BMI by 1SD, increases HOMA-IR by 1.8 and 2.0 units in males and females, respectively.

Figure 4.9. Odds ratios for type 2 diabetes (T2D) risk in (A) Arab-Emirati (EM1), (B) South Asian (SA), and (C) Afro-Caribbean (AC) individuals compared to Caucasians (CAU) at every level of general obesity; (1) Normal (BMI = 18.0-24.9 Kg/m^2) (2) Overweight (BMI = 25.0-29.9 Kg/m^2), (3) Class I/II Obese (BMI = 30.0-39.9 Kg/m^2); and (4) Class III Obese (BMI ≥ 40.0 Kg/m^2). The EM1 sample population included 2,704 Arab-Emirati males and 4,579 Arab-Emirati females who attended Imperial College London in 2015 for general health screening. The UKBB baseline data included 423,803 CAU, 7,841 SA, and 5,455 AC individuals. Error bars present 95% confidence intervals.



4.5.7.2 Waist and WHR

EM2 and UKBB data were used to perform trans-ethnic comparisons of overall and diabetes status-specific central adiposity prevalence rates (assessed by WHR) and waist circumference and WHR measures.

Central adiposity rates in Arab-Emirati males (EM2) (WHR > 0.9) with either NGT or T2D were significantly higher than their AC counterparts only [NGT: 74.2% (74.0-74.3) in Arab-Emiratis and 48.1% (47.9-48.3) in AC ($p < 0.001$); T2D: 91.3% (91.2-91.4) in Arab-Emiratis and 79.6% (79.5-79.7) in AC ($p < 0.001$)] (**Supplementary Table 4.5**). The results were different in females: Arab-Emirati females with NGT had significantly higher central adiposity rates (measured by WHR > 0.8) than their CAU counterparts, but lower than their SA counterparts [61.1% (61.0-61.3) in Arab-Emiratis, 54.0 (53.9-54.2) in CAU (0.035), 68.2% (68.0-68.3) ($p = 0.032$), and 63.8 (63.6-63.7) ($p = 0.092$)]. No significant differences were observed across ethnic groups among females with T2D (**Supplementary Table 4.5**).

EM2 females with NGT had significantly higher waist circumference ($p <0.001$) and WHR ($p < 0.05$) than their CAU counterparts (**Table 4.9**) (**Figures 4.10**). Arab-Emirati females also had significantly higher waist circumference than SA females, irrespective of diabetes status ($p < 0.001$). On the other hand, no differences in either waist or WHR were observed between all Arab-Emirati and AC females with or without T2D. In addition, no significant differences in WHR were observed among all females with T2D across ethnic groups (**Table 4.9**) (**Figure 4.11**).

Different results were observed in males. Unlike females, Arab-Emirati males had significantly higher

waist circumference and WHR compared AC males ($p < 0.001$), irrespective of diabetes status. Additionally, Arab-Emirati males had significantly higher waist circumference, but similar WHR compared to SA males, irrespective of diabetes status. No differences were observed in waist or WHR between Arab-Emirati and CAU males with or without T2D (**Table 4.10 and Figures 4.12 and 4.13**).

After adjustment for differences in age, waist circumference and WHR ORs for T2D were similar across ethnic groups and sexes (**Table 4.11**). Despite that, the ORs of central adiposity as measured by WHR (WHR > 0.9 in males and WHR > 0.8 in females) for T2D were highest among Arab-Emirati males [OR (95%CI): 2.7 (2.3-3.3) in Arab-Emiratis, 1.2 (1.1-1.2) in CAU, 2.1 (1.9-2.3) in SA, and 2.5 (2.3-2.8) in AC], and females [OR (95%CI): 6.6 (5.4-8.0) in Arab-Emiratis, 3.3 (3.2-3.4) in CAU, 3.73 (3.3-4.2) in SA, and 5.4 (4.5-6.3) in AC] (**Table 4.7**). Furthermore, the T2D population risk (%) (adjusted for age) attributable to central adiposity was highest in Arab-Emirati males and females. After adjustment for BMI, T2D risk attributable to central adiposity was 55.3%, 11.2%, 45.9%, and 44.8% in Arab-Emirati, CAU, SA, and AC males. Overall, estimates were higher in females; 71.2%, 55.9%, 66.0%, 75.1% in Arab-Emirati, CAU, SA, and AC females (**Table 4.7**).

As with BMI, the association of waist circumference and WHR with HOMA-IR and HOMA-B was only investigated in Arab-Emiratis (EM2). After adjustment for age, BMI, and T2D status, waist and WHR were significantly associated with HOMA-IR in males ($p = 0.006$) and females ($p < 0.001$). An increase in waist circumference by 1SD increases HOMA-IR by 2.6 and 2.2 units in males and females, respectively. Association of waist circumference and WHR with HOMA-B appear to be dependent on BMI (**Table 4.8**).

Figure 4.10. Box plots demonstrating differences in adiposity measures including (A) body mass index (BMI), (B) Waist circumference, (C) waist-to-hip ratio (WHR), and (D) anthropometry-derived visceral adipose tissue (VAT) in females with normal glucose tolerance (NGT) from the EM2 sample population compared to the UK biobank (UKBB) cohort. The EM2 sample population was collected in a case-control study design. UKBB data included individuals of Caucasian (CAU), South (SA), and Afro-Caribbean (AC) descent. VAT was estimated using the following formula: VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

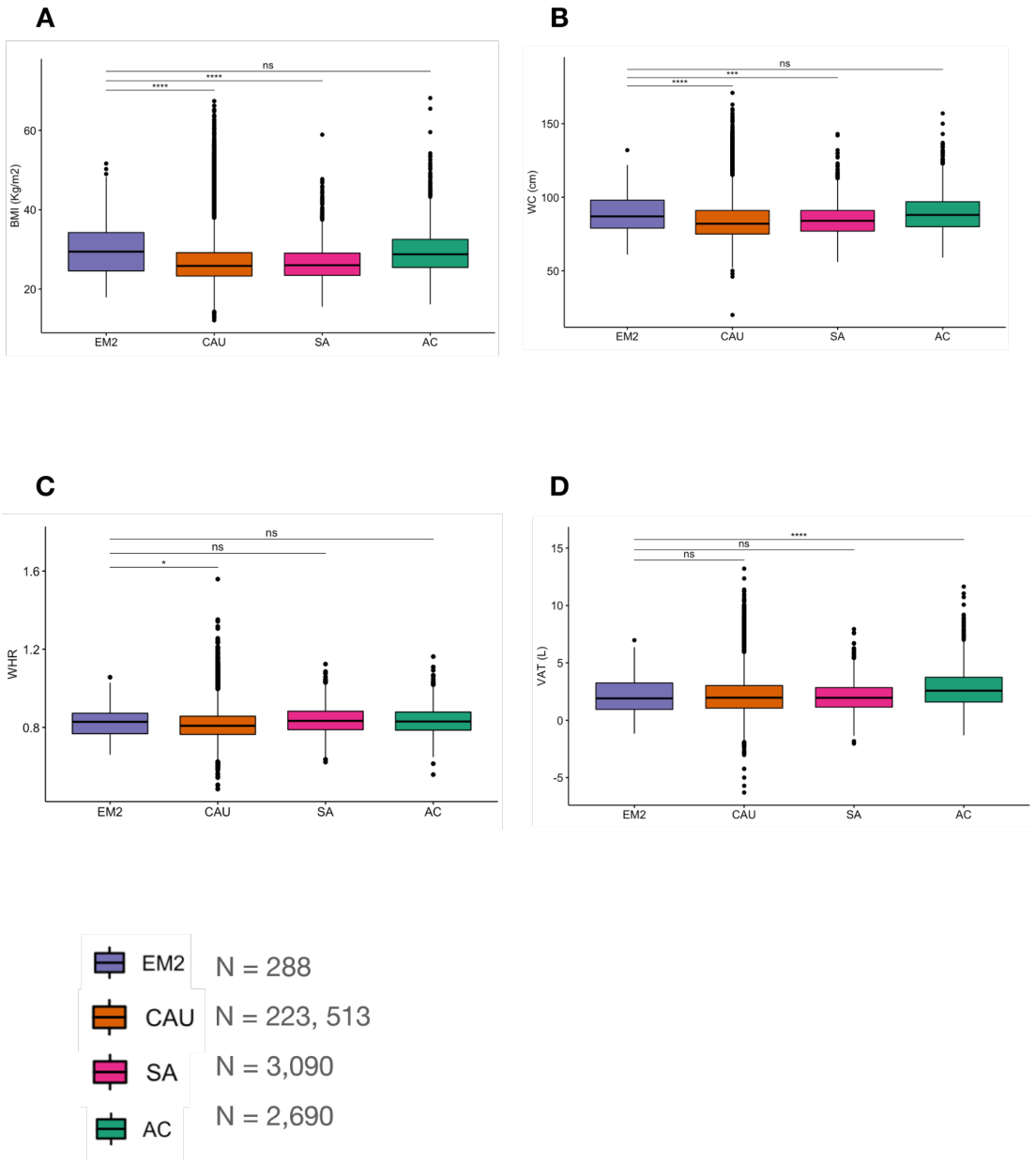
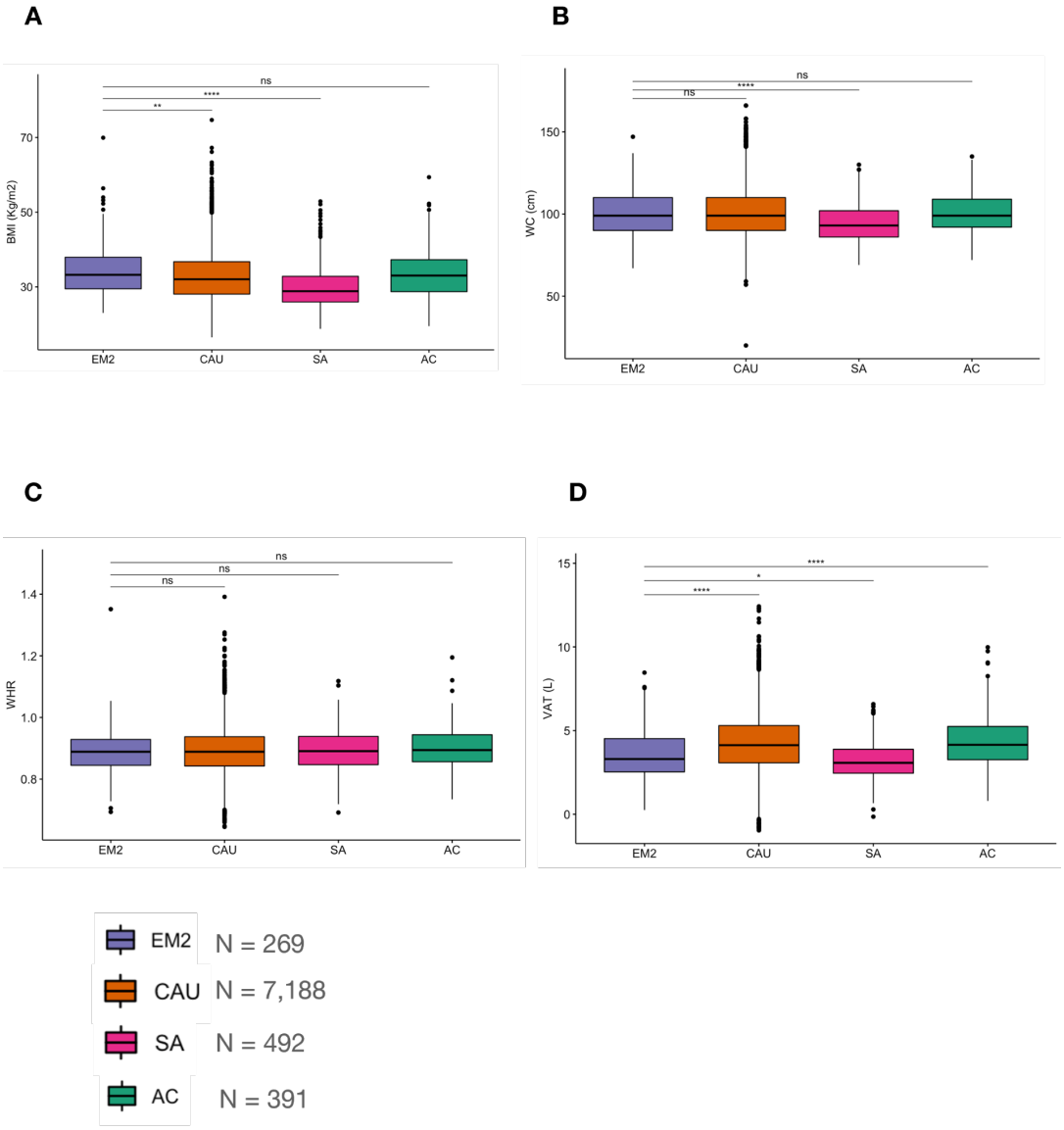


Figure 4.11. Box plots demonstrating differences in adiposity measures including (A) body mass index (BMI), (B) Waist circumference, (C) waist-to-hip ratio (WHR), and (D) anthropometry-derived visceral adipose tissue (VAT) in females with type 2 diabetes (T2D) from the EM2 sample population compared to the UK biobank (UKBB) cohort. The EM2 sample population was collected in a case-control study design. UKBB data included individuals of Caucasian (CAU), South (SA), and Afro-Caribbean (AC) descent. VAT was estimated using the following formula: VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).



4.5.7.3 Anthropometry-derived VAT volume

The anthropometry-derived VAT volume was significantly lower in Arab-Emirati males (EM2) compared to their CAU counterparts ($p < 0.001$), irrespective of diabetes status (**Table 4.10 and Figures 4.12 and 4.13**). No significant differences in VAT volume between Arab-Emirati, SA, and AC males with NGT were observed. Arab-Emirati males with T2D, however, had significantly higher VAT compared to SA males ($p < 0.001$). On the other hand, Arab-Emirati females with T2D, had significantly lower VAT compared to both their CAU and AC counterparts ($p < 0.001$) (**Table 4.9**, **Figure 4.11**). No significant differences observed in VAT volume among individuals with NGT.

After adjustment for age and BMI, VAT was significantly associated with T2D across all ethnic groups ($p < 0.001$), with the highest ORs among Arab-Emiratis, particularly in females [ORs (95% CI): 2.1 (1.9-2.3)] in Arab-Emiratis, 1.8 (1.7-1.8) in CAU, 1.8 (1.7-1.9) in SA, and 1.7 (1.6-1.7) in AC females] (**Table 4.11**). ORs for T2D were significantly higher among Arab-Emirati females compared to each other ethnic group ($p < 0.05$). ORs for T2D among males were significantly higher in Arab-Emiratis compared to CAU and SA ($p < 0.05$), but significantly lower than AC males.

As with waist circumference and WHR, the association of VAT with HOMA-IR and HOMA-B was only tested in Arab-Emiratis (EM2). After the adjustment for age, BMI, and T2D status, VAT was significantly associated with HOMA-IR in males ($p = 0.006$) and females ($p < 0.001$). An increase in VAT volume by 1SD, increases HOMA-IR by 2.7 and 2.2 units in males and females, respectively. Like waist and WHR, association of VAT with HOMA-B also appears to be dependent on BMI (**Table 4.8**).

Figure 4.12. Box plots demonstrating differences in adiposity measures including (A) body mass index (BMI), (B) Waist circumference, (C) waist-to-hip ratio (WHR), and (D) anthropometry-derived visceral adipose tissue (VAT) in males with normal glucose tolerance (NGT) from the EM2 sample population compared to the UK biobank (UKBB) cohort. The EM2 sample population was collected in a case-control study design. UKBB data included individuals of Caucasian (CAU), South (SA), and Afro-Caribbean (AC) descent. VAT was estimated using the following formula:

VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

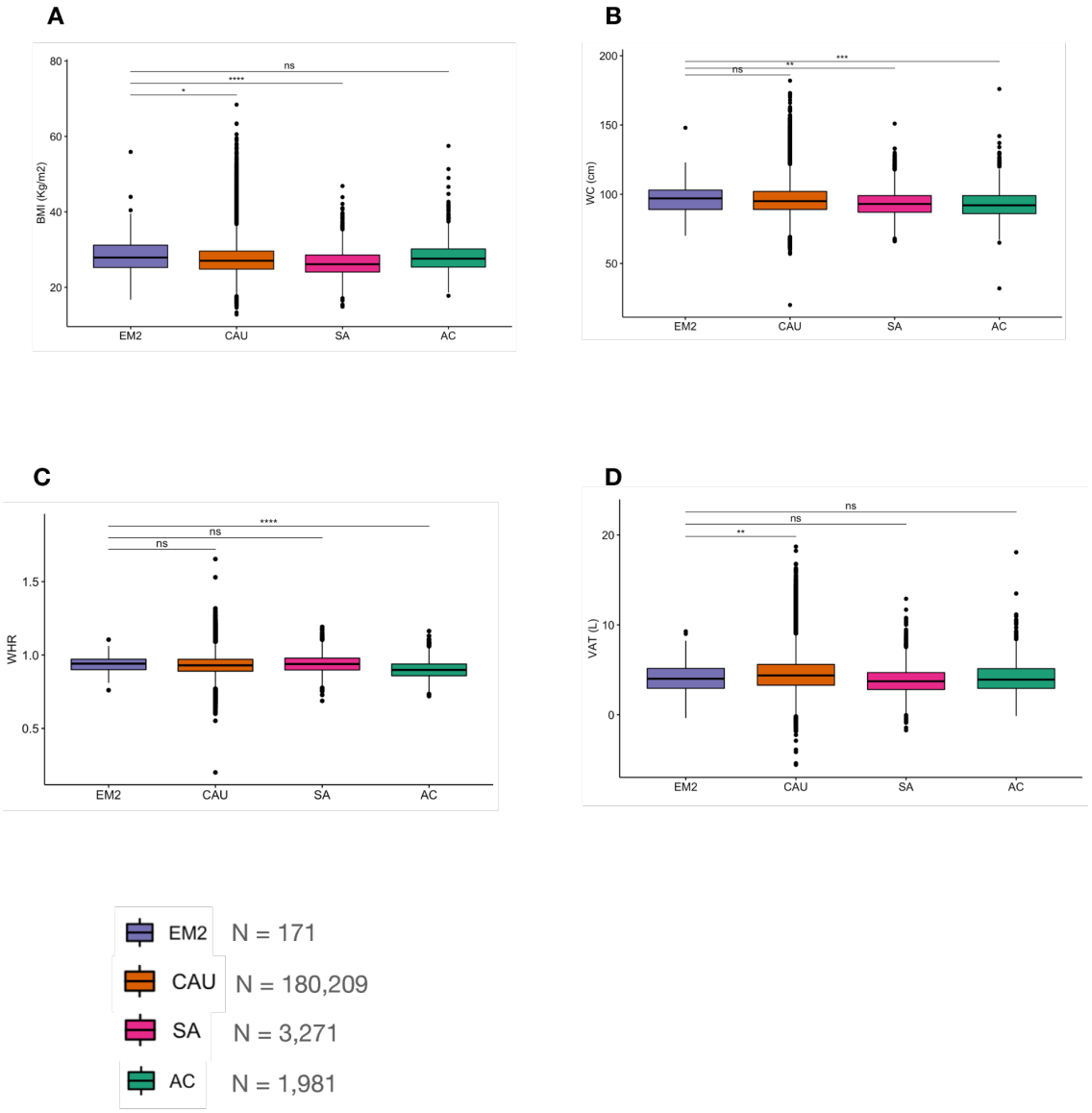
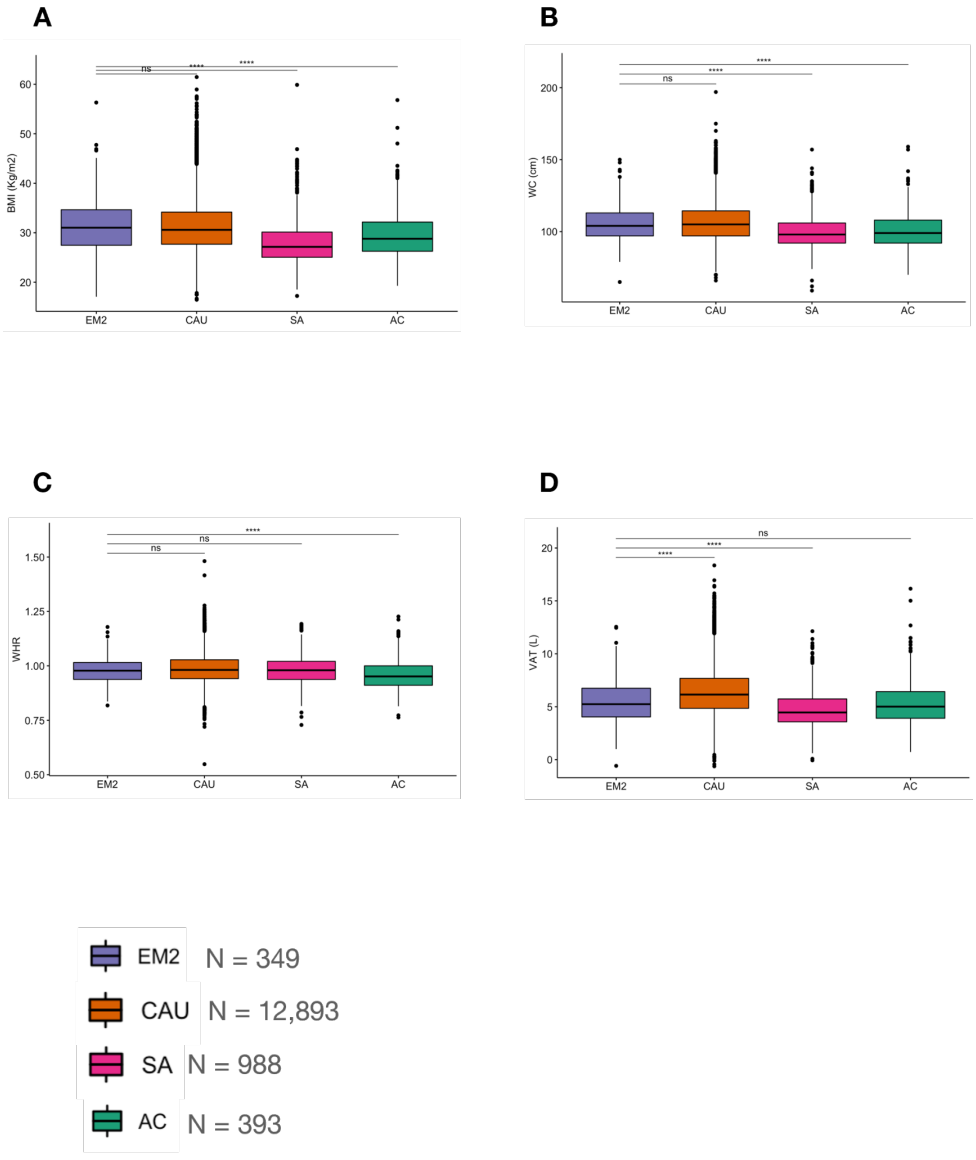


Figure 4.13. Box plots demonstrating differences in adiposity measures including (A) body mass index (BMI), (B) Waist circumference, (C) waist-to-hip ratio (WHR), and (D) anthropometry-derived visceral adipose tissue (VAT) in males with type 2 diabetes (T2D) from the EM2 sample population compared to the UK biobank (UKBB) cohort. The EM2 sample population was collected in a case-control study design. UKBB data included individuals of Caucasian (CAU), South (SA), and Afro-Caribbean (AC) descent. VAT was estimated using the following formula: VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).



4.5.8 Sensitivity analyses

Given that the diagnosis of T2D among UKBB participants was based on self-reported diabetes and age at diagnosis (Chapter 2, section 2.2.3.1), this might have led to an underestimation of T2D prevalence rates in this cohort. However, my sensitivity analysis showed that none of the individuals originally defined as having NGT had HbA1c > 47.9 mmol/mol (i.e. all are healthy with no diabetes).

Several factors could be confounding the associations presented in this chapter, including diet, socioeconomic status, and smoking. Unfortunately, data on dietary factors and socioeconomic status in EM1, EM2, or UKBB were unavailable for analyses. However, further adjustments for smoking status have not affected the associations reported in this part of the study (**Supplementary Tables 4.7 and 4.8**). Furthermore, given the detrimental influence of menopausal transition (MT) on women's body composition[309] , I performed stratified analyses on females undergoing MT in EM2 (i.e., females aged 45-55 years, n = 62) (**Supplementary Table 4.9**). I found that the associations of adiposity measures with T2D or HOMA-IR in this subgroup were like those observed in other females, albeit slightly weaker (**Supplementary Table 4.9**).

4.6 DISCUSSION

The results I present in this chapter confirm the large ethnic variation in T2D prevalence rates, with a striking 12- to 14-fold increased T2D risk among Arab-Emiratis compared to Europeans (CAU in the UKBB), which cannot be explained by differences in BMI between the two populations. Results presented here also indicate that, unlike BMI, central adiposity measures may have a bigger contribution to this excess risk. Firstly, the T2D risk that is attributable to central adiposity as measured by WHR was higher in Arab-Emiratis, even after adjustment for differences in BMI. Secondly, central adiposity prevalence rates were higher in Arab-Emiratis than in Europeans. Thirdly, the association of

anthropometry-derived VAT volume with T2D was stronger among Arab-Emiratis than in all other ethnic groups (Europeans, South Asians, and Afro-Caribbeans) studied, particularly in females.

The Arab-Emiratis were younger than all other ethnic groups studied. After considering the differences in age, Arab-Emirati males and females had a 12- and 20-fold higher risk of T2D compared to Europeans, respectively. This risk decreased for females to 14-fold after adjusting for differences in BMI between the two populations, highlighting the contribution of increased BMI to the excess T2D risk among Arab-Emirati females.

As in Europeans, I demonstrate that BMI, waist circumference, and WHR are independent risk factors for T2D among Arab-Emiratis, with similar effect sizes to Europeans and other ethnic groups. My results also show ~35-fold excess risk of T2D among lean Arab-Emirati individuals compared to Europeans (BMI <25 Kg/m^2). This excess risk decreases in overweight individuals, and even further in class I/II obese and class III obese individuals. Arab-Emiratis with class III obesity had similar T2D risk to that of Europeans with class III obesity. Furthermore, the T2D risk in Europeans attributable to BMI ≥ 30 Kg/m^2 was higher than that observed among Arab-Emiratis. Altogether, these results challenge the role of obesity (as assessed by BMI) in explaining the increased risk of T2D in Arab-Emiratis.

The limitations of BMI must be acknowledged when interpreting these results. Although BMI is a simple non-invasive method for assessing obesity within populations, people with similar BMI may have heterogeneous obesity status, with remarkable differences in their comorbidities and health risks[310, 311] . Relying on BMI as a measure of obesity could in fact lead to misclassification of cardiometabolic health risks[312]. Failure of BMI to account for the increased risk of T2D amongst Arab-Emiratis could represent its reduced sensitivity to fully capture the complex distribution of fat as well as the extent of adipocyte cell

hypertrophy.

Not all fat depots are equally detrimental for health, and their anatomical location makes a substantial difference. Visceral fat and ectopic fat accumulation in or around the liver, pancreas, and muscles can disrupt glucose-insulin metabolism, leading to the development of hepatic insulin resistance[313] and β-cell dysfunction[314]. Effective fat storage in subcutaneous adipose tissue, on the other hand, can separate the potential detrimental lipids from ectopic fat storage or reduce lipotoxicity in cells or tissues lacking effective fat storage capacity[101]. These divergent regulatory functions of visceral and subcutaneous fat depots may explain why some people are protected against diseases such as T2D.

Although waist and WHR indices cannot distinguish between visceral and abdominal subcutaneous fat depots, they were more reliable than BMI in evaluating overall fat distribution between ethnic groups[67, 68]. Overall, the results presented here highlight the potential contribution of differences in fat distribution between Arab-Emiratis and Europeans in explaining differences in the magnitude of T2D risk. After considering differences in BMI, central adiposity as measured by WHR was found to be attributable to 55.3% and 71.2% of T2D among Arab-Emirati males and females, compared to 11.2% and 55.9% in European males and females, respectively. The low estimate in European males suggests that the association of central adiposity with T2D risk in this subgroup is dependent on BMI (PAR % before BMI adjustment was 67% in European males). My results also show that anthropometry-derived VAT volume is an independent risk (sex-stratified, age and BMI adjusted) factor for T2D across all ethnic groups, with bigger effect sizes among Arab-Emiratis.

Despite these results, anthropometry-derived VAT levels were generally higher in Europeans in this study. This could be due to different reasons. Firstly, the equation I used here was derived and validated

in Europeans only, and whether it is applicable to other ethnic groups is unknown. Secondly, it is possible that Arab-Emiratis store fat more ectopically (around the liver, heart, pancreas, muscle) rather than deep inside their abdomen (visceral). The UKBB yields valuable trans-ethnic data including MRI-measured fat depots to investigate differences in body fat deposition and ectopic fat content between ancestry groups. Recently, Alenaini et al. assessed MRI data from the UKBB and Hammersmith cohorts of individuals of South Asian, African, and European descent[315]. Their results, however, presented a limited number of ethnic differences in the distribution of MRI-measured body fat depots (VAT and liver fat). Consistent with my anthropometry-derived findings, they also found that Europeans presented higher VAT/(abdominal subcutaneous adipose tissue + VAT) values compared to individuals of South Asian or African descent[315]; an effect that would seem to be at odds with the increased susceptibility to metabolic disease observed in South Asian populations[213]. Altogether, these results call for investigations of alternative mechanisms to explain ethnic disparity in metabolic risk. This will be discussed more in Chapter seven.

Both insulin resistance and pancreatic β-cell function can lead to impaired insulin secretion and are important in the pathogenesis of T2D[23]. Unfortunately, insulin data was not available in the UKBB cohort and, thus, investigating ethnic differences was not possible. However, compared to data from the London Life Sciences Prospective Population Study (LOLIPOP) [283], Arab-Emiratis with NGT had 2- and 3-fold higher insulin resistance (HOMA-IR) than that observed among Indian Asians and Europeans, respectively. When comparing the association of adiposity measures with HOMA-IR among Arab-Emiratis in this study to results based on the LOLIPOP cohort in Europeans and South Asians, I found: (1) the association of BMI and waist circumference with HOMA-IR was strongest in Europeans and lowest in South Asian men and women, but (2) the association of WHR was strongest among Arab-Emiratis, compared to Europeans and South Asians[283]. These observations, however, need to be validated in a

European, South Asian, and Arab-Emirati trans-ethnic analysis.

I also showed here that BMI is strongly and positively associated with β-cell function. Data in the literature support these results: although T2D is characterised by decreased pancreatic β-cell mass and function (which is also observed here), in non-diabetic obesity, an expansion in β-cell mass occurs to provide sufficient insulin to prevent hyperglycaemia[316]. In fact, obesity has been shown to increase basal and postprandial β-cell insulin secretion, even in the absence of insulin resistance[317].

The results presented in my study so far suggest an important role of unfavourable fat distribution in Arab-Emiratis, as reflected by waist and WHR in explaining the increased risk of T2D better than BMI. These findings call for future investigations employing state-of-the-art imaging techniques to fully characterise fat distribution and ectopic deposition, allowing better delineation of the contribution of adiposity to the increased risk of T2D in the UAE population. My findings also make a case for a possible increased predisposition (genetic/epigenetic) to insulin resistance and β-cell dysfunction underlying the increased susceptibility to T2D in Arab-Emiratis and provide a rationale for genome-wide association studies using (in addition to T2D) HOMA-IR and HOMA-B phenotypes.

Although the representativeness of EM1 of the general UAE population has been evaluated, it is also important to consider the "healthy volunteer bias" that the UKBB cohort is subject to, which could have led to overestimating the increased risk of T2D in Arab-Emiratis. With that said, the overall obesity and diabetes rates observed in the UKBB were very similar to those reported in more representative samples/studies like IDF and NCD-RisC[3, 205]. In addition, I reported similar results when comparing T2D risk in Arab-Emiratis (EM1) and Europeans in the LOLIPOP study (Abstract 1547-P, American Diabetes Association Conference 76th Session, 2016).

In conclusion, I show that there is a 12- to 14-fold excess risk of T2D among Arab-Emiratis compared to Europeans, which is independent of differences in BMI. In addition to being simple and clinically-applicable measuring tools, waist and WHR appear to contribute more to differences in T2D risk across ethnic populations and highlight the potential characterisation of unfavourable fat distribution in the Arab-Emirati population. There is indeed a growing interest and need for fast, more reliable, and clinically-applicable tools to monitor specific fat depots (e.g., visceral, ectopic), for improved risk stratification in obese (and non-obese) individuals. The gold standard methods for fat depots quantification (MRI or CT) are costly and dependent on scanner availability. Thus, deriving and validating Arab-Emirati-specific (e.g., VAT volume) estimation formulas would be valuable for future research and clinical practice in the UAE.

Table 4.1. Basic characteristics of individuals in the EM1 sample population. EM1 included 2,704 Arab-Emirati males and 4,579 females who attended Imperial College London in 2015 for general health screening.

	Males			
N	**2,704**			
	Mean	**SD**	**Median**	**IQR**
Age (years)	37.07	14.56	34.00	27.0-44.0
BMI (Kg/m^2)	28.27	5.70	27.50	24.5-31.3
HbA1c (mmol/mol)	38.20	13.69	34.43	31.2-38.8
TC (mmol/l)	4.55	0.98	4.53	3.9-5.1
LDL (mmol/l)	3.07	0.91	3.05	2.4-3.7
HDL (mmol/l)	1.19	0.31	1.16	1.0-1.4
Triglycerides (mmol/l)	1.31	0.86	1.09	0.6-1.2
SBP (mmHg)	239.66	6115.05	120.00	112.0-131.0
DBP (mmHg)	72.15	14.73	72.00	66.0-77.3
	Females			
N	**4,579**			
	Mean	**SD**	**Median**	**IQR**
Age (years)	35.11	13.51	32	25.0-41.0
BMI (Kg/m^2)	28.56	6.67	27.78	23.8-32.7
HbA1c (mmol/mol)	35.85	11.4	33.33	30.1-36.6
TC (mmol/l)	4.46	0.84	4.38	3.9-5.0
LDL (mmol/l)	2.8	0.76	2.73	2.3-3.3
HDL (mmol/l)	1.46	0.35	1.43	1.2-1.7
Triglycerides (mmol/l)	1	0.69	0.83	0.8-1.6
SBP (mmHg)	114.6	15.41	112	104.0-122.0
DBP (mmHg)	65.5	12.38	64	58.0-71.0

Abbreviations: BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, SBP; systolic blood pressure, DBP; diastolic blood pressure

Table 4.2 Adiposity measures, glycaemic profile, lipid profile, and liver and kidney functions in 693 Arab-Emirati males with NGT, prediabetes, or T2D from the EM2 sample population. The EM2 sample population was collected in a case-control study design.

	NGT				Prediabetes					T2D				
N	**171**				**173**					**349**				
	Mean	**SD**	**Median**	**IQR**	**Mean**	**SD**	**Median**	**IQR**	**P**	**Mean**	**SD**	**Median**	**IQR**	**P**
Age (years)	36.74	5.34	36.77	33.0-40.2	39.67	5.91	38.77	35.6-43.0	<0.001	37.24	5.39	38.03	34.0-41.3	0.068
SBP (mmHg)	118.39	12.45	117.00	111.0-127.0	122.05	13.57	120.00	113.5-130.0	0.049	124.80	14.57	123.00	114.0-135.0	<0.001
DBP (mmHg)	74.50	9.04	75.00	68.0-79.0	77.73	8.45	76.00	72.0-84.0	0.013	77.96	9.42	78.00	71.0-84.0	<0.001
Adiposity														
BMI (Kg/m^2)	28.29	4.96	27.83	25.3-31.2	30.61	5.91	29.76	26.7-33.2	0.005	31.58	5.49	31.01	27.5-34.6	<0.001
Weight (Kg)	84.72	16.49	83.70	73.3-93.3	91.96	18.75	89.70	80.0-100.0	0.006	94.00	18.32	91.00	80.5-104.1	<0.001
Waist (cm)	96.43	11.58	97.00	89.0-103.0	102.30	13.36	100.00	93.0-110.0	0.004	105.49	13.53	104.00	97.0-113.0	<0.001

Hip (cm)	103.00	10.88	102.00	95.0-108.2	106.37	11.55	104.00	99.0-112.5	0.112	108.05	12.29	106.00	99.0-114.0	<0.001
WHR	0.94	0.05	0.94	0.90-0.97	0.96	0.05	0.96	0.93-0.99	<0.001	0.98	0.06	0.98	0.94-1.0	<0.001
VAT (L)	4.01	1.85	3.97	2.9-5.1	5.06	2.12	4.89	3.6-6.2	0.002	5.48	2.12	5.23	4.0-6.7	<0.001
Glycaemic profile														
Glucose (mmol/L)	5.37	0.35	5.39	5.1-5.5	5.83	0.48	5.80	5.6-6.0	<0.001	8.37	3.01	7.49	6.4-9.5	<0.001
HbA1c (mmol/mol)	33.72	4.19	34.43	32.2-36.6	37.24	4.64	37.71	35.3-39.9	<0.001	55.17	17.95	50.82	43.2-62.8	<0.001
Insulin (pmol/L)	94.66	65.28	78.20	54.6-112.9	119.31	60.56	106.50	72.3-158.0	<0.001	141.93	99.56	116.35	80.8-180.5	<0.001
HOMA-IR	3.31	2.45	2.72	1.8-3.9	4.51	2.49	3.80	2.6-6.0	<0.001	7.84	6.87	6.12	3.6-9.7	<0.001
HOMA-B	145.65	90.15	127.11	88.3-174.7	149.96	77.05	133.62	91.9-182.0	0.901	104.90	79.86	87.74	50.6-133.6	<0.001
Lipid profile														
TC (mmol/L)	4.77	0.82	4.68	4.2-5.3	4.93	0.83	4.98	4.4-5.6	0.670	4.38	1.20	4.22	3.5-5.0	<0.001
LDL (mmol/L)	3.28	0.77	3.25	2.7-3.7	3.43	0.79	3.41	2.8-4.0	0.146	2.89	1.04	2.72	2.1-3.5	<0.001

HDL (mmol/L)	1.28	0.31	1.26	1.1-1.5	1.23	0.32	1.20	1.0-1.4	0.820	1.06	0.25	1.05	0.9-1.2	<0.001
Trig (mmol/L)	1.19	0.87	0.94	0.7-1.3	1.37	0.83	1.16	0.9-1.6	<0.001	1.98	1.98	1.54	1.1-2.2	<0.001
Liver & kidney function														
AST (U/L)	22.74	8.65	21.90	17.9-25.6	25.62	12.88	23.70	18.9-28.1	0.080	24.63	16.10	21.15	17.6-26.5	0.986
ALT (IU/L)	26.57	14.37	22.60	17.3-31.5	33.55	20.28	28.60	21.3-41.7	0.002	35.95	21.10	29.95	22.0-42.6	<0.001
Bilirubin (umol/L)	11.15	6.83	9.00	6.5-13.6	8.68	4.60	7.60	5.8-10.9	0.002	8.96	5.19	7.70	5.7-10.4	0.001
Creatinine (umol/L)	80.39	12.10	79.10	71.7-86.8	81.21	12.46	80.35	72.8-88.5	0.871	77.05	23.79	73.90	66.9-82.0	<0.001
Urea (mmol/L)	4.83	1.25	4.71	4.0-5.7	4.74	1.09	4.61	4.0-5.2	1.000	4.77	1.58	4.58	3.8-5.4	0.594

Abbreviations: NGT; normal glucose tolerance, T2D; type 2 diabetes, BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, SBP; systolic blood pressure, DBP; diastolic blood pressure, HOMA-IR; homeostatic model assessment of insulin resistance, HOMA-B; homeostatic model assessment of β-cell function, AST; aspartate transaminase, ALT; alanine transaminase.

P: p values for Wilcoxon rank-sum test performed comparing parameters in prediabetes and T2D with NGT. Adjusted for multiple testing.

Table 4.3. Adiposity measures, glycaemic profile, lipid profile, and liver and kidney functions in 755 Arab-Emirati females with NGT, prediabetes, or T2D from the EM2 sample population. The EM2 sample population was collected in a case-control study design.

	NGT				Prediabetes					T2D				
N	**288**				**198**					**269**				
	Mean	**SD**	**Median**	**IQR**	**Mean**	**SD**	**Median**	**IQR**	**P**	**Mean**	**SD**	**Median**	**IQR**	**P**
Age (years)	37.51	6.15	36.91	33.4-41.1	39.65	6.27	39.42	35.6-43.1	0.016	38.38	5.59	38.86	35.2-42.5	0.016
SBP (mmHg)	110.74	14.28	109.00	101.0-118.0	116.64	16.43	116.00	105.0-123.0	<0.001	119.19	12.90	119.00	111.0-128.2	<0.001
DBP (mmHg)	66.22	9.27	65.00	59.0-73.0	69.86	10.18	70.00	64.0-75.0	<0.001	70.93	9.24	71.00	64.0-78.0	<0.001
Adiposity														
BMI (Kg/m^2)	29.94	6.47	29.45	24.7-34.2	32.42	5.62	32.44	28.1-36.1	<0.001	34.30	6.89	33.15	29.4-37.9	<0.001
Weight (Kg)	75.61	17.48	73.50	62.0-86.8	81.76	14.64	81.40	70.5-91.1	<0.001	86.30	18.67	82.20	73.0-95.7	<0.001
Waist (cm)	87.95	13.28	87.00	79.0-98.0	94.77	12.86	93.00	86.0-103.0	<0.001	100.58	13.85	99.00	90.8-110.0	<0.001

Hip (cm)	106.61	13.15	105.00	96.0-115.0	110.99	11.38	110.00	103.0-117.0	<0.001	113.06	13.87	111.00	104.0-121.0	<0.001
WHR	0.82	0.07	0.83	0.77-0.87	0.86	0.08	0.86	0.81-0.90	<0.001	0.89	0.07	0.89	0.85-0.93	<0.001
VAT (L)	2.10	1.54	1.92	1.0-3.2	2.96	1.48	2.88	2.0-3.8	<0.001	3.58	1.57	3.30	2.6-4.5	<0.001
Glycaemic profile														
Glucose (mmol/L)	5.22	0.31	5.20	5.0-5.4	5.73	0.54	5.70	5.4-6.0	<0.001	7.79	2.81	7.01	5.9-8.7	<0.001
HbA1c (mmol/mol)	33.73	3.83	34.43	31.2-36.6	37.45	4.62	37.71	34.4-40.2	<0.001	52.82	16.26	49.73	41.0-60.7	<0.001
Insulin (pmol/L)	85.21	50.99	75.82	54.0-102.0	121.71	75.50	105.50	72.9-149.6	<0.001	127.63	88.99	107.70	75.0-160.7	<0.001
HOMA-IR	2.85	1.75	2.44	1.8-3.5	4.56	3.17	3.79	2.6-5.4	<0.001	6.55	5.75	5.08	3.4-7.9	<0.001
HOMA-B	140.70	76.65	126.99	92.6-165.7	158.18	85.81	138.58	100.4-196.8	0.403	116.64	116.55	91.09	47.4-150.9	<0.001
Lipid profile														
TC (mmol/L)	4.63	0.92	4.50	4.1-5.1	4.79	0.77	4.73	4.3-5.4	0.031	4.61	1.00	4.56	4.0-5.2	0.816
LDL (mmol/L)	2.99	0.80	2.91	2.5-3.4	3.17	0.74	3.12	2.6-3.7	0.028	3.06	0.93	3.01	2.4-3.6	0.334

HDL (mmol/L)	1.52	0.36	1.48	1.3-1.7	1.39	0.31	1.31	1.2-1.6	<0.001	1.29	0.33	1.24	1.1-1.5	<0.001
Trig (mmol/L)	1.00	0.81	0.83	0.6-1.1	1.20	0.57	1.07	0.8-1.4	<0.001	1.52	0.88	1.31	1.0-1.8	<0.001
Liver & kidney function														
AST (U/L)	18.49	5.10	17.50	14.8-21.1	18.70	5.64	17.20	15.4-21.2	1.000	20.38	9.51	17.15	14.3-23.5	1.000
ALT (IU/L)	15.82	8.56	13.80	10.9-17.8	19.19	12.40	16.00	12.2-22.0	0.007	23.60	17.68	17.50	13.0-27.5	<0.001
Bilirubin (umol/L)	6.91	3.33	5.90	4.8-8.6	6.00	2.86	5.40	4.1-7.4	0.019	6.18	2.85	5.50	4.2-7.5	0.025
Creatinine (umol/L)	57.90	8.75	57.10	51.8-63.9	59.70	9.37	58.85	52.5-65.5	0.296	56.24	12.27	55.60	48.9-61.8	0.015
Urea (mmol/L)	3.76	1.00	3.65	3.1-4.4	3.89	1.08	3.75	3.1-4.5	0.778	3.70	1.10	3.61	3.1-4.2	0.778

Abbreviations: NGT; normal glucose tolerance, T2D; type 2 diabetes, BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, SBP; systolic blood pressure, DBP; diastolic blood pressure, HOMA-IR; homeostatic model assessment of insulin resistance, HOMA-B; homeostatic model assessment of β-cell function, AST; aspartate transaminase, ALT; alanine transaminase.
P: p values for Wilcoxon rank-sum test performed comparing parameters in prediabetes and T2D with NGT. Adjusted for multiple testing.

Table 4.4. Characteristics of individuals from the EM1 sample population compared to those from the UK Biobank of Caucasian (CAU), South Asian (SA), and Afro-Caribbean (AC) descent. The EM1 sample population included 2,704 males and 4,579 females who attended Imperial College London in 2015 for general health screening. The UKBB cohort included 423,803 CAU, 7,841 SA, and 5,455 AC individuals. Fasting blood glucose data was available on 18,548 individuals in the UKBB and 3,776 individuals from the EM1.

	All														
	EM1			**CAU**				**SA**				**AC**			
N	**7,283**			**423,803**				**7,841**				**5,455**			
	Mean SD		**Median IQR**	**Mean SD**		**Median IQR**	**P**	**Mean SD**		**Median IQR**	**P**	**Mean SD**		**Median IQR**	**P**
Age (years)	35.84	13.94	33.00 (26.0-42.0)	56.62	8.05	58.00 (50.0-63.0)	<0.001	53.03	8.46	52.00 (46.0-60.0)	<0.001	51.50	7.99	50.00 (45.0-57.0)	<0.001
BMI (Kg/m^2)	28.45	6.33	27.64 (24.1-32.2)	27.27	4.68	26.61 (24.1-29.7)	<0.001	26.96	4.35	26.41 (24.0-29.2)	<0.001	29.26	5.37	28.48 (25.6-31.9)	<0.001
HbA1c (mmol/mol)	36.72	12.35	34.43 (31.2-36.7)	35.44	5.70	35.00 (32.6-37.3)	<0.001	39.29	9.70	37.20 (34.5-40.1)	<0.001	37.92	9.67	37.00 (33.8-39.7)	<0.001
FBG (mmol/l)	5.94	2.24	5.30 (5.1-5.5)	5.07	1.08	4.92 (4.6-5.3)	<0.001	5.65	2.04	4.98 (4.7-5.5)	<0.001	4.98	1.34	4.65 (4.4-5.1)	<0.001
	Males														

	EM1			CAU				SA				AC			
N	**2,704**			**193,102**				**4,259**				**2,374**			
	Mean SD		**Median IQR**	**Mean SD**		**Median IQR**	**P**	**Mean SD**		**Median IQR**	**P**	**Mean SD**		**Median IQR**	**P**
Age (years)	37.07	14.56	34.00 (27.0-44.0)	56.86	8.14	58.00 (50.0-64.0)	<0.001	53.32	8.77	53.00 (46.0-61.0)	<0.001	51.47	8.20	50.00 (45.0-57.0)	<0.001
BMI (Kg/m^2)	28.27	5.70	27.50 (24.5-31.3)	27.74	4.16	27.23 (24.9-29.9)	<0.001	26.83	3.86	26.42 (24.3-29.0)	<0.001	28.24	4.27	27.82 (25.5-30.5)	0.094
HbA1c (mmol/mol)	38.20	13.69	34.43 (31.2-38.8)	35.72	6.42	35.00 (32.6-37.4)	<0.001	39.99	10.37	37.40 (34.7-40.6)	<0.001	38.50	11.06	37.30 (34.1-40.1)	<0.001
FBG (mmol/l)	6.31	2.64	5.39 (5.2-5.6)	5.12	1.19	4.94 (4.6-5.3)	<0.001	5.76	2.19	4.90 (4.7-5.8)	<0.001	4.97	1.16	4.65 (4.4-5.2)	<0.001
	Females														
	EM1			**CAU**				**SA**				**AC**			
N	**4,579**			**230,701**				**3,582**				**3,081**			
	Mean SD		**Median IQR**	**Mean SD**		**Median IQR**	**P**	**Mean SD**		**Median IQR**	**P**	**Mean SD**		**Median IQR**	**P**

Age (years)	35.11	13.51	32.00 (25.0-41.0)	56.42	7.96	58.00 (50.0-63.0)	<0.001	52.70	8.08	52.00 (46.0-59.0)	<0.001	51.52	7.82	50.00 (45.0-57.0)	<0.001
BMI (Kg/m^2)	28.56	6.67	27.78 (23.8-32.7)	26.88	5.05	25.96 (23.4-29.4)	<0.001	27.11	4.85	26.40 (23.7-29.6)	<0.001	30.04	5.96	29.27 (25.7-33.2)	<0.001
HbA1c (mmol/mol)	35.85	11.40	33.33 (30.1-36.6)	35.20	5.00	35.00 (32.6-37.4)	<0.001	38.48	8.79	37.00 (34.3-39.7)	<0.001	37.47	8.42	36.60 (33.6-39.4)	<0.001
FBG (mmol/l)	5.75	1.99	5.26 (5.0-5.5)	5.03	0.97	4.90 (4.6-5.3)	<0.001	5.56	1.91	5.01 (4.7-5.5)	<0.001	4.98	1.45	4.67 (4.4-5.1)	<0.001

Abbreviations: BMI; body mass index, HbA1c; glycated haemoglobin, FBG; fasting blood (plasma) glucose.

P: p values for Wilcoxon rank-sum test performed comparing parameters for each ethnic group to EM1. Adjusted for multiple testing

Table 4.5. Multinomial regression analysis to test the contribution of BMI towards increased risk of T2D in Arab-Emirati (EM1), South Asian (SA), and Afro-Caribbean (AC) males and females, in comparison to Caucasians (CAU, reference category). The EM1 sample population included 2,704 males and 4,579 females who attended Imperial College London in 2015 for general health screening. Data on 423,803 CAU, 7,841 SA, and 5,455 AC individuals were retrieved from the baseline UK Biobank cohort.

	Males						
		Model 1		**Model 2**		**Model 3**	
	N	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**
CAU	193,102	**Reference category**					
EM1	2,704	3.18 (2.88-3.51)	<0.001	11.94 (10.62-13.43)	<0.001	12.49 (11.03-14.14)	<0.001
AC	2,374	2.77 (2.48-3.09)	<0.001	4.26 (3.80-4.77)	<0.001	4.45 (3.94-5.02)	<0.001
SA	4,259	4.22 (3.92-4.54)	<0.001	5.72 (5.30-6.18)	<0.001	8.19 (7.55-8.90)	<0.001
	Females						
		Model 1		**Model 2**		**Model 3**	
	N	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**
CAU	230,720	**Reference category**					
EM1	4,579	4.39 (4.01-4.81)	<0.001	20.98 (18.74-23.49)	<0.001	14.75 (13.08-16.64)	<0.001

AC	3,081	4.52 (4.05-5.04)	<0.001	7.13 (6.37-7.98)	<0.001	4.64 (4.10-5.25)	<0.001
SA	3,582	4.95 (4.50-5.46)	<0.001	7.04 (6.36-7.80)	<0.001	8.22 (7.38-9.20)	<0.001

Model 1: Unadjusted odds ratios (OR); Model 2: Adjusted OR for age; Model 3: Adjusted OR for age and BMI.

Table 4.6. Odds ratios for T2D among Arab-Emirati (EM1), South Asian (SA), and Afro-Caribbean (AC) compared to Caucasian (CAU) (reference category) males and females stratified by obesity classes. Obesity classes were defined according to WHO classification.

	CAU (Reference category)	**EM1**			**AC**			**SA**		
	N	**N**	**OR (95%CI)**	**P**	**N**	**OR (95%CI)**	**P**	**N**	**OR (95%CI)**	**P**
Normal BMI	142,132	2,080			1,088			2,629		
Model 1			5.02 (4.17-6.04)	<0.001		5.80 (4.57-7.36)	<0.001		5.80 (9.26-11.92)	<0.001
Model 2			37.32 (29.3-47.48)	<0.001		9.94 (7.77-12.72)	<0.001		14.69 (12.86-16.77)	<0.001
Overweight	180,858	2,386			2,248			3,481		
Model 1			4.79 (4.28-5.37)	<0.001		3.79 (3.34-4.31)	<0.001		6.21 (5.69-6.79)	<0.001
Model 2			21.91 (19.01-25.26)	<0.001		6.49 (5.68-7.42)	<0.001		9.29 (8.46-10.20)	<0.001
Class I/II obese	90,909	2,282			1,786			1,444		
Model 1			2.28 (2.06-2.53)	<0.001		1.92 (1.70-2.17)	<0.001		3.36 (2.99-3.78)	<0.001

Model 2			7.77 (6.88-8.77)	<0.001		2.85 (2.51-3.22)	<0.001		4.73 (4.19-5.34)	<0.001
Class III obese	7,170	352			237			89		
Model 1			0.95 (0.74-1.22)	<0.001		1.37 (1.04-1.80)	<0.001		4.08 (2.67-6.24)	<0.001
Model 2			2.63 (1.97-3.50)	<0.001		1.76 (1.32-2.34)	<0.001		5.58 (3.59-8.66)	<0.001

Model 1: Unadjusted odds ratios (OR); Model 2: Adjusted OR for age; Normal BMI = 18.0-24.9 Kg/m^2; Overweight = 25.0-29.9 Kg/m^2; Class I/II obese = 30.0-39.9 Kg/m^2; Class III obese ≥ 40.0 Kg/m^2.

Table 4.7. Type 2 diabetes (T2D) risk attributed to BMI and WHR among Arab-Emiratis (7,283 from EM1 and 1,077 from EM2) and 423,803 Caucasian (CAU), 7,841 South Asian (SA), and 5,455 Afro-Caribbean (AC) individuals from the UK Biobank cohort. The EM2 sample population was collected in a case-control study. The EM1 sample population included 2,704 Arab-Emirati males and 4,579 females who attended Imperial College London in 2015 for general health screening. Data on CAU, SA, and AC individuals were retrieved from the UKBB baseline cohort.

	Males														
Exposure	**Pe**					**OR (95%CI)**					**PAR (%)**				
	EM		**UKBB**			**EM**		**UKBB**			**EM**		**UKBB**		
	EM2	EM1	CAU	SA	AC	EM2	EM1	CAU	SA	AC	EM2	EM1	CAU	SA	AC
BMI ≥ 30 Kg/m^2															
Model 1	0.58	0.33	0.24	0.17	0.28	2.94 (2.6-3.4)	1.63 (1.5-1.7)	4.19 (4.1-4.2)	2.09 (2.0-2.2)	1.76 (1.6-1.9)	38.00	17.21	43.36	15.63	17.55
Model 2	0.58	0.33	0.24	0.17	0.28	2.96 (2.6-3.4)	2.52 (2.3-2.8)	4.31 (4.3-4.4)	2.32 (2.2-2.5)	1.83 (1.7-2.0)	38.14	33.40	44.27	18.33	17.55
WHR > 0.9															

Model 1	0.91		0.70	0.77	0.53	3.73 (3.1-4.5)		4.49 (4.4-4.6)	3.71 (3.4-4.0)	4.21 (3.8-4.6)	66.67		70.96	67.60	62.98
Model 2	0.91		0.70	0.77	0.53	3.71 (3.1-4.5)		3.92 (3.8-4.0)	2.92 (2.7-3.2)	3.15 (2.9-3.5)	66.47		67.15	59.65	53.26
Model 3	0.91		0.70	0.77	0.53	2.74 (2.3-3.3)		1.18 (1.1-1.2)	2.10 (1.9-2.3)	2.53 (2.3-2.8)	57.79		11.19	45.86	44.78
	Females														
Exposure	**Pe**					**OR (95%CI)**					**PAR (%)**				
	EM		**UKBB**			**EM**		**UKBB**			**EM**		**UKBB**		
	EM2	EM1	CAU	SA	AC	EM2	EM1	CAU	SA	AC	EM2	EM1	CAU	SA	AC
BMI ≥ 30 Kg/m^2															
Model 1	0.71	0.38	0.22	0.23	0.44	2.89 (2.6-3.3)	3.38 (3.2-3.6)	6.56 (6.5-6.7)	3.13 (2.9-3.4)	2.96 (2.7-3.2)	46.43	47.49	55.02	32.88	46.31
Model 2	0.71	0.38	0.22	0.23	0.44	2.83 (2.5-3.2)	2.26 (2.1-2.5)	6.49 (6.4-6.6)	3.29 (3.1-3.5)	2.89 (2.7-3.1)	45.91	32.38	54.71	34.50	45.40
WHR > 0.9															

Model 1	0.92		0.55	0.71	0.69	7.99 (6.6-9.7)		7.00 (6.8-7.2)	5.66 (5.0-6.4)	9.54 (8.1-11.2)	80.49		76.74	76.79	85.49
Model 2	0.92		0.55	0.71	0.69	7.88 (6.5-9.5)		6.19 (6.0-6.3)	4.87 (4.3-5.5)	7.41 (6.3-8.7)	80.32		74.06	73.32	81.56
Model 3	0.92		0.55	0.71	0.69	6.56 (5.4-8.0)		3.30 (3.2-3.4)	3.73 (3.3-4.2)	5.36 (4.5-6.3)	77.98		55.85	65.97	75.05

Abbreviations: Pe; proportion exposed, OR: odds ratios for type 2 diabetes; PAR; population attributable risk; BMI, body mass index, WHR; waist-to-hip ratio. Population attributable risk (PAR) for EM2 = Pe (proportion of cases exposed) * (OR-1) / OR. Population attributable risk (PAR) for EM1 and UKBB = Pe (proportion of population exposed) * (OR-1) / 1+(Pe (OR-1)). Model 1: Unadjusted ORs, Model 2: Adjusted for age, Model 3: Adjusted for age and BMI.

Table 4.8. Linear regression analysis showing the association of adiposity measures with insulin resistance and β-cell dysfunction as measured by homeostatic assessment model (HOMA-IR and HOMA-B) in Arab-Emirati males and females from the EM2 sample population. The EM2 sample population was collected in a case-control study design and included 693 Arab-Emirati males and 755 females.

	HOMA-IR											
	Males						**Females**					
BMI (Kg/m^2)	**B**	**95% CI**		**Standardised coefficient**	**% Change**	**P**	**B**	**95% CI**		**Standardised coefficient**	**% Change**	**P**
BMI + Age	0.026	0.025	0.027	0.407	50.26	<0.001	0.018	0.017	0.019	0.377	53.65	<0.001
BMI + Age + T2D	0.018	0.016	0.019	0.277	31.88	<0.001	0.012	0.011	0.014	0.258	47.64	<0.001

Waist (cm)												
Waist + Age	0.011	0.010	0.012	0.416	55.74	<0.001	0.009	0.009	0.010	0.419	55.95	<0.001
Waist + Age + BMI	0.007	0.005	0.008	0.249	47.21	0.002	0.008	0.007	0.009	0.370	53.25	<0.001
Waist + Age + BMI + T2D	0.006	0.005	0.008	0.230	46.32	0.006	0.008	0.007	0.009	0.347	52.06	<0.001
WHR												
WHR + Age	0.017	0.015	0.019	0.268	48.10	<0.001	0.013	0.012	0.014	0.315	50.42	<0.001
WHR + Age + BMI	0.010	0.009	0.012	0.162	43.25	<0.001	0.010	0.009	0.011	0.245	47.02	<0.001
WHR + Age + BMI + T2D	0.008	0.007	0.010	0.132	41.96	<0.001	0.008	0.007	0.009	0.188	44.38	<0.001
VAT (L)												
VAT + Age	0.073	0.069	0.077	0.435	56.84	<0.001	0.086	0.081	0.091	0.444	57.32	<0.001
VAT + Age + BMI	0.063	0.052	0.073	0.376	53.59	<0.001	0.074	0.066	0.082	0.383	53.98	<0.001
VAT + Age + BMI + T2D	0.045	0.034	0.055	0.268	48.08	0.006	0.064	0.055	0.073	0.329	51.14	<0.001

	HOMA-B											
	Males						Females					
	B	95% CI		Standardised coefficient	% Change	P	B	95% CI		Standardised coefficient	% Change	P
BMI (Kg/m^2)												
BMI + Age	0.020	0.018	0.021	0.317	37.29	<0.001	0.010	0.009	0.011	0.183	44.19	<0.001
BMI + Age + T2D	0.021	0.019	0.022	0.332	39.31	<0.001	0.012	0.011	0.013	0.224	46.03	<0.001
Waist (cm)												
Waist + Age	0.008	0.007	0.008	0.295	34.35	<0.001	0.003	0.003	0.004	0.132	41.99	0.002
Waist + Age + BMI	0.002	0.000	0.003	0.074	7.66	0.389	-0.002	-0.003	-0.001	-0.080	33.95	0.289
Waist + Age + BMI + T2D	0.004	0.002	0.005	0.140	15.01	0.098	0.005	0.003	0.006	0.188	44.41	0.013
WHR												
WHR + Age	0.005	0.004	0.007	0.083	37.29	0.038	-0.000	-0.001	0.001	-0.002	36.71	0.979

							1					
WHR + Age + BMI	-0.001	-0.002	0.001	-0.011	37.29	0.779	-0.002	-0.003	-0.001	-0.045	35.17	0.290
WHR + Age + BMI + T2D	0.004	0.002	0.005	0.062	37.29	0.118	0.005	0.004	0.006	0.105	40.85	0.011
VAT (L)												
VAT + Age	0.051	0.046	0.055	0.313	37.29	<0.001	0.028	0.022	0.034	0.129	41.86	0.002
VAT + Age + BMI	0.018	0.007	0.029	0.111	37.29	0.272	-0.011	-0.021	-0.001	-0.051	34.97	0.459
VAT + Age + BMI + T2D	0.023	0.012	0.033	0.141	37.29	0.146	0.034	0.025	0.044	0.158	43.07	0.015

Abbreviations: BMI; body mass index, WHRX100; (waist to hip ratio) *100, VAT; estimated visceral adipose tissue volume. VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm). HOMA-IR and HOMA-B were $\log_{10}$ transformed for linear regression analyses. Standardised coefficient = β * (SD of predictor/SD of outcome). Percentage change = [exp(standardised coefficient) - 1] *100. Change in outcome = [exp(standardised coefficient) - 1] * SD of outcome

Table 4.9. Characteristics of 557 females from the EM2 sample population with normal glucose tolerance (NGT) or type 2 diabetes (T2D), compared to females from the UK Biobank cohort; 230,701 Caucasians (CAU), 3,582 South Asians (SA), and 3,081 Afro-Caribbeans (AC). The EM2 sample population was collected in a case-control study design. UKBB baseline cohort data were used.

	NGT										
	EM2		**CAU**			**SA**			**AC**		
N	**288**		**223,513**			**3,090**			**2,690**		
	Mean SD	**Median IQR**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**
Age (years)	37.46 (6.13)	36.89 (33.4-41.0)	56.29 (7.97)	57.00 (50.0-63.0)	<0.001	52.00 (8.02)	51.00 (45.0-58.0)	<0.001	50.68 (7.46)	49.00 (45.0-55.0)	<0.001
BMI (Kg/m^2)	29.92 (6.52)	29.43 (24.6-34.2)	26.69 (4.87)	25.84 (23.3-29.2)	<0.001	26.65 (4.55)	26.02 (23.5-29.1)	<0.001	29.52 (5.73)	28.78 (25.5-32.5)	0.493
Waist (cm)	87.95 (13.39)	87.00 (79.0-98.0)	83.67 (11.84)	82.00 (75.0-91.0)	<0.001	84.48 (11.02)	84.00 (77.0-91.0)	<0.001	89.16 (12.78)	88.00 (80.0-97.0)	0.241
WHR	0.83 (0.07)	0.83 (0.77-0.87)	0.81 (0.07)	0.81 (0.76-0.86)	0.021	0.84 (0.07)	0.83 (0.79-0.88)	0.073	0.83 (0.07)	0.83 (0.79-0.88)	0.163
VAT (L)	2.1 (1.56)	1.91 (1.0-3.2)	2.12 (1.48)	1.96 (1.1-3.0)	1.000	2.02 (1.26)	1.95 (1.1-2.8)	1.000	2.75 (1.65)	2.58 (1.6-3.7)	<0.001

TC (mmol/l)	4.63 (0.93)	4.5 (4.1-5.1)	5.93 (1.10)	5.87 (5.2-6.6)	<0.001	5.57 (1.01)	5.49 (4.9-6.2)	<0.001	5.34 (1.04)	5.28 (4.6-6.0)	<0.001
LDL (mmol/l)	3.00 (0.81)	2.91 (2.5-3.4)	3.66 (0.86)	3.60 (3.1-4.2)	<0.001	3.48 (0.77)	3.41 (2.9-4.0)	<0.001	3.28 (0.79)	3.21 (2.7-3.8)	<0.001
HDL (mmol/l)	1.52 (0.36)	1.48 (1.3-1.7)	1.61 (0.37)	1.57 (1.3-1.8)	<0.001	1.43 (0.33)	1.39 (1.2-1.6)	<0.001	1.56 (0.36)	1.52 (1.3-1.8)	0.068
Triglycerides (mmol/l)	1.01 (0.81)	0.84 (0.6-1.2)	1.52 (0.82)	1.31 (1.0-1.8)	<0.001	1.66 (0.95)	1.42 (1.0-2.0)	<0.001	1.04 (0.54)	0.91 (0.7-1.2)	0.003
	T2D										
	EM2		**CAU**			**SA**			**AC**		
N	**269**		**7,188**			**492**			**391**		
	Mean SD	**Median IQR**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**
Age (years)	38.39 (5.56)	38.86 (35.3-42.5)	60.38 (6.65)	62.00 (56.0-66.0)	<0.001	57.11 (6.98)	58.00 (52.0-63.0)	<0.001	57.29 (7.83)	58.00 (51.0-64.0)	<0.001
BMI (Kg/m^2)	34.34 (6.91)	33.24 (29.5-38.0)	32.81 (6.67)	32.05 (28.1-36.8)	0.030	30.01 (5.64)	28.86 (26.0-32.8)	<0.001	33.62 (6.33)	33.04 (28.7-37.3)	0.306
Waist (cm)	100.64 (13.94)	99.00 (90.0- 110.0)	100.05 (14.77)	99.00 (90.0-110.0)	1.000	94.41 (11.31)	93.00 (86.0-102.0)	<0.001	101.10 (12.80)	99.00 (92.0-109.0)	1.000

WHR	0.89 (0.07)	0.89 (0.84-0.92)	0.89 0.07	0.89 (0.84-0.93)	1.000	0.89 (0.07)	0.89 (0.85-0.94)	1.000	0.90 (0.07)	0.89 (0.86-0.94)	0.840
VAT (L)	3.58 (1.58)	3.30 (2.5-4.5)	4.20 (1.73)	4.13 (3.1-5.3)	<0.001	3.20 (1.15)	3.07 (2.5-3.9)	0.012	4.35 (1.58)	4.15 (3.3-5.2)	<0.001
TC (mmol/l)	4.61 (1.00)	4.56 (4.0-5.2)	4.73 (1.04)	4.60 (4.0-5.3)	0.748	4.66 (1.00)	4.53 (3.9-5.2)	1.000	4.62 (1.01)	4.56 (4.0-5.1)	1.000
LDL (mmol/l)	3.06 (0.93)	3.00 (2.4-3.6)	2.81 (0.78)	2.70 (2.3-3.2)	<0.001	2.79 (0.78)	2.68 (2.2-3.2)	<0.001	2.76 (0.77)	2.66 (2.2-3.1)	<0.001
HDL (mmol/l)	1.29 (0.33)	1.24 (1.1-1.5)	1.31 (0.34)	1.25 (1.1-1.5)	0.786	1.27 (0.29)	1.23 (1.1-1.4)	0.786	1.39 (0.36)	1.34 (1.2-1.6)	<0.001
Triglycerides (mmol/l)	1.51 (0.89)	1.30 (1.0-1.8)	2.11 (1.13)	1.87 (1.4-2.6)	<0.001	1.99 (1.07)	1.76 (1.4-2.3)	<0.001	1.28 (0.66)	1.09 (0.8-1.6)	<0.001

Abbreviations: NGT; normal glucose tolerance, T2D; type 2 diabetes, BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, VAT; estimated visceral adipose tissue volume. VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm). P: p values for Wilcoxon rank-sum test performed comparing parameters for each ethnic group to EM2. Adjusted for multiple testing.

Table 4.10. Characteristics of 520 males from the EM2 sample population with normal glucose tolerance (NGT) or type 2 diabetes (T2D), compared to males from the UK Biobank cohort; 193,102 Caucasians (CAU), 4,259 South Asians (SA), and 2,374 Afro-Caribbeans (AC). The EM2 sample population was collected in a case-control study design. UKBB baseline cohort data were used.

	NGT									
	EM2		**CAU**			**SA**			**AC**		
N	**171**		**180,209**			**3,271**			**1,981**		
	Mean SD	**Median IQR**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**
Age (years)	36.76 (5.35)	36.87 (33.0-40.2)	56.58 (8.17)	58 (50.0-63.0)	<0.001	51.96 (8.52)	51 (44.5-59.0)	<0.001	50.48 (7.77)	49 (44.0-56.0)	<0.001
BMI (Kg/m^2)	28.3 (4.97)	27.89 (25.3-31.2)	27.49 (3.95)	27.05 (24.8-29.6)	0.025	26.47 (3.57)	26.13 (24.1-28.6)	<0.001	27.96 (4.04)	27.61 (25.4-30.20)	0.382
Waist (cm)	96.46 (11.62)	97.00 (89.0-103.0)	96.03 (10.66)	95 (89.0-102.0)	0.482	93.79 (9.71)	93 (87.0-99.0)	0.005	93.05 (10.58)	92 (86.0-99.0)	<0.001
WHR	0.94 (0.06)	0.94 (0.90-0.97)	0.93 (0.06)	0.93 (0.89-0.97)	0.251	0.94 (0.06)	0.94 (0.90-0.98)	0.814	0.9 (0.06)	0.9 (0.86-0.94)	<0.001
VAT (L)	4.02 (1.86)	4.00 (2.9-5.1)	4.53 (1.84)	4.37 (3.3-5.6)	<0.001	3.82 (1.56)	3.72 (2.8-4.7)	0.242	4.11 (1.74)	3.91 (2.9-5.1)	0.776

TC (mmol/l)	4.76 (0.82)	4.68 (4.2-5.2)	5.58 (1.09)	5.55 (4.8-6.3)	<0.001	5.42 (1.07)	5.42 (4.7-6.1)	<0.001	5.26 (1.06)	5.18 (4.5-5.9)	<0.001
LDL (mmol/l)	3.28 (0.78)	3.25 (2.7-3.7)	3.55 (0.83)	3.53 (3.0-4.1)	<0.001	(3.48) (0.81)	3.5 (2.9-4.0)	<0.001	3.34 (0.82)	3.27 (2.8-3.9)	0.299
HDL (mmol/l)	1.28 (0.31)	1.26 (1.1-1.5)	1.3 (0.31)	1.26 (1.1-1.5)	0.731	1.18 (0.28)	1.14 (1.0-1.3)	<0.001	1.33 (0.33)	1.28 (1.1-1.5)	0.448
Triglycerides (mmol/l)	1.19 (0.88)	0.94 (0.7-1.3)	1.95 (1.12)	1.67 (1.2-2.4)	<0.001	2.12 (1.25)	1.79 (1.3-2.6)	<0.001	1.36 (0.86)	1.13 (0.8-1.6)	<0.001
	T2D										
	EM2		**CAU**			**SA**			**AC**		
N	**349**		**12,893**			**988**			**393**		
	Mean SD	**Median IQR**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**	**Mean SD**	**Median IQR**	**P**
Age (years)	37.16 (5.42)	37.92 (33.9-41.3)	60.64 (6.53)	62 (57.0-66.0)	<0.001	57.79 (8.05)	59 (52.0-65.0)	<0.001	56.5 (8.49)	56 (50.0-65.0)	<0.001
BMI (Kg/m^2)	31.63 (5.53)	31.02 (27.5-34.7)	31.29 (5.24)	30.59 (30.6-34.2)	0.281	28.04 (4.47)	27.15 (25.0-30.1)	<0.001	29.67 (5.01)	28.79 (26.3-32.2)	<0.001
Waist (cm)	105.59 (13.6)	104 (97.0-113.0)	106.59 (13.23)	105 (97.0-114.4)	0.189	99.61 (11.49)	98 (92.0-106.0)	<0.001	100.69 (12.93)	99 (92.0-108.0)	<0.001

WHR	0.98 (0.06)	0.98 (0.94-1.01)	0.99 (0.07)	0.98 (0.94-1.03)	0.096	0.98 (0.06)	0.98 (0.94-1.02)	0.407	0.96 (0.07)	0.95 (0.91-1.00)	<0.001
VAT (L)	5.49 (2.13)	5.24 (4.0-6.7)	6.37 (2.21)	6.15 (4.8-7.7)	<0.001	4.78 (1.77)	4.46 (3.6-5.7)	<0.001	5.36 (2.23)	5.01 (3.9-6.4)	0.187
TC (mmol/l)	4.39 (1.19)	4.22 (3.5-5.0)	4.32 (0.97)	4.21 (3.7-4.8)	1.000	4.29 (1.05)	4.14 (3.5-4.9)	1.000	4.35 (1.16)	4.2 (3.6-4.9)	1.000
LDL (mmol/l)	2.9 (1.03)	2.72 (2.1-3.5)	2.6 (0.71)	2.5 (2.1-3.0)	<0.001	2.61 (0.77)	2.48 (2.1-3.0)	<0.001	2.65 (0.86)	2.52 (2.1-3.0)	0.004
HDL (mmol/l)	1.06 (0.25)	1.06 (0.9-1.2)	1.11 (0.28)	1.06 (0.9-1.2)	0.149	1.09 (0.26)	1.05 (0.9-1.2)	0.584	1.18 (0.31)	1.16 (1.0-1.3)	<0.001
Triglycerides (mmol/l)	1.98 (2.0)	1.54 (1.1-2.2)	2.25 (1.3)	1.94 (1.4-2.8)	<0.001	2.01 (1.16)	1.73 (1.2-2.5)	0.002	1.51 (1.09)	1.2 (0.8-1.9)	<0.001

Abbreviations: NGT; normal glucose tolerance, T2D; type 2 diabetes, BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, VAT; estimated visceral adipose tissue volume. VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm). P: p values for Wilcoxon rank-sum test performed comparing parameters for each ethnic group to EM2. Adjusted for multiple testing.

Table 4.11. Logistic regression analysis showing the association of adiposity measures with type 2 diabetes (T2D) in 1,077 Arab-Emirati males and females from the EM2 sample population in comparison to 423,803 Caucasian (CAU), 7,841 South Asian (SA) and 5,455 Afro-Caribbean (AC) data from baseline UK Biobank cohort. The EM2 sample population was collected in a case-control study design. Odds ratios are presented for one unit change in each predictor/explanatory variable but for 0.01 change in waist-to-hip ratio.

	Males															
	EM2				**CAU**				**SA**				**AC**			
N	**520**				**193,102**				**4,259**				**2,374**			
	OR	**95% CI**		**P**	**OR**	**95% CI**		**P**	**OR**	**95% CI**		**P**	**OR**	**95% CI**		**P**
BMI + Age	1.14	1.12	1.16	<0.001	1.20	1.20	1.20	<0.001	1.13	1.12	1.13	<0.001	1.10	1.09	1.11	<0.001
Waist + Age	1.06	1.06	1.07	<0.001	1.08	1.07	1.08	<0.001	1.05	1.05	1.05	<0.001	1.05	1.05	1.06	<0.001
Waist + Age + BMI	1.05	1.04	1.06	0.0093	1.04	1.04	1.04	<0.001	1.05	1.04	1.05	<0.001	1.10	1.09	1.11	<0.001
WHR + Age	1.14	1.13	1.16	<0.001	1.13	1.13	1.13	<0.001	1.10	1.10	1.11	<0.001	1.11	1.11	1.12	<0.001
WHR + Age + BMI	1.12	1.10	1.13	<0.001	1.07	1.07	1.07	<0.001	1.08	1.08	1.09	<0.001	1.11	1.10	1.12	<0.001

VAT + Age	1.47	1.41	1.53	<0.001	1.52	1.52	1.53	<0.001	1.36	1.34	1.39	<0.001	1.34	1.31	1.37	<0.001
VAT + Age + BMI	1.50	1.37	1.66	0.004	1.28	1.27	1.29	<0.001	1.39	1.33	1.45	<0.001	1.87	1.77	1.98	<0.001
	Females															
	EM2				**CAU**				**SA**				**AC**			
N	**557**				**230,701**				**3,582**				**3,081**			
	OR	**95% CI**		**P**	**OR**	**95% CI**		**P**	**OR**	**95% CI**		**P**	**OR**	**95% CI**		**P**
BMI + Age	1.10	1.09	1.12	<0.001	1.18	1.18	1.19	<0.001	1.14	1.13	1.15	<0.001	1.11	1.10	1.12	<0.001
Waist + Age	1.07	1.07	1.08	<0.001	1.09	1.09	1.09	<0.001	1.07	1.07	1.08	<0.001	1.07	1.06	1.07	<0.001
Waist + Age + BMI	1.10	1.09	1.11	<0.001	1.09	1.08	1.09	<0.001	1.07	1.07	1.08	<0.001	1.09	1.09	1.10	<0.001
WHR + Age	1.15	1.14	1.16	<0.001	1.15	1.15	1.15	<0.001	1.11	1.10	1.12	<0.001	1.13	1.12	1.14	<0.001
WHR + Age + BMI	1.14	1.12	1.15	<0.001	1.11	1.11	1.11	<0.001	1.10	1.09	1.10	<0.001	1.11	1.10	1.12	<0.001
VAT + Age	1.88	1.79	1.97	<0.001	2.02	2.01	2.03	<0.001	1.92	1.87	1.98	<0.001	1.62	1.58	1.65	<0.001

VAT + Age + BMI	2.09	1.94	2.26	<0.001	1.80	1.78	1.81	<0.001	1.77	1.69	1.85	<0.001	1.65	1.58	1.72	<0.001

Abbreviations: BMI; body mass index, WHRX100; (waist to hip ratio)*100, VAT; estimated visceral adipose tissue volume. VAT = 0.03*age + 0.40*sex - 0.26*height + 0.09*weight - 4518*weight/height^2 + 0.22*(height·weight)^0.5 + 0.24*waist circumference - 0.20*hip circumference - 15.16*waist/hip + 37.74. Units for the variables: age (year), sex (female = 0, male = 1), height (cm), weight (kg), WC (cm), HC (cm).

Chapter Five: The contribution of physical inactivity and sedentary time towards T2D risk among Arab-Emiratis

Lack of physical activity that populations in the Gulf region are characterised by, is reasonably well-documented in the literature. It may be assumed that a large proportion of T2D risk within these populations is attributable to this lack of physical activity, but that is yet to be evidenced. In this chapter, I investigate the contribution of physical inactivity and sedentary time towards T2D risk and its related metabolic disturbances, as well as to anthropometric measures among Arab-Emiratis (study population of the UAE), in comparison to other population groups, including Europeans, South Asians, and Afro-Caribbeans. Supplementary figures and tables can be accessed through Zenodo using this link: https://zenodo.org/record/7825396#.ZDfrqC0Rro4.

5.1 OVERVIEW

I used the Global Physical Activity Questionnaire (GPAQ) to collect physical activity data from 1,433 Arab-Emirati men and women (EM2 sample population). Physical activity was also quantified in 106 individuals from EM2 using Actigraph GT9X. Cross-sectional baseline physical activity questionnaire data on 344,616 Caucasians (CAU), 6,029 South Asians (SA), and 4,201 Afro-Caribbeans (AC) were retrieved from the UK Biobank (UKBB). Univariate and multiple regression approaches were used to investigate the relationship of physical activity parameters with T2D and its related metabolic disturbances.

Strength of these associations were assessed between ethnic groups. Arab-Emiratis living in the UAE have approximately 2-fold lower levels of physical activity compared to Europeans living in the UK (CAU). Sedentary behaviour and low physical activity appear to contribute to the increased risk of T2D in Arab-Emiratis through increased insulin resistance and central adiposity, and unfavourable lipid profile. Despite that, the extent to which lack of physical activity can explain differences in T2D risk between ethnic populations appears to be limited.

5.2 AIMS

To investigate the contribution of physical inactivity towards risk of T2D and associated glycaemic disturbances in Arab-Emiratis, and towards differences in magnitude of risk across ethnic groups.

5.2.1 Objectives

1. To collect physical activity data using the Global Physical Activity Questionnaire (GPAQ) from a sample of Arab-Emirati men and women with or without T2D, collected in a case-control study (**EM2**).
2. To assess the validity of questionnaire-specific physical activity data collected using a validated accelerometer, Actigraph GT9X in a subset of the EM2 sample population.
3. To investigate the contribution of physical inactivity towards differences in risk of T2D among ethnic groups. This includes:
    a. Determination of total, work-, leisure-, and transport-specific physical activity levels as well as sedentary time in Arab-Emiratis.
    b. Assessment of the relationship between sedentary time, overall low physical activity (LPA) and domain-specific physical inactivity with T2D and its related metabolic disturbances.

c. Evaluation of the strength of association of physical inactivity with risk of T2D in Arab-Emiratis compared to individuals of Caucasian (CAU), South Asian (SA), and Afro-Caribbean (AC) descent from the UK Biobank (UKBB) cohort.
d. Investigation of the magnitude of T2D risk attributable to physical inactivity in each ethnic group.

5.3 INTRODUCTION

Physical inactivity has been recognised as one of the leading risk factors for global mortality contributing to 3.2 million deaths per year[104]. The relationship of physical inactivity with increased risk of developing T2D is well established, predominantly in European populations[124-127]. Although the majority of those studies focused on vigorous activities or physical activity of various intensities combined, adherence to moderate intensity activities such as brisk walking was also proven to substantially reduce the risk of T2D[318]. In fact, individuals who regularly engage in moderate intensity activities have approximately 30% lower risk of T2D, compared to those with sedentary lifestyles[318]. Moreover, randomised controlled trials have shown that physical activity alone, or in conjunction with dietary changes can reduce incidence of T2D[87, 88, 319].

Physical activity guidelines and recommendations for dyslipidaemia, prediabetes, T2D, and general health, were historically emphasised on moderate-vigorous physical activity levels[107, 110, 320]. More recently, studies are highlighting the role of reducing sedentary behaviour, which takes up the majority of the waking day, rather than just focussing on moderate-vigorous physical activity performed during only a small part of the week (~150 min/week)[321, 322]. In support of that, physiological studies have indicated that health consequences of modern lifestyles, which are shaped by excessive hours of

muscular inactivity, is distinct from not performing exercise[323], and therefore, guidelines targeting sedentary behaviours specifically are needed to improve human physiology and prevent disease.

Overall increased physical activity levels have also been suggested to improve insulin sensitivity and its related metabolic disturbances, independent of obesity[132, 320]. A 30-minute increase in moderate-intensity physical activity has been reported to reduce fasting insulin levels by approximately 7%[324]. In addition, increased physical activity levels can improve glucose metabolism, reduce triglyceride concentration, and increase high-density lipoprotein (HDL) levels[325], even if weight loss is not achieved[326, 327]. Physical activity interventions were also proven to be effective in improving adiposity measures including body mass index (BMI), and body fat composition; for a given BMI, people who participate in more physical activity have lower body fat percentage, compared to those who are less active[328].

Populations of the Middle East, particularly those in the Gulf region, have undergone rapid socioeconomic development that has resulted in striking demographic and epidemiological transitions, with obesity and T2D becoming some of the leading causes of morbidity and mortality[259]. The increased risk of T2D within these populations is often thought to be largely attributable to lack of physical activity, but that is yet to be investigated. According to WHO global estimates, the UAE is among the top 10 countries with some of the highest rates of physical inactivity; 39% in males and 49% in females[105]. Although limited, cross-sectional studies in populations of the Gulf region, including Saudi Arabia and Kuwait have reported higher rates of physical inactivity that ranged between 58%-61% in men and 72%-74% in women [262]. These figures raise the possibility of the contribution of physical inactivity to the increased risk of T2D in populations in the region in general, and form the basis of my hypothesis that

lack of physical activity and sedentary time may be attributable to a marked proportion of T2D risk in Arab-Emiratis.

In this study, I investigated the relationship of physical inactivity with the risk of T2D and its related metabolic disturbances in a sample population of Arab-Emiratis compared to other ethnic groups, including Europeans, South Asians, and Afro-Caribbeans living in the UK. Results presented here provide some insight into the mechanisms through which low physical activity can contribute to the increased T2D burden in the UAE.

5.4 MATERIALS AND METHODS

5.4.1 Study design and participant characterisation

I investigated 1,433 Arab-Emiratis from the EM2 sample population, 334,616 CAU, 6,029 SA, and 4,201 AC from the UKBB cohort. Study design, participant recruitment, screening and diagnosis were described in detail in Chapter 2 (UKBB; section 2.2, and EM2; section 2.4). Here I describe them briefly.

The EM2 sample population included Arab-Emirati men and women with T2D or free from diabetes, recruited in a case-control study design from ICLDC between 2017 and 2020 (**Chapter4**, **Figure 4.2**). Data from CAU, SA, and AC individuals were retrieved from the baseline UKBB cohort. **Figure 2.3** in Chapter 2 describes the UKBB data processing.

5.4.2 Physical activity

5.4.2.1 Physical activity in Arab-Emiratis

I collected physical activity data from the EM2 sample population using the validated Arabic version of the Global Physical Activity Questionnaire[329] (N = 1,448) and Actigraph GT9X accelerometers (N = 106).

5.4.2.1.1 Global Physical Activity Questionnaire

A detailed description of how GPAQ measures activity levels can be found in Chapter 2 section 2.4.4.1. The GPAQ scoring protocol was used to create the following indicators: total (TPA) and domain-specific physical activity (i.e., work, transport, leisure) (**Supplementary Table 5.1**). According to GPAQ guidelines[287], individuals were categorised into having: (1) Low- (LPA), (2) Moderate- (MPA), or High- (HPA) physical activity (**Supplementary Table 5.2**). Physical inactivity prevalence was defined as the prevalence of LPA. Domain-specific physical inactivity was calculated following GPAQ guidelines[287]. For each domain, 'active' individuals were defined as those who participated in moderate-vigorous activities for ≥ 10 minutes in a typical week. UKBB data-comparable parameters were computed, and these included the following:

MET_mod (MET minutes/week) *= (minutes spent in moderate intensity activities at work and leisure per week) * 4*

MET_vig (MET minutes/week) *= (minutes spent in vigorous intensity activities at work and leisure per week) * 8*

MET_walk (MET minutes/week) *= (minutes spent in transport activities per week) * 4*

5.4.2.1.2 Actigraph GT9X

One hundred and forty-three participants were recruited for this part of the study. Actigraph GT9X was used to quantify physical activity in a subset of individuals from the EM2 sample population. Actigraph GT9X is a small (3.5 x 3.5 x 1) and lightweight (14g) triaxial accelerometer-based physical activity monitor. A detailed description of monitor initialisation, activity data collection, and wear-time validation can be found in Chapter 2 section 2.4.4.2. Data scoring was performed in ActiLife using the Freedson Combination (1998) algorithm to calculate energy expenditure, Freedson Adult (1998) algorithm[288] to calculate METs, cut points and MVPA. The Freedson Combination algorithm uses the work-energy theorem to calculate caloric expenditure below 1951 counts and the Freedson '98 equation to calculate energy expenditure above 1952 counts.

Freedson Combination Algorithm:

If CPM > 1951 then,

*kcals/min = Scale * (0.00094 * CPM + (0.1346 * BM - 7.37418))*

Else,

*kcals/min = CPM * 0.0000191 * BM*

Where:

CPM = *Counts per minute*

BM = *Body mass in kilograms*

Freedson Adult Algorithm:

*1 MET = 1kcal/kg*h*

As described in Chapter 2 in more detail, counts are a result of summing post-filtered accelerometer values (30Hz) into epoch "chunks". The value of the counts varies based on the intensity of the raw acceleration. **Supplementary Table 5.3** describes the cut-off points used to define sedentary time, light, moderate, vigorous, and very vigorous activities[288]. End-point parameters computed were: (1) sedentary time (minutes/day), (2) time spent in light activities (minutes/day), (3) time spent in moderate activities (minutes/day), (4) time spent in vigorous or very vigorous activities (minutes/day), (5) total physical activity (TPA) as measured by time spent in moderate-vigorous activities (minutes/day), and calories burned (Kcal/day).

5.4.2.2 Physical activity in the UKBB

5.4.2.2.1 Short International Physical Activity Questionnaire

Physical activity in the UKBB baseline cohort was measured using the short International Physical Activity Questionnaire (IPAQ)[281, 282]. Data retrieval and processing is described in more detail in Chapter 2 section 2.2.4.1. Although the physical activity data collection questionnaire used for Arab-Emiratis was different from that used for the UKBB cohort, GPAQ and IPAQ have similar properties, pattern of non-response, data quality, and produce comparable results[330].

5.4.3 Diagnosis and definitions

5.4.3.1 T2D diagnosis

This was previously described in Chapter 4 (section 4.4.2). Briefly, for the EM2 sample population, T2D and prediabetes were diagnosed by ICLDC's endocrinologists following ADA guidelines. Insulin resistance and β-cell dysfunction were estimated by the Homeostatic Model Assessment for Insulin Resistance (HOMA-IR) and β-cell function (HOMA-B). Individuals in the UKBB were categorised into T2D and

controls according to two data fields: age diabetes diagnosed (Data field 2976) and self-reported diabetes (Data field 20002).

5.4.3.2 Adiposity

All adiposity measures calculations and definitions were described in Chapter 4 (section 4.4.2.3). These include BMI, waist circumference (WC), waist-to-hip ratio (WHR), and anthropometry-derived visceral adipose tissue volume (VAT).

5.4.4 Power Calculations

A Wald test was used to compute power. To investigate the association of physical inactivity with T2D with a sample size of 1,068 (individuals with prediabetes excluded from association analyses with T2D), and a proportion of exposed individuals (Prx = Pr(x=1)) of 0.6 and proportion of T2D cases (Pry=Pr(y=1|x=0)) of 0.5, I had 78% power to detect an odds ratio of 1.4. Prx estimate was based on previous studies conducted in populations in the Gulf region[262].

5.4.5 Statistical Analysis

R studio (Version 1.4.1106) was used for all statistical analysis presented. Figures were plotted in R studio or Graphpad Prism (Version 9.2.0). Results are expressed as mean (SD) and median (interquartile range) for continuous variables and as proportions and/or percentages for categorical variables with confidence intervals as necessary. Data normality was tested by Shapiro Normality test (**Supplementary Figures 1 and 2, Supplementary table 5.4**). The Wilcoxon rank-sum test was used to compare continuous parameters between sexes and diabetes status groups. The Chi-squared test was used to compare categorical parameters between sexes and diabetes status groups. Regression analyses were used to test the association of LPA (categorical; MPA/HPA = 0 and LPA = 1), MET_mod, MET_vig, and MET_walk with T2D, lipid and adiposity measures among Arab-Emirati, CAU, SA, and AC individuals.

Regression models included: unadjusted, age adjusted, age and sex adjusted, age sex and BMI adjusted. Unadjusted and adjusted odds ratios (ORs) and β coefficients are presented for: (1) 80 MET minutes/week change in TPA, which is equivalent to an increase of 10 minutes of moderate activity and 5 minutes of vigorous activity, or a combination of the two, (2) 80 MET minutes/week change in MET_mod, which is equivalent to an increase of 20 minutes of moderate activity, (3) 80 MET minutes/week change in MET_vig, which is equivalent to an increase of 10 minutes of vigorous activity, and (4) 80 MET minutes/week change in MET_walk, which is equivalent to an increase of 20 minutes of transport (walk/cycle) activity. Direction of effect and effect sizes were compared across ethnic groups. In addition, in the Arab-Emirati sample (EM2), the association of domain-specific (work, transport, leisure) physical activity (for every 1minute increase per day) and sedentary time (for every 1 hour increase per day) with T2D and related anthropometric and metabolic measures were tested using logistic and linear regressions. Differences between odds ratios or β coefficients across ethnic groups were considered significant ($p < 0.05$) if overlap between confidence intervals was less than 50%[307]. The association of Actigraph-computed average sedentary time with T2D and HOMAIR was also tested in Arab-Emiratis. To calculate the magnitude of T2D risk attributed to LPA and sedentary time (> 12 hours/day) [population attributable risk (PAR)] amongst Arab-Emiratis (case-control study design), I used the following formula:

PAR (%) = Pe * [(OR-1)/OR] * 100

Pe = proportion of cases exposed

To calculate PAR for UKBB, I used the following equation:

PAR = [Pe(RR-1)*100/Pe(RR-1)+1]

Pe = proportion of population exposed

RR = Relative risk (estimated as the OR)

Lastly, as a sensitivity analysis, I used multinomial logistic regression (MLR) to test for the relationship of ethnicity with T2D risk, and to determine the contribution of physical inactivity, as measured by LPA, to increased T2D risk among SA and AC individuals, compared to CAU individuals (reference category).

5.5 RESULTS

5.5.1 The contribution of physical inactivity towards risk of T2D, glycaemic disturbances, lipid, and adiposity measures in Arab-Emiratis

5.5.1.1 Physical activity levels as measured by GPAQ in the EM2 sample population

I collected physical activity data using the GPAQ from 1,485 Arab-Emirati men and women. Individuals with gestational, T1D, or MODY were excluded (n = 37). Individuals with incomplete data (not answering one or more of the 16 questions in the questionnaire) were also excluded (n = 15). 1,433 individuals had complete data; 455 with normal glucose tolerance (NGT), 365 with prediabetes, and 613 with T2D (**Figure 5.1B**). These included 684 males and 749 females. Mean (SD) age in this sample was 37.9 (5.9) years and mean (SD) BMI was 31.3 (6.3) Kg/m^2.

WHO recommends adults to participate in at least 150 minutes of moderate-intensity physical activity or 75 minutes of vigorous intensity activities per week[110]. In this sample, 59.3%, 59.7%, and 60.5% of individuals with NGT, prediabetes, and T2D met these recommendations, respectively.

GPAQ takes into consideration both the intensity and frequency of physical activity performed in a week. According to GPAQ guidelines, the overall activity of 65.9%, 67.4%, and 65.4% of individuals with

NGT, prediabetes, and T2D, respectively, was defined as low (LPA) (**Table 5.1, Figure 5.2**). Overall, males had lower rates of LPA compared to females, irrespective of diabetes status ($p < 0.001$). Whereas, in both males and females, LPA rates were not significantly different in individuals with prediabetes or T2D in comparison to their counterparts with NGT (**Table 5.1, Figure 5.2**). Similarly, the proportion of individuals who participated in moderate/vigorous activities at work, transport, or leisure time were not different across individuals with NGT, prediabetes, or T2D (**Table 5.1, Figure 5.3**).

Mean (SD) TPA (MET minutes/week) in this sample was 1,187.7 (2349.8); 769.9 (1,553.2) in females, and 1,616.0 (2929.9) in males. Males with T2D had significantly lower TPA compared to their counterparts with NGT ($p = 0.035$) (**Table 5.2**), but no significant differences were observed among females.

Figure 5.1 Flowchart demonstrating data collection and analysis performed using the Global Physical Activity Questionnaire in the EM2 sample population. (A) Study design for the case-control study conducted to collect the EM2 sample population (Figure 4.1, Chapter 4). (B) GPAQ data scoring and physical activity measures computed.

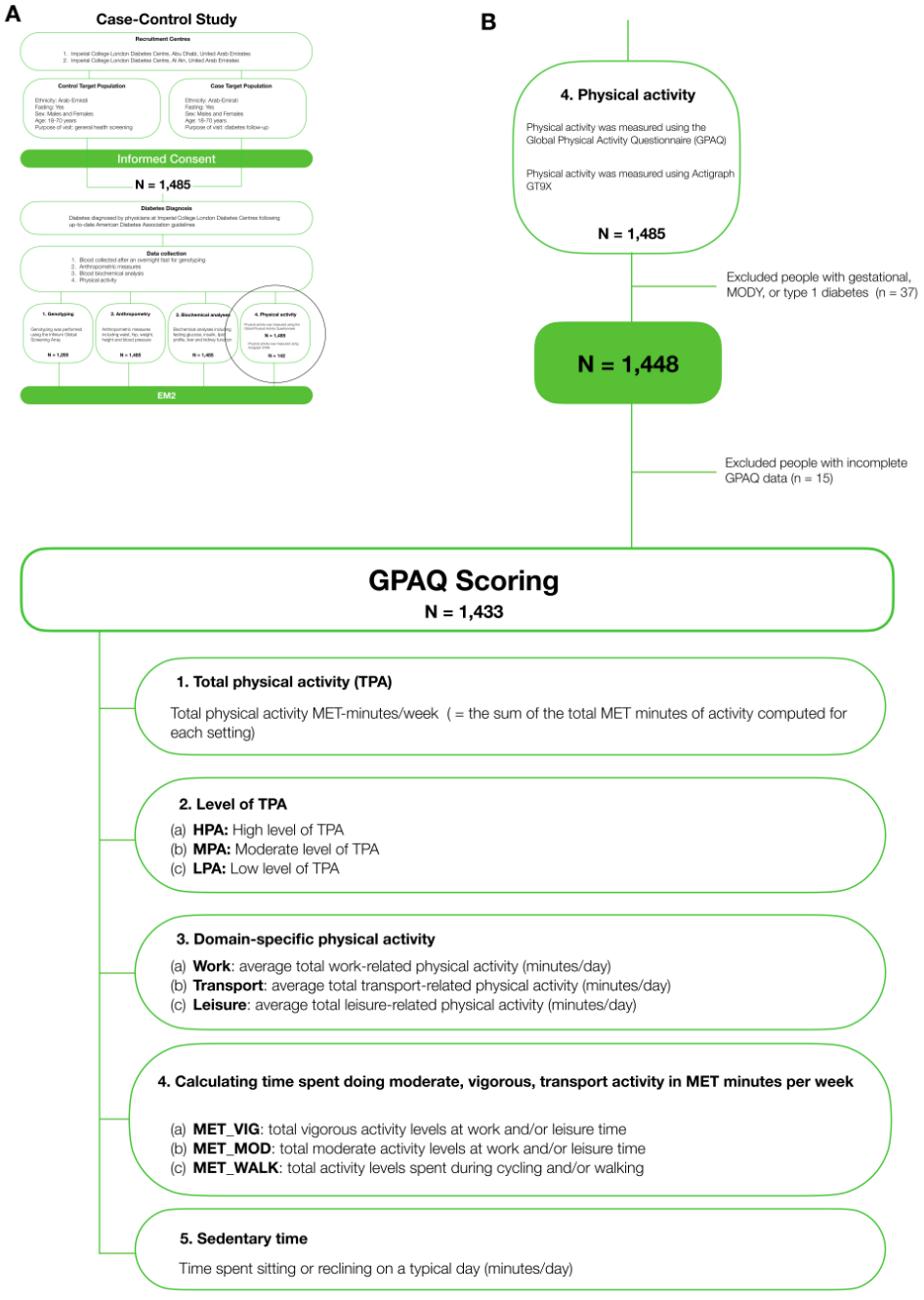
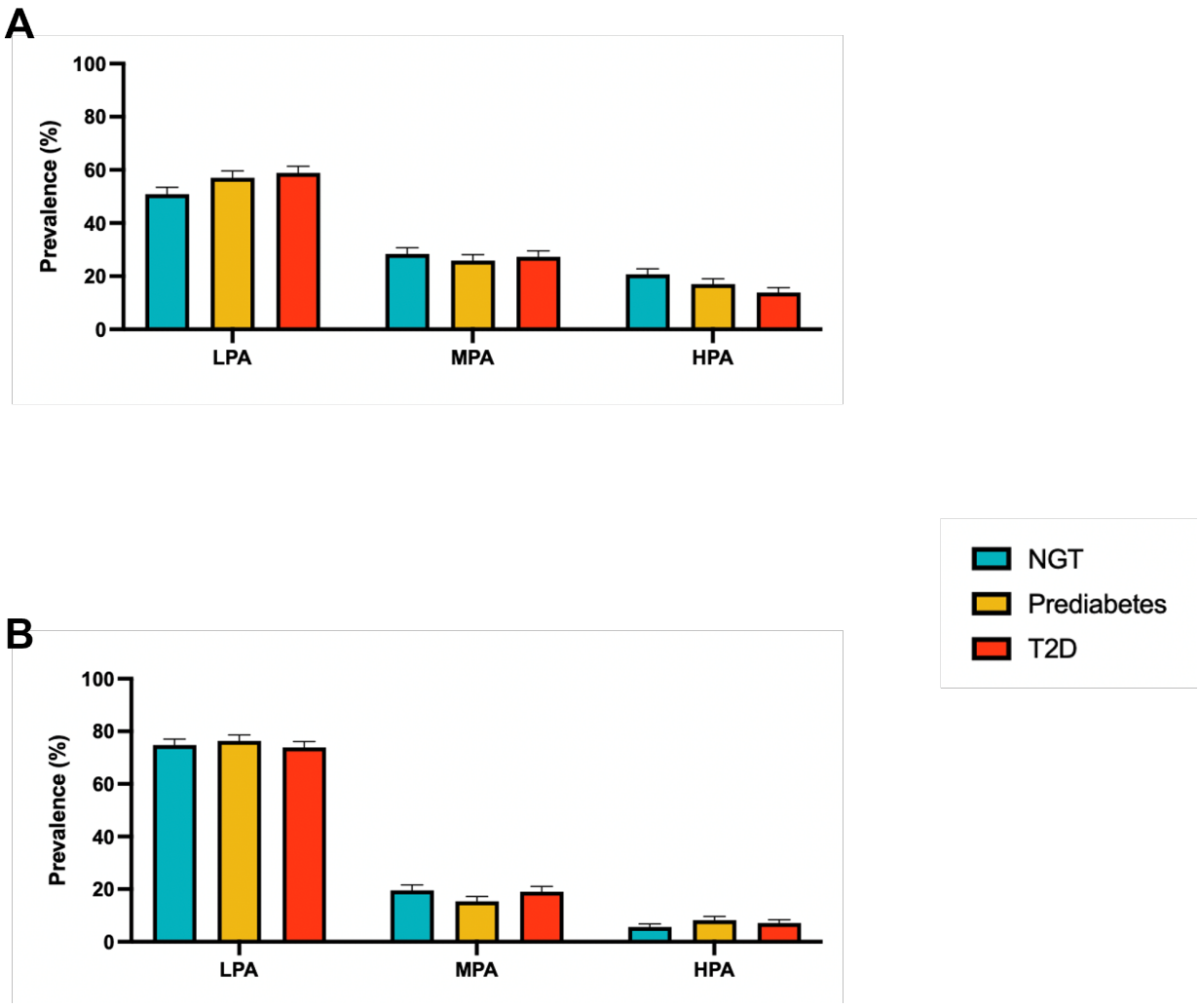
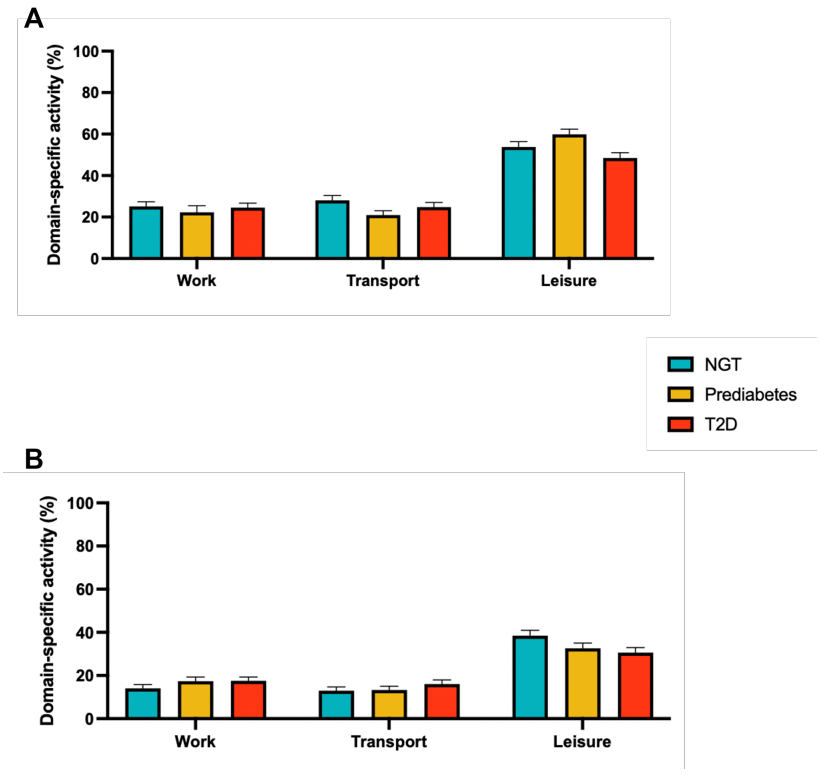


Figure 5.2 Bar charts demonstrating prevalence of low physical activity (LPA), moderate physical activity (MPA), and high physical activity (HPA) in Arab-Emirati males (A) and females (B) with Normal glucose tolerance (NGT) compared to those with prediabetes or type 2 diabetes (T2D). Physical activity was measured in 684 Arab-Emirati males and 749 females from the EM2 sample population using the Global Physical activity questionnaire (GPAQ). Individuals' total physical activity (TPA) was computed as the sum of total MET minutes of activity in a typical week across three domains: work, transport, and leisure.



Mean (SD) sedentary time in this sample was 683.3 (245.3) minutes/day with no significant differences between sexes; 684.6 (239.7) minutes/day in females and 681.9 (251.4) minutes/day in males. Sedentary time among females with T2D, however, was significantly lower in comparison to their NGT counterparts (p = 0.020) (**Table 5.2**). No significant differences in sedentary time were observed among males across diabetes status groups (**Table 5.2**).

Figure 5.3 Bar charts demonstrating prevalence of Arab-Emirati males (A) and females (B) who participate in moderate/vigorous activities at work, transport or leisure time. Males and females with Normal glucose tolerance (NGT) were compared to those with prediabetes or type 2 diabetes (T2D). Physical activity was measured in 684 Arab-Emirati males and 749 females from the EM2 sample population using the Global Physical activity questionnaire (GPAQ). Individuals' total physical activity (TPA) was computed as the sum of total MET minutes of activity in a typical week across three domains: work, transport, and leisure. Domain-specific physical activity prevalence rates were calculated based on the proportion of individuals who participated in any moderate or vigorous physical activity at work, transport, or leisure for at least 10 minutes in a typical day of the week.



5.5.1.2 The impact physical inactivity on glycaemic, lipid, and adiposity measures in Arab-Emiratis with or without T2D

Next, I investigated the impact of physical inactivity on glycaemic disturbances, lipid and adiposity measures among Arab-Emiratis. Overall, inactive individuals (LPA) were significantly older, with higher BMI, waist circumference, WHR, fasting insulin levels, and insulin resistance (HOMA-IR) (**Table 5.3**). Similar results were observed irrespective of diabetes status.

Individuals with NGT and low physical activity levels (LPA) had significantly higher insulin resistance as measured by HOMA-IR ($p = 0.020$), as well as higher insulin ($p = 0.022$) and triglyceride levels ($p = 0.031$), and WHR ($p = 0.005$), in comparison to those who had moderate or high activity levels (MPA/HPA) (**Table 5.3**). Similarly, individuals with T2D and LPA had significantly higher BMI ($p = 0.001$), waist circumference ($p = 0.005$), insulin resistance (HOMA-IR) ($p < 0.001$), and insulin levels ($p < 0.001$), compared to their counterparts who had moderate/high physical activity levels (**Table 5.3**). No significant differences were observed across activity groups in individuals with prediabetes (**Table 5.3**).

5.5.1.3 The association of physical activity parameters with T2D, glycaemic disturbances and anthropometric measures

Regression analyses showed no significant association of LPA or domain (work, transport, leisure)-specific physical activity with T2D (**Table 5.4**). However, in females, LPA was significantly associated with increased HOMA-IR, independently of BMI ($p = 0.028$). After adjustment for age and BMI, females with LPA have 5.7% higher HOMA-IR compared to those with MPA/HPA (**Table 5.4**). A positive association of LPA with HOMA-IR was also observed in males, but results were not significant. Moreover, higher leisure time physical activity was associated with decreased HOMA-IR, independently of BMI, among both

males (p = 0.002) and females (p = 0.010). After adjustment for age and BMI, a 1 minute/day increase in leisure-time activity is associated with a decrease of 0.15% and 0.13% in HOMA-IR in males and females, respectively (**Table 5.4**). LPA was also significantly associated with increased triglyceride levels in females (p = 0.004) (**Tables 5.5**). After the adjustment for age and BMI, females with LPA have 0.18 mmol/L higher triglyceride levels, compared to active individuals (MPA or HPA).

Sedentary time was significantly associated with T2D with overall small effect sizes (OR < 1.1) (**Table 5.4**). Increased sedentary time was also significantly associated with increased HOMA-IR, independently of BMI. After adjustment for age and BMI, a 1 hour/day increase in sedentary time increases HOMA-IR by 0.42% (p = 0.036) in males and 0.55% (p = 0.043) in females (**Table 5.4**). Regression analyses also showed that sedentary time was associated with decreased HDL levels and increased triglycerides (only significant in females) (**Table 5.5**). After adjustment for age and BMI, a 1 hour/day increase in sedentary time decreases HDL by 0.01 mmol/L (p = 0.036) and increases triglycerides by 0.02 mmol/L (p = 0.004). Higher sedentary time was also significantly associated with increased BMI, waist circumference, and VAT in both sexes. Overall, a 1 hour/day increase in sedentary time increases BMI by 0.15-0.19 Kg/m^2, waist circumference by 0.18-0.19 cm, and VAT volume by 0.02 L (**Table 5.6**).

5.5.1.4 Physical activity levels as measured by Actigraph GT9X in a subset of Arab-Emiratis

I collected data on a subset of individuals from the EM2 sample population using Actigraph GT9X activity monitors to test the validity of GPAQ-derived parameters. 143 Arab-Emirati participants from the EM2 sample population participated in this part of the study. Nine individuals failed to return the Actigraph monitors (6.3%). Valid data as defined by wear time of at least 3 days including at least 1 weekend day was available for 106 participants (74.1%). These included 47 males and 59 females, and overall, 35 with

NGT, 35 with prediabetes, and 36 with T2D. Individuals with T2D were significantly older than those with NGT, with a mean (SD) age of 40.4 (3.8) and 37.2 (4.8) years, respectively (**Table 5.7**). Overall, participants in this part of the study (n = 106) had similar age and BMI to the rest of the sample (n = 1,327), with no significant differences observed. Furthermore, no significant differences were observed across the 3 diabetes status groups (NGT, prediabetes, T2D) in BMI, calories burned (Kcal/day), sedentary time (minutes/day), or time spent in light, moderate, or vigorous activities (minutes/day) (**Table 5.7**). Overall, individuals burned mean (SD) 1,324 (653) Kcal/day. Mean (SD) sedentary time, and time spent in light, moderate, and vigorous activities across all participants was: 440 (198.9), 459 (122.2), 111 (55.1), and 0.44 (2.0) minutes/day, respectively.

5.5.1.5 Data validity: GPAQ versus Actigraph GT9X

Overall, individuals tended to over-report sedentary time: sedentary time reported by individuals in GPAQ was significantly higher than that measured by Actigraph GT9X ($p < 0.001$), with an average difference of 195 minutes (**Supplementary Table 5.5**). In contrast, GPAQ-reported time spent in moderate activities was significantly lower than that measured by Actigraph ($p < 0.001$).

There was also significant over-reporting of time spent in vigorous activities, which was significantly higher than that measured by Actigraph ($p < 0.001$) (**Supplementary Table 5.5**).

These differences were not observed among individuals with T2D: in my diabetes status groups (NGT, prediabetes, T2D)-specific comparisons; self-reported levels of physical activity among individuals with T2D were similar to those measured by Actigraph. Regression analyses in this sample (n = 106) showed no association of sedentary time with T2D or insulin resistance (HOMA-IR) (**Supplementary Table 5.6**).

5.5.2 Cross-ancestry comparisons

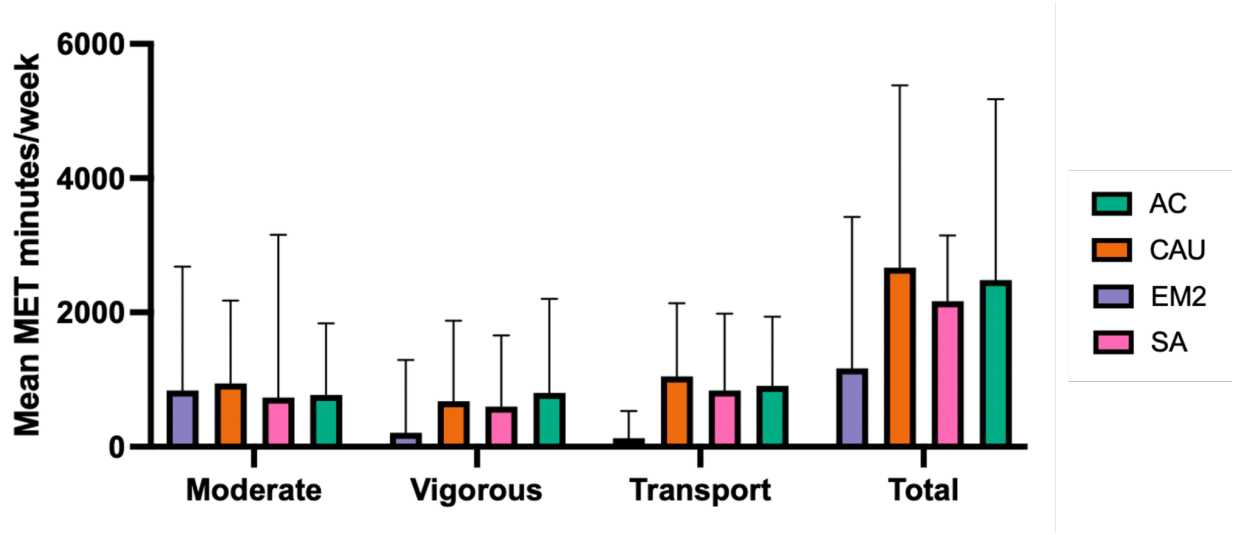
5.5.2.1 Differences in physical activity levels in Arab-Emiratis compared to other ethnic groups

Cross-sectional baseline physical activity questionnaire data on 344,616 Caucasians (CAU), 6,029 South Asians (SA), and 4,201 Afro-Caribbeans (AC) were retrieved from the UK Biobank (UKBB).

Overall, Arab-Emiratis had approximately 2-fold lower total physical activity levels (TPA) compared to CAU, SA, and AC individuals, with the largest differences in time spent in transport activities (MET_walk) and vigorous activities (MET_vig) (**Table 5.8, Figure 5.4**). Results were slightly different when stratified by sex and diabetes status.

TPA levels were significantly lower in Arab-Emirati individuals with or without T2D, compared to their counterparts of CAU, SA, and AC descent ($p < 0.001$) (**Table 5.8**). TPA was highest among CAU males and females, both with and without T2D, with an overall mean (SD) of 2665 (2718) MET minutes/week, compared to 2166 (2429) MET minutes/week in SA, 2482 (2697) MET minutes/week in AC, and 1170 (2254) MET minutes/week in Arab-Emiratis. Similarly, Arab-Emiratis had the lowest frequency of vigorous and transport activities per week across sexes and diabetes status groups ($p < 0.001$) (**Table 5.9**).

Figure 5.4 Bar charts demonstrating mean MET minutes per week spent in total, moderate, vigorous, and transport activities in 1,433 Arab-Emirati (EM2), 344,616 Caucasian (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) (A) males and (B) females Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort.



The frequency of moderate activities in males with T2D, however, was not different between ethnic groups. In addition, Arab-Emirati and SA males with NGT reported similar frequencies of moderate activities (p = 0.20); 983 (1991) and 714 (1058) MET minutes/week, respectively (**Table 5.9**).

That was not the case in females; Arab-Emirati females reported the lowest frequencies of moderate intensity activities, in comparison to the other ethnic groups (p < 0.001), irrespective of diabetes status (**Table 5.9**).

5.5.2.2 The contribution of physical inactivity towards risk of T2D and its related metabolic and anthropometric measures across ethnic groups

As in Arab-Emiratis, regression analyses showed no association of TPA with T2D in all other ethnic groups. An increase of 80 MET minutes/week in total, moderate (MET_mod), vigorous physical activity (MET_vig), or transport (MET_walk) physical activity levels is not associated with T2D, with ORs of ~ 1 in all the 4 ethnic groups (**Table 5.11**). LPA was significantly associated with T2D in CAU and SA individuals. After adjustment for age, sex, and BMI, ORs of LPA for T2D was 1.4 (CI 95%: 1.3-1.4) and 1.2 (CI 95%: 1.2-1.3) among individuals of CAU and SA descent, respectively (**Table 5.11**). That association was not observed among Arab-Emiratis and AC individuals.

In addition to T2D, LPA was significantly associated with increased HbA1c levels among CAU and SA individuals (**Table 5.12**). Although the β coefficients were numerically larger in Arab-Emiratis, this did not reach statistical significance.

LPA was also associated with decreased HDL and increased triglyceride levels in CAU, AC, and SA individuals, independently of BMI ($p < 0.001$) (**Table 5.12**). An inverse association with HDL was also observed in Arab-Emiratis, but this was not significant.

Lastly, the association of LPA with increased adiposity was observed across all ethnic groups, and that included increased BMI, waist circumference and VAT (**Table 5.13**). Larger effects (as measured by β coefficients) for waist circumference and VAT were observed in Arab-Emiratis, compared to other ethnic groups. Regression coefficients for VAT were significantly higher in Arab-Emiratis compared to all other ethnic groups ($p < 0.05$). Whereas regression coefficients for waist were significantly higher in Arab-Emiratis compared to CAU and AC individuals ($p < 0.05$), but not their SA counterparts (**Table 5.13**).

The magnitude of T2D risk that is attributable to physical inactivity, as measured by LPA, was highest among Arab-Emirati males (**Table 5.14**). After adjustment for age, sex, and BMI, LPA can explain 8.6%, 6.9%, 6.3%, and 0.4% of T2D risk among Arab-Emirati, CAU, SA, and AC males, respectively (**Table 5.14**). Unlike in males, LPA cannot explain T2D risk among Arab-Emirati females (PAR%: -7.3%).

Conversely, sedentary time of over 12 hours per day is attributable to 8.9% and 5.3% of T2D risk among Arab-Emirati males and females, respectively (**Table 5.14**).

5.5.3 Sensitivity analyses

Given that physical activity data was not available for the EM1 sample population, I could not test the contribution of lack of physical activity towards the excess risk of T2D in Arab Emiratis. However, my sensitivity analysis demonstrated the weak contribution of physical inactivity towards the excess risk of T2D among SA and AC individuals, compared to CAU (**Supplementary Table 5.7**).

As part of my sensitivity analyses, I also evaluated the possibility of selection bias of the sub-sample of individuals who agreed to participate and wear the Actigraph GT9X, as compared to the rest of the EM2 sample population. Overall, individuals who had valid Actigraph GT9X data (n = 106) had significantly lower adiposity measures (BMI, waist, WHR, and VAT), higher GPAQ-measured TPA, and lower GPAQ-measured sedentary time, as compared to the rest of the sample (**Supplementary Table 5.8**). This may explain the lack of variation in Actigraph-measured activity levels across diabetes-status groups (NGT vs. Prediabetes vs. T2D)

5.6 DISCUSSION

My results show that Arab-Emiratis living in the UAE have approximately 2-fold lower levels of physical activity than Europeans living in the UK. Overall, the prevalence rates of physical inactivity, as measured by LPA and sedentary time among Arab-Emiratis are similar, irrespective of their diabetes status. Despite that, my results suggest that lack of physical activity may contribute to T2D burden in this population through increased central adiposity, insulin resistance and an unfavourable lipid profile. Moreover, sedentary time appears to be a better indicator of T2D risk among Arab-Emiratis, than overall low physical activity.

Populations in the Gulf region share similar cultural, political, and religious characteristics and beliefs[331]. Although the extreme climate, particularly in summer months, is thought to be one of the biggest barriers for living a physically active lifestyle in the Gulf region, that is not the only one. A study conducted in Saudi Arabia, for example, showed that the lack of physical activity in the population is largely due to ways of commuting, working environments, long working hours, and lack of sports facilities particularly for women[332]. Another study, which was conducted in summer months in the UAE, reported some major barriers for physical activity that the population is facing, and these include; (1) cultural challenges particularly for women, with only limited women-only public gyms, sports clubs or recreational centres, (2) traditional clothes, for both men and women, that hinder physical activity, (3) daily house-hold activities are mostly carried out by domestic helpers, and (4) the main mode of transportation is by car (from door to door)[333].

The prevalence of physical inactivity among Arab-Emirati men and women in this study was 74.3% and 55.7%, respectively. These estimates were similar to those reported previously in the region, despite the differences in study design (case-control vs. nationally representative sample)[262]. In addition, the small

differences in physical inactivity rates between people with or without T2D in this study was also consistent with previous reports in the UAE and the Gulf region[262].

Although a direct association of physical inactivity with T2D was not observed among Arab-Emiratis, results here highlighted the impact of lack of physical activity on T2D-related metabolic disturbances and anthropometric measures. Overall, more active individuals had lower insulin, insulin resistance, and adiposity measures, irrespective of diabetes status. Leisure time physical activity was associated with decreased insulin resistance, independently of BMI in both men and women.

On the other hand, sedentary time was associated with increased insulin resistance. In line with this finding, biological mechanisms have been identified for the beneficial effects of physical activity in glucose metabolism independent of body fatness. Exercise increases insulin-stimulated glycogen synthesis by GLUT4 glucose transporters and through increased glycogen synthase activity[133]. Exercise-related elevated capillary proliferation in muscles, with increased muscle mass, may also contribute to the beneficial effects of physical activity on insulin sensitivity[136, 334]. Other mechanisms relating physical inactivity to insulin resistance include increasing risk of dyslipidaemia and obesity, mitochondrial dysfunction, oxidative damage, and inducing chronic low-grade systemic inflammation[138, 335].

We also found sedentary time to be associated with decreased HDL and increased triglyceride levels among Arab-Emiratis, particularly in females. Consistent with these findings, previous studies have also highlighted the detrimental effect of sedentary behaviour on lipid profile, particularly on HDL and triglyceride levels[139, 336]. In a cross-sectional population-based study of European adults, screen time, as defined by watching television or using a computer, was inversely correlated with HDL-cholesterol, independently of age, sex, education, profession type, income, dietary factors, and smoking[336]. They also

found an inverse relationship between intense physical activity time and triglyceride levels.

Physical activity rates among Arab-Emiratis in this study were 2- to 3-fold lower than in individuals from the UKBB cohort, irrespective of diabetes status. The association of physical inactivity (as measured by LPA) with T2D was only observed among SA and CAU, but not Arab-Emirati or AC individuals. Several reasons can explain these observations. Firstly, it could be a problem of small sample size and insufficient power. Secondly, the impact of physical inactivity on risk of T2D could differ across ethnic groups. More recent multi-ethnic studies support this possible explanation[296, 337]. Thirdly, the cross-sectional design of this study allowed us to investigate only a 'snapshot' of individuals' current physical activity levels, with no knowledge on their physical activity patterns through longer periods of time, or even early-life. Several studies have highlighted the impact of early-life sport practice on risk of T2D, independent of obesity and current physical activity levels[338]. Another source of bias associated with cross-sectional data that may have influenced my findings includes changes in behaviour after diabetes diagnosis. T2D cases, particularly those attending a diabetes centre like ICLDC on a regular basis, may have taken measures to reduce weight, become more active, and eat healthier following diagnosis, especially as it is recommended by their physicians or dieticians every visit. Lastly, although differences in age, sex, and BMI within each ethnic group were controlled for, several other confounding factors that can impact our results need to be taken into consideration, including dietary factors, education, smoking status, and income.

Physical inactivity (LPA) appears to explain slightly a bigger proportion of T2D cases among Arab-Emirati males compared to the other ethnic groups, especially when compared to AC males (8.6% in Arab-Emirati, 6.9% in CAU, 6.3% in SA, and 0.4% in AC males). Moreover, sedentary time of greater than 12 hours per day appears to better explain T2D risk in Arab-Emiratis (as compared to LPA), particularly in

females. In line with this, physiological studies have highlighted the importance of paying attention to all active and inactive behaviours throughout the whole waking day[323]. There are unique molecular, physiological, and clinical effects of long periods of sitting (inactivity physiology), that is separate from the responses caused by structured exercise (exercise physiology)[323]. Nevertheless, evidence at present supports the central role of total duration of sedentary time in mediating risk for T2D, rather than for the pattern of how sedentary time is accumulated[323].

Compared to Europeans (CAU), Arab-Emiratis have lower HDL, higher BMI, waist circumference, WHR, and anthropometry-estimated VAT volume. Our observations here suggest that physical inactivity could contribute to the decreased levels of HDL and the increased central adiposity (not BMI) in Arab-Emiratis, as compared to CAU individuals.

Unfortunately, because insulin data is not available for the UKBB cohort, it was not possible to make trans-ethnic comparisons of the impact of physical inactivity on insulin resistance; a crucial analysis that needs to be done and taken into consideration.

The use of a self-reported measure of sedentary time and physical activity levels as opposed to objective, accelerometer-derived measures must be acknowledged. The use of questionnaires may underestimate the strength of some risk factor relationships with disease[339], and is subject to response bias[340]. Despite that, self-report measures remain the most feasible and affordable instruments for population studies like the current one. On the other hand, it is important to consider the possible influence of wearing an accelerometer on encouraging increased physical activity. In addition, as indicated by my sensitivity analyses, accelerometer-measured activity levels could be subject to selection bias, and thus may produce estimates that are not representative of the general population. In

the Arab-Emirati sample population studied here, I found that participants who agree to wear an accelerometer are less obese and more active, as compared to the rest of the sample.

Estimates in the UKBB could be affected by "healthy volunteer" bias and that is another study limitation that needs to be acknowledged. Overall physical inactivity rates among CAU, SA, and AC individuals in the UKBB cohort was 18.6%. These estimates were lower than those of the WHO[105], but very similar to national statistics published by the National Health Service (NHS)[341]. According to the WHO estimates in 2016, insufficient physical activity rates were 31.5% in men and 40.0% in women in the UK, with an overall prevalence of 35.9%[105]. Whereas, age-specific national estimates by the NHS, reported physical inactivity rates of approximately 20% and 22% among adults aged 45-54 years and 55-64 years, respectively[341].

In conclusion, I show that physical activity levels are about 2-fold lower among Arab-Emiratis and may contribute to the increased central adiposity in the population, compared to Europeans. My results highlight the potential for increased physical activity to reduce insulin levels, increase insulin sensitivity, and improve lipid profile among Arab-Emiratis and, thus, potentially help reduce some of the excess risk of T2D.

A public health mandate that is predominantly focussed on replacing sedentary time with non-fatiguing and potentially more achievable types of light activity in Arab-Emiratis, can help increase overall physical activity levels, and thereby have a high health impact at a population level. Existing evidence indeed highlights the advantageous physiological effects of acquiring behaviours that demand the lowest perceived exertion possible, frequently throughout the entire day, and habitually each day. Nevertheless, my results so far suggest that physical inactivity and increased adiposity cannot fully

account for the excess risk of T2D among Arab-Emiratis as compared to other populations, and a possible increased genetic susceptibility also needs to be investigated.

Table 5.1 Prevalence of physical activity measures in 684 Arab-Emirati males and 749 females from the EM2 sample population. The EM2 sample population was collected in a case-control study design. Physical activity in a typical week was measured using the Global Physical activity questionnaire (GPAQ). Individuals were categorised into those having low physical activity (LPA), moderate physical activity (MPA), or high physical activity (HPA). Domain-specific physical activity prevalence rates were calculated based on the proportion of individuals who participated in any moderate or vigorous physical activity at work, transport, or leisure for at least 10 minutes in a typical day of the week.

	Males										
	NGT			**Prediabetes**				**T2D**			
N	**169**			**170**				**345**			
	Prevalence (%)	**Upper limit**	**Lower limit**	**Prevalence (%)**	**Upper limit**	**Lower limit**	**P**	**Prevalence (%)**	**Upper limit**	**Lower limit**	**P**
LPA	50.88	53.46	48.31	57.06	59.61	54.51	0.30	58.84	61.37	56.31	0.11
MPA	28.40	30.73	26.08	25.88	28.14	23.63	0.69	27.25	29.53	24.95	0.87
HPA	20.71	22.80	18.62	17.06	19.00	15.12	0.47	13.91	15.70	12.13	0.07
Work activity	25.15	27.38	22.91	22.26	25.43	21.07	0.78	24.50	26.71	22.28	0.96

Transport activity	28.07	30.38	25.76	20.93	23.03	18.83	0.16	24.78	27.00	22.56	0.49
Leisure activity	53.80	56.37	51.23	59.88	62.41	57.36	0.30	48.42	51.00	45.84	0.29
	Females										
	NGT			**Prediabetes**				**T2D**			
	286			**195**				**268**			
	Prevalence (%)	**Upper limit**	**Lower limit**	**Prevalence (%)**	**Upper limit**	**Lower limit**	**P**	**Prevalence (%)**	**Upper limit**	**Lower limit**	**P**
LPA	74.83	77.06	72.60	76.41	78.60	74.22	0.77	73.88	76.14	71.62	0.88
MPA	19.58	21.62	17.54	15.38	17.24	13.53	0.29	19.03	21.05	17.01	0.96
HPA	5.59	6.78	4.41	8.21	9.62	6.79	0.35	7.09	8.41	5.77	0.58
Work activity	14.00	15.77	12.20	17.35	19.30	15.40	0.38	17.54	19.30	15.40	0.3
Transport activity	12.94	14.67	11.21	13.27	15.01	11.51	1.00	16.04	17.94	14.15	0.36

Leisure activity	38.46	40.97	35.96	32.65	35.07	30.23	0.23	30.60	32.97	28.22	0.06

P: p values for Chi-squared test performed comparing categorical parameters in prediabetes and T2D groups with NGT. Adjusted for multiple testing.

Table 5.2 Physical activity measures in 684 Arab-Emirati males and 749 females from the EM2 sample population. The EM2 sample population was collected in a case-control study design. Physical activity in a typical week was measured using the Global Physical activity questionnaire (GPAQ). Individuals' total physical activity (TPA) was computed as the sum of total MET minutes of activity in a typical week across three domains: work, transport, and leisure. Work-, transport-, and leisure-specific physical activity was computed in minutes spent participating in moderate/vigorous physical activity in a typical day. Sedentary time was reported for a typical day.

	Males									
	NGT			**Prediabetes**				**T2D**		
	169			**170**				**345**		
	Median IQR	**Mean SD**		**Median IQR**	**Mean SD**		**P**	**Median IQR**	**Mean SD**		**P**
Sedentary time (minutes/day)	702.00 (480.0-864.0)	653.92	274.52	660.00 (536.2-840.0)	663.78	221.22	0.779	720.00 (540.0-900.0)	704.75	252.21	0.106
Total physical activity (MET minutes/week)	1080.00 (80.0-2200.0)	1757.00	2742.60	802.00 (0.0-1915.0)	1732.2	3400.19	0.494	600.00 (0.0-1720.0)	1503.00	2793.89	0.035
Work physical activity (minutes/day)	0.00 (0.0-0.0)	23.79	71.56	0.00 (0.0-0.0)	23.19	76.61	0.822	0.00 (0.0-0.0)	24.37	83.47	0.938

Transport physical activity (minutes/day)	0.00 (0.0-0.0)	9.15	24.86	0.00 (0.0-0.0)	5.18	13.92	0.109	0.00 (0.0-1.4.0)	6.87	15.76	0.459
Leisure physical activity (minutes/day)	8.57 (0.0-38.6)	21.45	27.49	14.20 (0.0-41.7)	23.45	31.68	0.499	0.00 (0.0-30.0)	16.88	23.19	0.100
	Females										
	NGT			**Prediabetes**				**T2D**			
	286			**195**				**268**			
	Median IQR	**Mean SD**		**Median IQR**	**Mean SD**		**P**	**Median IQR**	**Mean SD**		**P**
Sedentary time (minutes/day)	690.00 (510.0-840.0)	660.34	240.42	720.00 (540.0-857.2)	671.26	254.91	0.481	750.00 (600.0-900.0)	720.24	223.57	0.02
Total physical activity (MET minutes/week)	100.00 (0.0-960.0)	758.00	1504.49	120.00 (0.0-849.0)	848.97	1646.1	0.987	0.00 (0.0-960.0)	809.40	1559.54	0.735
Work physical activity (minutes/day)	0.00 (0.0-0.0)	10.65	39.82	0.00 (0.0-0.0)	16.59	51.3	0.322	0.00 (0.0-0.0)	15.62	49.22	0.215

Transport physical activity (minutes/day)	0.00 (0.0-0.0)	1.81	7.83	0.00 (0.0-0.0)	2.73	12.3	0.87	0.00 (0.0-0.0)	1.33	4.88	0.389
Leisure physical activity (minutes/day)	0.00 (0.0-21.0)	12.55	21.9	0.00 (0.0-17.0)	10.06	18.8	0.179	0.00 (0.0-12.9)	11.07	23.52	0.083

P: p values for Wilcoxon rank-sum test performed comparing continuous parameters in prediabetes and T2D with NGT. P values shown were adjusted for multiple testing.

Table 5.3 Adiposity, glycaemic, and lipid measures in 1,433 Arab-Emirati individuals from the EM2 sample population. The EM2 sample population was collected in a case-control study. Individuals with low physical activity (LPA) were compared to those who had moderate or high physical activity (MPA/HPA) in a typical week. Physical activity was measured using the Global Physical activity Questionnaire (GPAQ).

	All								
N	**947**				**486**				
	LPA				**MPA/HPA**				
	Mean	**SD**	**Median**	**IQR**	**Mean**	**SD**	**Median**	**IQR**	**P**
Age (years)	38.20	6.00	38.18	34.4-42.1	37.39	5.68	37.65	33.8-41.4	0.040
BMI (Kg/m^2)	31.71	6.39	31.07	27.3-35.5	30.46	5.91	29.59	26.3-33.6	<0.001
Waist (cm)	98.79	15.02	99.00	89.0-108.0	97.28	13.58	96.00	88.0-105.0	0.030
WHR	0.900	0.090	0.910	0.85-0.97	0.920	0.090	0.930	0.86-0.97	0.015
VAT (L)	3.90	2.25	3.66	2.4-5.3	3.89	2.09	3.66	2.4-5.0	0.968
Glucose (mmol/L)	6.70	2.42	5.80	5.3-7.0	6.79	2.66	5.76	5.3-7.3	0.770
HbA1c (mmol/moll)	44.93	15.87	39.89	34.4-50.3	44.70	16.66	38.80	35.5-49.7	0.418
Insulin (pmol/L)	119.47	81.74	100.50	68.5-148.8	106.53	78.18	88.47	60.2-128.8	<0.001
HOMAIR	5.37	5.10	3.96	2.5-6.6	4.82	4.67	3.42	2.2-5.9	<0.001

HOMAB	132.65	89.96	119.50	75.9-169.9	121.02	91.68	102.21	66.0-153.5	<0.001
TC (mmol/L)	4.65	0.99	4.59	4.0-5.3	4.61	0.99	4.57	4.0-5.3	0.493
LDL (mmol/L)	3.10	0.91	3.04	2.5-3.6	3.07	0.89	3.02	2.4-3.6	0.629
HDL (mmol/L)	1.29	0.35	1.26	1.0-1.5	1.27	0.35	1.23	1.1-1.5	0.463
Triglycerides (mmol/L)	1.48	1.04	1.16	0.83-1.68	1.42	1.56	1.09	0.75-1.65	0.038
	NGT								
N	**300**				**155**				
	LPA				**MPA/HPA**				
	Mean	**SD**	**Median**	**IQR**	**Mean**	**SD**	**Median**	**IQR**	**P**
Age (years)	37.68	6.17	36.97	33.4-41.5	36.4	5.19	36.4	33.2-39.7	0.052
BMI (Kg/m^2)	29.71	6.28	29.3	24.8-33.5	28.7	5.4	27.75	25.2-31.8	0.107
Waist (cm)	91.14	13.87	92	81.0-101.0	91.55	12.29	92	83.0-99.0	0.827
WHR	0.880	0.080	0.870	0.80-0.92	0.860	0.090	0.900	0.82-0.95	0.005
VAT (L)	2.76	1.89	2.61	1.2-4.0	3.04	1.95	2.93	1.5-4.1	0.234
Glucose (mmol/L)	5.29	0.34	5.3	5.1-5.5	5.24	0.32	5.24	5.0-5.5	0.195

HbA1c (mmol/moll)	33.85	3.99	34.43	32.0-36.6	33.56	3.92	34.43	31.2-36.6	0.847
Insulin (pmol/L)	91.81	56.3	79.03	56.9-109.9	83.63	57.93	67.82	52.4-100.2	0.022
HOMAIR	3.14	2.08	2.71	1.9-3.8	2.81	1.99	2.27	1.7-3.4	0.020
HOMAB	146.43	80.52	130.68	95.9-173.5	136.44	84.53	119.41	87.7-162.2	0.064
TC (mmol/L)	4.68	0.96	4.54	4.1-5.1	4.7	0.73	4.66	4.2-5.3	0.342
LDL (mmol/L)	3.09	0.85	2.97	2.5-3.5	3.14	0.71	3.09	2.7-3.6	0.132
HDL (mmol/L)	1.43	0.36	1.4	1.2-1.6	1.41	0.37	1.4	1.2-1.6	0.809
Triglycerides (mmol/L)	1.12	0.91	0.92	0.7-1.3	1.01	0.67	0.79	0.6-1.1	0.031
	Prediabetes								
N	**246**				**119**				
	LPA				**MPA/HPA**				
	Mean	**SD**	**Median**	**IQR**	**Mean**	**SD**	**Median**	**IQR**	**P**
Age (years)	39.5	6.54	39.17	35.1-43.0	38.18	6.05	38.11	34.8-42.1	0.169
BMI (Kg/m^2)	31.62	5.87	31.01	27.6-35.6	30.54	5.97	29.57	26.4-34.1	0.054
Waist (cm)	98.32	14.19	97	88.0-107.0	96.52	12.24	95	89.5-102.5	0.293

WHR	0.900	0.090	0.910	0.84-0.96	0.920	0.080	0.920	0.86-0.96	0.214
VAT (L)	3.82	2.18	3.44	2.3-5.2	3.78	1.91	3.63	2.4-4.7	0.894
Glucose (mmol/L)	6.12	1.86	5.75	5.5-6.1	6.28	2.1	5.78	5.6-6.1	0.585
HbA1c (mmol/moll)	40.1	11.8	38.8	34.4-41.0	40.7	12.83	38.25	35.5-41.0	0.933
Insulin (pmol/L)	112.43	60.62	103.1	68.4-144.6	116.4	86.55	93.45	63.9-142.8	0.575
HOMAIR	4.27	2.5	3.76	2.5-5.5	4.49	3.53	3.61	2.4-5.5	0.607
HOMAB	142.78	73.98	135.27	92.6-182.6	147.19	106.61	130.69	81.2-190.2	0.449
TC (mmol/L)	4.85	0.81	4.74	4.-5.5	4.8	0.83	4.79	4.2-5.3	0.464
LDL (mmol/L)	3.28	0.79	3.21	2.7-3.9	3.28	0.8	3.33	2.7-3.8	0.816
HDL (mmol/L)	1.33	0.33	1.29	1.1-1.5	1.3	0.34	1.27	1.1-1.5	0.429
Triglycerides (mmol/L)	1.28	0.66	1.13	0.8-1.6	1.24	0.73	1.04	0.8-1.5	0.163
	T2D								
N	**401**				**212**				
	LPA				**MPA/HPA**				
	Mean	**SD**	**Median**	**IQR**	**Mean**	**SD**	**Median**	**IQR**	**P**

Age (years)	37.79	5.39	38.42	34.6-41.9	37.67	5.74	38.36	34.2-42.1	0.899
BMI (Kg/m^2)	33.26	6.37	32.69	28.7-36.9	31.71	5.94	30.81	27.3-34.1	0.001
Waist (cm)	104.32	13.87	103	94.0-113.0	101.46	13.64	100	91.0-109.0	0.005
WHR	0.940	0.070	0.940	0.89-0.99	0.940	0.080	0.940	0.89-0.99	0.819
VAT (L)	4.74	2.16	4.52	3.1-6.0	4.52	2.06	4.26	2.9-5.8	0.174
Glucose (mmol/L)	8.11	2.83	7.3	6.2-9.1	8.17	3.14	7.36	6.1-9.0	0.751
HbA1c (mmol/moll)	54.65	16.7	50.82	43.2-62.8	53.44	18.28	49.73	39.9-60.7	0.149
Insulin (pmol/L)	144.39	99.73	120.9	82.3-183.1	117.41	82.58	100.06	66.8-142.7	<0.001
HOMAIR	7.71	6.68	5.77	3.8-9.0	6.42	5.85	4.66	2.9-7.9	<0.001
HOMAB	116.09	102.33	92.87	53.9-151.0	95.75	80.93	79.34	40.6-126.1	0.003
TC (mmol/L)	4.5	1.09	4.42	3.7-5.1	4.43	1.2	4.31	3.6-5.1	0.298
LDL (mmol/L)	2.99	0.99	2.92	2.3-3.6	2.91	1.02	2.74	2.1-3.5	0.193
HDL (mmol/L)	1.16	0.31	1.11	0.9-1.3	1.15	0.29	1.12	1.0-1.3	0.798
Triglycerides (mmol/L)	1.74	1.22	1.44	1.1-2.0	1.83	2.17	1.35	1.1-2.0	0.491

Abbreviations: NGT; normal glucose tolerance, T2D; type 2 diabetes, BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, VAT; estimated visceral adipose tissue volume. P: p values for Wilcoxon rank-sum test performed comparing continuous

parameters between individuals with LPA to those with MPA/HPA in diabetes status group (adjusted for multiple testing).

Table 5.4 Logistic and linear regression analysis showing the association of physical activity measures with type 2 diabetes (T2D) and insulin resistance as measured by homeostatic model assessment (HOMA-IR) in 684 Arab-Emirati males and 749 females from the EM2 sample population. Physical activity in a typical week was measured using the Global Physical activity questionnaire (GPAQ). Individuals' total physical activity (TPA) was computed as the sum of total MET minutes of activity in a typical week across three domains: work, transport, and leisure. Individual's physical activity was categorised into low physical activity (LPA), moderate physical activity (MPA), or high physical activity (HPA). Work-, transport-, and leisure-specific physical activity was computed in minutes spent participating in moderate/vigorous physical activity in a typical day. Sedentary time was reported for a typical day.

	T2D					
	All		**Males**		**Females**	
N	**1,068**		**554**		**514**	
	OR 95% CI	**P**	**OR 95% CI**	**P**	**OR 95% CI**	**P**
LPA						
LPA	0.977 (0.895-1.067)	0.860	1.380 (1.215-1.567)	0.088	0.952 (0.835-1.085)	0.799
LPA + Age	0.968 (0.886-1.057)	0.802	1.357 (1.194-1.543)	0.108	0.958 (0.840-1.093)	0.828
LPA + Age + BMI	0.863 (0.788-0.947)	0.281	1.165 (1.018-1.333)	0.444	0.914 (0.796-1.051)	0.664
TPA (MET minutes/week)						
TPA	1.000 (1.000-1.000)	0.695	1.000 (1.000-1.000)	0.259	1.000 (1.000-1.000)	0.657
TPA + Age	1.000 (1.000-1.000)	0.681	1.000 (1.000-1.000)	0.267	1.000 (1.000-1.000)	0.700
TPA + Age + BMI	1.000 (1.000-1.000)	0.210	1.000 (1.000-1.000)	0.593	1.000 (1.000-1.000)	0.356

Sedentary (hour/day)						
Sedentary	1.054 (1.044-1.065)	0.001	1.045 (1.030-1.061)	0.038	1.069 (1.053-1.086)	0.003
Sedentary + Age	1.056 (1.046-1.067)	0.000	1.047 (1.032-1.062)	0.032	1.075 (1.059-1.092)	0.002
Sedentary + Age + BMI	1.041 (1.030-1.052)	0.010	1.028 (1.013-1.044)	0.211	1.058 (1.042-1.076)	0.017
Work (mins/day)						
Work	1.001 (1.001-1.002)	0.214	1.000 (0.999-1.001)	0.938	1.003 (1.001-1.004)	0.197
Work + Age	1.001 (1.001-1.002)	0.228	1.000 (0.999-1.001)	0.956	1.002 (1.001-1.004)	0.232
Work + Age + BMI	1.001 (1.001-1.003)	0.103	1.000 (1.000-1.001)	0.799	1.004 (1.002-1.005)	0.084
Transport (mins/day)						
Transport	1.000 (0.997-1.002)	0.909	0.994 (0.991-0.997)	0.216	0.988 (0.979-0.998)	0.399
Transport + Age	1.000 (0.997-1.003)	0.994	0.995 (0.991-0.998)	0.249	0.989 (0.980-0.998)	0.414
Transport + Age + BMI	1.001 (0.999-1.004)	0.745	0.995 (0.992-0.999)	0.383	0.985 (0.975-0.995)	0.317
Leisure (mins/day)						
Leisure	0.997 (0.996-0.999)	0.299	0.993 (0.990-0.995)	0.050	0.997 (0.995-1.000)	0.445
Leisure + Age	0.997 (0.996-0.999)	0.317	0.993 (0.990-0.995)	0.053	0.997 (0.995-1.000)	0.449

Leisure + Age + BMI	0.999 (0.997-1.001)	0.734	0.996 (0.993-0.998)	0.235	0.997 (0.994-1.000)	0.444
	HOMA-IR					
	All		**Males**		**Females**	
N	**1,433**		**684**		**749**	
	B 95% CI	**P**	**B 95% CI**	**P**	**B 95% CI**	**P**
LPA						
LPA	4.749 (3.407-6.109)	0.015	7.788 (5.802-9.812)	0.007	5.950 (4.039-7.907)	0.033
LPA + Age	4.707 (3.365-6.067)	0.016	8.080 (6.078-10.120)	0.005	5.908 (3.998-7.864)	0.034
LPA + Age + BMI	2.470 (1.227-3.717)	0.177	3.925 (2.143-5.749)	0.135	5.739 (3.956-7.562)	0.028
TPA (MET minutes/week)						
TPA	0.000 (0.000-0.000)	0.245	0.000 (0.000-0.000)	0.247	0.000 (0.000-0.000)	0.032
TPA + Age	0.000 (0.000-0.000)	0.243	0.000 (0.000-0.000)	0.248	0.000 (0.000-0.000)	0.031
TPA + Age + BMI	0.000 (0.000-0.000)	0.915	0.000 (0.000-0.000)	0.638	0.000 (0.000-0.000)	0.095
Sedentary (hour/day)						
Sedentary	0.823 (0.672-0.975)	<0.001	0.793 (0.572-1.015)	0.015	0.864 (0.662-1.066)	0.004
Sedentary + Age	0.833	<0.001	0.783	0.016	0.894	0.003

	(0.682-0.975)		(0.562-1.005)		(0.692-0.996)	
Sedentary + Age + BMI	0.501 (0.361-0.642)	0.017	0.421 (0.220-0.632)	0.036	0.552 (0.361-0.743)	0.043
Work (mins/day)						
Work	0.000 (0.000-0.010)	0.741	0.010 (-0.010-0.020)	0.707	-0.020 (-0.040-0.000)	0.457
Work + Age	0.000 (-0.010-0.010)	0.750	0.010 (0.000-0.020)	0.689	-0.020 (-0.040-0.000)	0.422
Work + Age + BMI	0.010 (0.000-0.020)	0.397	0.010 (0.000-0.020)	0.443	-0.010 (-0.030-0.000)	0.603
Transport (mins/day)						
Transport	0.040 (0.000-0.090)	0.495	-0.040 -0.100 0.010	0.552	0.120 0.020 0.210	0.399
Transport + Age	0.050 (0.000-0.090)	0.479	-0.050 -0.100 0.000	0.515	0.120 0.020 0.210	0.394
Transport + Age + BMI	0.080 (0.040-0.120)	0.199	-0.020 -0.070 0.030	0.767	0.100 0.010 0.190	0.433
Leisure (mins/day)						
Leisure	-0.160 (-0.180--0.130)	<0.001	-0.210 (-0.250--0.180)	<0.001	-0.160 (-0.200--0.120)	0.003
Leisure + Age	-0.160 (-0.180--0.130)	<0.001	-0.220 (-0.250--0.180)	<0.001	-0.160 (-0.190--0.120)	0.004
Leisure + Age + BMI	-0.100 (-0.130--0.080)	0.003	-0.150 (-0.180--0.120)	0.002	-0.130 (-0.160--0.100)	0.010

HOMA-IR was log_{10} transformed for linear regression analyses. β values (presented here as % change) for association with HOMA-IR are used to compute percentage change in phenotype in predictor with the following formula: (exp(β) - 1)*100. Odds ratios and β coefficients were computed for every change in: 80 MET minute/week in TPA, 60 minutes/day in sedentary time, 1 minute/day in work activity, 1 minute/day in transport activity, 1 minute/day in leisure activity, and for low physical activity (LPA) versus moderate/high physical activity.

Table 5.5 Linear regression analysis showing the association of Low physical activity (LPA) and sedentary time with glycaemic and lipid measures in 684 Arab-Emirati males and 749 females from the EM2 sample population. Physical activity in a typical week was measured using the Global Physical activity questionnaire (GPAQ). Individuals' total physical activity was computed as the sum of total MET minutes of activity in a typical week across three domains: work, transport, and leisure. Individual's physical activity was categorised into low physical activity (LPA), moderate physical activity (MPA), or high physical activity (HPA).

	HbA1c (mmol/mol)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	1.594	0.676	2.513	0.242	0.388	-0.521	1.297	0.774
LPA + Age	1.770	0.847	2.693	0.196	0.367	-0.541	1.275	0.785
LPA + Age + BMI	1.436	0.503	2.369	0.299	0.408	-0.492	1.309	0.760
Sedentary (hour/day)								
Sedentary	0.032	0.022	0.042	0.033	0.023	0.014	0.031	0.086
Sedentary + Age	0.031	0.021	0.041	0.036	0.021	0.012	0.030	0.108

Sedentary + Age + BMI	0.029	0.019	0.039	0.049	0.015	0.006	0.023	0.268
	Glucose (mmol/L)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	0.111	-0.029	0.251	0.594	-0.021	-0.151	0.110	0.914
LPA + Age	0.153	0.012	0.294	0.463	-0.017	-0.147	0.114	0.931
LPA + Age + BMI	0.135	-0.007	0.278	0.522	-0.028	-0.159	0.102	0.883
Sedentary (hour/day)								
Sedentary	0.049	0.033	0.066	0.045	0.032	0.018	0.047	0.125
Sedentary + Age	0.048	0.031	0.064	0.052	0.030	0.016	0.044	0.155
Sedentary + Age + BMI	0.046	0.030	0.063	0.060	0.022	0.007	0.036	0.308
	LDL (mmol/L)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**

LPA								
LPA	0.010	-0.040	0.061	0.890	0.074	0.025	0.123	0.312
LPA + Age	-0.001	-0.051	0.049	0.990	0.072	0.023	0.121	0.324
LPA + Age + BMI	0.007	-0.044	0.058	0.926	0.066	0.017	0.116	0.365
Sedentary (hour/day)								
Sedentary	-0.006	-0.012	0.000	0.474	-0.003	-0.008	0.003	0.731
Sedentary + Age	-0.006	-0.012	0.000	0.504	-0.002	-0.007	0.004	0.845
Sedentary + Age + BMI	-0.005	-0.011	0.001	0.553	-0.002	-0.008	0.003	0.779
	HDL (mmol/L)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	0.010	-0.040	0.061	0.890	-0.051	-0.072	-0.030	0.102
LPA + Age	-0.001	-0.051	0.049	0.990	-0.051	-0.072	-0.030	0.105
LPA + Age + BMI	0.007	-0.044	0.058	0.926	-0.052	-0.073	-0.032	0.088
Sedentary (hour/day)								

Sedentary	-0.006	-0.008	-0.004	0.028	-0.009	-0.011	-0.006	0.010
Sedentary + Age	-0.006	-0.008	-0.004	0.030	-0.009	-0.011	-0.007	0.007
Sedentary + Age + BMI	-0.005	-0.006	-0.003	0.096	-0.007	-0.009	-0.005	0.036
	TC (mmol/L)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	-0.007	-0.064	0.049	0.930	0.063	0.009	0.118	0.435
LPA + Age	-0.024	-0.081	0.032	0.770	0.061	0.007	0.116	0.449
LPA + Age + BMI	-0.007	-0.064	0.050	0.935	0.053	-0.002	0.108	0.512
Sedentary (hour/day)								
Sedentary	-0.006	-0.013	0.000	0.509	-0.002	-0.008	0.004	0.816
Sedentary + Age	-0.006	-0.012	0.001	0.544	-0.001	-0.007	0.005	0.938
Sedentary + Age + BMI	-0.005	-0.011	0.002	0.644	-0.001	-0.007	0.005	0.942
	Triglycerides (mmol/L)							
	Males				**Females**			

N	684				749			
	B	95% CI		P	B	95% CI		P
LPA								
LPA	0.010	-0.072	0.092	0.935	0.182	0.134	0.230	0.011
LPA + Age	0.000	-0.082	0.082	1.000	0.180	0.132	0.228	0.012
LPA + Age + BMI	-0.015	-0.099	0.068	0.903	0.178	0.130	0.226	0.012
Sedentary (hour/day)								
Sedentary	0.008	-0.001	0.018	0.559	0.023	0.017	0.028	0.003
Sedentary + Age	0.009	-0.001	0.018	0.545	0.024	0.019	0.029	0.002
Sedentary + Age + BMI	0.008	-0.002	0.017	0.600	0.022	0.017	0.027	0.004

Abbreviations: HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein. β values represent change in outcome. They were computed for: low physical activity versus moderate/high physical activity, 60 minutes/day increase in sedentary time.

Table 5.6 Linear regression analysis showing the association of Low physical activity (LPA) and sedentary time with adiposity measures in 684 Ara-Emirati males and 749 females from the EM2 sample population. Physical activity in a typical week was measured using the Global Physical activity questionnaire (GPAQ). Individuals' total physical activity was computed as the sum of total MET minutes of activity in a typical week across three domains: work, transport, and leisure. Individual's physical activity was categorised into low physical activity (LPA), moderate physical activity (MPA), or high physical activity (HPA). Sedentary time was reported for a typical day.

	BMI (Kg/m^2)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	1.613	1.320	1.906	<0.001	0.118	-0.263	0.498	0.835
LPA + Age	1.668	1.374	1.962	<0.001	0.095	-0.283	0.473	0.866
Sedentary (hour/day)								
Sedentary	0.146	0.112	0.181	0.004	0.178	0.137	0.219	0.004
Sedentary + Age	0.145	0.111	0.180	0.005	0.192	0.151	0.232	0.002
	Waist (cm)							
	Males				**Females**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
N	**684**				**749**			
LPA								
LPA	4.604	3.870	5.339	<0.001	1.606	0.713	2.499	0.225
LPA + Age	4.637	3.899	5.376	<0.001	1.549	0.667	2.432	0.237
LPA + Age + BMI	1.378	1.041	1.715	0.006	1.194	0.720	1.668	0.090
Sedentary (hour/day)								
Sedentary	0.515	0.428	0.602	<0.001	0.504	0.410	0.598	<0.001

Sedentary + Age	0.516	0.429	0.603	<0.001	0.537	0.444	0.629	<0.001
Sedentary + Age + BMI	0.177	0.138	0.217	0.003	0.193	0.142	0.243	0.010
	WHR							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	0.766	0.454	1.077	0.098	0.866	0.389	1.344	0.221
LPA + Age	0.666	0.354	0.978	0.150	0.834	0.362	1.305	0.233
LPA + Age + BMI	0.185	-0.117	0.487	0.680	0.781	0.321	1.242	0.253
Sedentary (hour/day)								
Sedentary	0.142	0.105	0.178	0.010	0.236	0.186	0.286	0.002
Sedentary + Age	0.145	0.108	0.182	0.008	0.254	0.204	0.303	0.001
Sedentary + Age + BMI	0.096	0.061	0.132	0.068	0.206	0.157	0.255	0.005
	VAT (L)							
	Males				**Females**			
N	**684**				**749**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA								
LPA	0.695	0.577	0.812	<0.001	0.180	0.077	0.283	0.239
LPA + Age	0.669	0.551	0.787	<0.001	0.168	0.068	0.268	0.258
LPA + Age + BMI	0.106	0.060	0.151	0.119	0.132	0.071	0.193	0.142
Sedentary (hour/day)								
Sedentary	0.079	0.065	0.092	<0.001	0.055	0.045	0.066	0.001
Sedentary + Age	0.080	0.066	0.093	<0.001	0.061	0.051	0.072	<0.001
Sedentary + Age + BMI	0.021	0.016	0.026	0.008	0.024	0.018	0.031	0.011

Abbreviations: BMI; body mass index, WHR; waist to hip ratio, VAT; estimated visceral adipose tissue volume
β values represent change in outcome. They were computed for: low physical activity versus moderate/high physical activity, 60 minutes/day increase in sedentary time.

Table 5.7 Activity parameters as measured by Actigraph GT9X in 106 Arab-Emirati individuals from the EM2 sample population. EM2 was collected in a case control study design. Individuals were asked to wear Actigraph GT9X for 7 days and only take it off before sleep. A valid day was defined as a wear time of 10 hours per day. Valid dataset was defined as a wear time of at least 3 days with at least 1 weekend day. Freedson Adult (1998) algorithm was used to calculate METs, cut points, and moderate-vigorous physical activity (MVPA).

	All													
	NGT				**Prediabetes**					**T2D**				
N	**35**				**35**					**36**				
Males (n, %)	**9 (25.71)**				**19 (54.28)**					**19 (52.78)**				
	Mean SD		**Median IQR**		**Mean SD**		**Median IQR**		**P**	**Mean SD**		**Median IQR**		**P**
Age (years)	37.21	4.80	38.18	32.93-40.98	38.93	4.40	38.41	35.77-41.24	0.32	39.99	3.83	40.44	36.80-43.63	0.024
BMI (Kg/m^2)	29.67	6.56	28.76	24.66-32.91	29.99	6.26	29.09	21.05-33.70	0.765	30.93	4.40	30.84	28.10-33.27	0.336
Calories (Kcal/day)	1270.26	639.81	1252.85	852.5-1480.4	1292.73	738.78	1141.87	851.80-1613.30	0.981	1406.53	583.22	1390.50	116.60-1722.60	0.480
Sedentary (min/day)	450.51	188.79	495.00	281.70-599.40	435.70	202.52	452.00	264.90-595.10	0.624	433.87	209.89	408.79	266.60-631.00	0.675
Light activity (min/day)	437.17	119.53	451.11	383.60-515.10	453.04	132.13	479.38	384.50-522.90	0.482	485.35	112.74	499.34	443.60-553.50	0.110
Moderate activity	115.62	56.16	106.12	83.50-142.46	104.82	54.30	95.00	64.57-137.20	0.362	113.35	55.72	105.62	75.66-148.98	0.991

(min/day)														
Vigorous activity (min/day)	0.88	2.89	0.00	0.00-0.00	0.16	0.46	0.00	0.00-0.00	0.873	0.18	0.95	0.00	0.00-0.00	0.875
Very vigorous activity (min/day)	0.09	0.46	0.00	0.00-0.00	0.00	0.02	0.00	0.00-0.00	0.879	0.02	0.14	0.00	0.00-0.00	0.879
MVPA (min/day)	116.59	54.87	107.30	83.50-142.40	104.98	54.11	95.00	64.60-137.20	0.984	113.56	55.79	105.60	75.65-150.0	1.000

Table 5.8 Physical activity parameters in 1,068 Arab-Emirati (EM2), 344,616 Caucasia (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population, which was collected in a case-control study. Physical activity data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort.

	EM2				
	1,068				
	Median	**IQR**	**Mean**	**SD**	
Total Physical activity (MET minutes/week)	409.88	0.0-1400.0	1169.81	2253.97	
Moderate activity (MET minutes/week)	0.00	0.0-960.0	835.01	1849.50	
Vigorous activity (MET minutes/week)	0.00	0.0-0.0	209.68	1082.54	
Walking cycling (MET minutes/week)	0.00	0.0-0.0	126.17	407.70	
	CAU				
	334,616				
	Median	**IQR**	**Mean**	**SD**	**P**
Total Physical activity (MET minutes/week)	1782.00	810.0-3573.0	2665.23	2717.69	<0.001
Moderate activity (MET minutes/week)	480.00	120.0-1200.0	942.78	1232.93	<0.001
Vigorous activity (MET minutes/week)	240.00	0.0-960.0	675.51	1202.66	<0.001
Walking cycling (MET minutes/week)	693.00	297.0-1386.0	1046.94	1088.70	<0.001
	SA				
	6,029				
	Median	**IQR**	**Mean**	**SD**	**P**
Total Physical activity (MET minutes/week)	1356.00	1356.0-2853.0	2165.92	2428.98	<0.001

Moderate activity (MET minutes/week)	360.00	80.0-840.0	730.00	1061.16	<0.001
Vigorous activity (MET minutes/week)	160.00	0.0-597.9	597.93	1142.08	<0.001
Walking cycling (MET minutes/week)	495.00	231.0-990.0	837.99	981.51	<0.001
	AC				
	4,201				
	Median	**IQR**	**Mean**	**SD**	**P**
Total Physical activity (MET minutes/week)	1586.00	691.5-3253.5	2481.51	2696.98	<0.001
Moderate activity (MET minutes/week)	360.00	120.0-960.0	772.66	1063.15	<0.001
Vigorous activity (MET minutes/week)	240.00	0.0-960.0	801.77	1401.11	<0.001
Walking cycling (MET minutes/week)	495.00	247.5-1155.0	907.08	1030.08	<0.001

P: p values for Wilcoxon rank-sum test performed comparing continuous parameters in EM2 compared to other ethnic groups. P values were adjusted for multiple testing.

Table 5.9 Sex- and diabetes status-specific comparisons of physical activity measures in 1,068 Arab-Emirati (EM2), 344,616 Caucasia (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort.

	Males (NGT)										
	EM2		**CAU**			**SA**			**AC**		
N	**169**		**153,498**			**2,609**			**1,546**		
	Median	**IQR**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**
Total Physical activity (MET minutes/week)	1080.00	80.0-2200.0	1869.50	836.8-3813.0	<0.001	1398	594.0-2997.0	<0.001	1671.00	695.0-3675.0	<0.001
Moderate activity (MET minutes/week)	360.20	0.0-1200.0	480.00	120.0-1200.0	<0.001	300	60.0-840.0	0.198	360.00	80.0-960.0	0.005
Vigorous activity (MET minutes/week)	0.00	0.0-180.0	240.00	0.0-960.0	<0.001	160	0.0-840.0	<0.001	360.00	0.0-1200.0	<0.001
Walking cycling (MET minutes/week)	0.00	0.0-120.0	693.00	297.0-1386.0	<0.001	495	231.0-1039.5	<0.001	495.00	247.5-1188.0	<0.001
	Males (T2D)										
	EM2		**CAU**			**SA**			**AC**		

N	345		10,594			745			302		
	Median	**IQR**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**
Total Physical activity (MET minutes/week)	600.00	0.0-1720.0	1386.00	518.0-3039.0	<0.001	1164.00	518-2424.0	<0.001	1540.00	627.0-3066.0	<0.001
Moderate activity (MET minutes/week)	340.80	0.0-1200.0	320.00	0.0-1000.0	0.152	280.00	0.0-720.0	0.79	400.00	108.0-960.0	0.15
Vigorous activity (MET minutes/week)	0.00	0.0-0.0	0.00	0.0-480.0	<0.001	0.00	0.0-560.0	<0.001	240.00	0.0-930.0	<0.001
Walking cycling (MET minutes/week)	0.00	0.0-60.0	594.00	247.5-1386.0	<0.001	462.00	198.0-924.0	<0.001	495.00	247.5-990.0	<0.001
	Females (NGT)										
	EM2		**CAU**			**SA**			**AC**		
N	**286**		**175,322**			**2,330**			**2,068**		
	Median	**IQR**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**
Total Physical activity (MET minutes/week)	100.00	0.0-960.0	1746.00	813.0-3426.0	<0.001	1357.50	613.1-2844.0	<0.001	1357.50	613.1-2844.0	<0.001
Moderate activity (MET	0.00	0.0-720.0	480.00	120.0-1200.0	<0.001	360.00	80.0-960.0	<0.001	360.00	120.0-960.0	<0.001

minutes/week)											
Vigorous activity (MET minutes/week)	160.00	0.0-0.0	160.00	0.0-720.0	<0.001	80.00	0.0-640.0	<0.001	240.00	0.0-960.0	<0.001
Walking cycling (MET minutes/week)	0.00	0.0-0.0	693.00	330.0-1386.0	<0.001	495.00	247.5-990.0	<0.001	495.00	247.5-1039.5	<0.001
	Females (T2D)										
	EM2		**CAU**			**SA**			**AC**		
N	**268**		**5,202**			**345**			**285**		
	Median	**IQR**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**	**Median**	**IQR**	**P**
Total Physical activity (MET minutes/week)	0.00	0-960	1335.00	518.0-2946.0	<0.001	1253.00	506.5-2400.0	<0.001	1386.00	624.0-2670.0	<0.001
Moderate activity (MET minutes/week)	0.00	0.0-795.0	360.00	40.0-960.0	<0.001	320.00	40.0-840.0	<0.001	400.00	120.0-840.0	<0.001
Vigorous activity (MET minutes/week)	0.00	0.0-0.0	0.00	0.0-480.0	<0.001	0.00	0.0-480.0	<0.001	120.00	0.0-720.0	<0.001
Walking cycling (MET minutes/week)	0.00	0.0-0.0	495.00	247.5-1188.0	<0.001	495.00	231.0-742.5	<0.001	231.00	495.0-1039.5	<0.001

P: p values for Wilcoxon rank-sum test performed comparing continuous parameters in EM2 compared to other ethnic groups. P values were adjusted for multiple testing.

Table 5.10 Prevalence of Low physical activity (LPA) across in 1,068 Arab-Emirati (EM2), 344,616 Caucasia (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort.

	Males (NGT)													
	EM2			**CAU**				**SA**				**AC**			
N	**169**			**153,498**				**2,609**				**1,546**			
	Prev. (%)	**95% CI**		**Prev. (%)**	**95% CI**		**P**	**Prev. (%)**	**95% CI**		**P**	**Prev. (%)**	**95% CI**		**P**
LPA (%)	50.88	53.46	48.31	18.05	20.03	16.07	<0.001	24.95	27.18	22.72	<0.001	21.99	24.13	19.86	<0.001
	Males (T2D)													
	EM2			**CAU**				**SA**				**AC**			
N	**345**			**10,594**				**745**				**302**			
	Prev. (%)	**95% CI**		**Prev. (%)**	**95% CI**		**P**	**Prev. (%)**	**95% CI**		**P**	**Prev. (%)**	**95% CI**		**P**
LPA (%)	58.84	61.37	56.31	27.85	30.15	25.54	<0.001	27.92	30.23	25.61	<0.001	22.19	24.33	20.05	<0.001
	Females (NGT)													
	EM2			**CAU**				**SA**				**AC**			

N	286			175,322				2,330				2,068			
	Prev. (%)	95% CI		Prev. (%)	95% CI		P	Prev. (%)	95% CI		P	Prev. (%)	95% CI		P
LPA (%)	74.83	77.06	72.60	17.99	19.97	16.01	<0.001	23.99	26.19	21.79	<0.001	21.52	23.64	19.40	<0.001
	Females (T2D)														
	EM2			CAU				SA				AC			
N	268			5,202				345				285			
	Prev. (%)	95% CI		Prev. (%)	95% CI		P	Prev. (%)	95% CI		P	Prev. (%)	95% CI		P
LPA (%)	73.88	76.14	71.62	28.20	30.52	25.88	<0.001	29.57	31.92	27.21	<0.001	22.81	24.97	20.65	<0.001

Prev.; prevalence rates. P: p values for Chi-squared test performed comparing categorical parameters in EM2 compared to other ethnic groups. P values were adjusted for multiple testing.

Table 5.11 Logistic regression analysis showing the association of physical activity measures with type 2 diabetes in 1,068 Arab-Emirati (EM2), 344,616 Caucasia (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort. Odds ratios are presented for one unit change in each variable.

	EM2			CAU			SA			AC		
N	1,068			344,616			6,029			4,201		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Low physical activity												

LPA	0.96	0.88	1.05	0.741	1.77	1.74	1.79	<0.001	1.22	1.16	1.29	0.007	1.05	0.97	1.12	0.677
LPA + Age + Sex	1.11	1.02	1.22	0.433	1.85	1.83	1.88	<0.001	1.36	1.29	1.43	<0.001	1.18	1.10	1.28	0.134
LPA + Age + Sex + BMI	1.02	0.92	1.12	0.899	1.35	1.33	1.36	<0.001	1.22	1.15	1.28	0.017	1.06	0.97	1.14	0.650
Total Physical activity (MET minutes/week)																
TPA	1.00	1.00	1.00	0.638	0.99	0.99	0.99	<0.001	0.99	0.99	1.00	<0.001	1.00	1.00	1.00	0.079
TPA + Age + Sex	1.00	1.00	1.00	0.524	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	<0.001	1.00	1.00	1.00	0.029
TPA + Age + Sex + BMI	1.00	1.00	1.00	0.913	1.00	1.00	1.00	<0.001	1.00	0.99	1.00	<0.001	1.00	1.00	1.00	0.206
Moderate activity (MET minutes/week)																
Met_mod	1.00	1.00	1.01	0.189	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	0.010	0.99	1.00	1.00	0.726
Met_mod + Age + Sex	1.00	1.00	1.00	0.490	0.9883	0.99	0.99	<0.001	0.99	0.99	0.99	<0.001	0.99	0.99	1.00	0.223
Met_mod + Age + Sex + BMI	1.00	1.00	1.01	0.314	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	<0.001	0.99	1.00	1.00	0.703
Vigorous activity (MET minutes/week)																
Met_vig	0.99	0.99	1.00	0.217	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	0.001	0.99	0.99	0.99	0.016
Met_vig + Age + Sex	0.99	0.99	0.99	0.055	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	<0.001	0.99	0.99	1.00	0.017
Met_vig + Age + Sex + BMI	0.99	0.99	1.00	0.217	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	0.003	0.99	0.99	1.00	0.079

Walking cycling (MET minutes/week)																
Met_walk	1.00	0.99	1.01	0.895	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	0.001	1.00	0.99	1.00	0.343
Met_walk + Age + Sex	0.98	0.97	0.99	0.191	0.99	0.99	0.99	<0.001	0.99	0.99	0.99	0.001	1.00	0.99	1.00	0.227
Met_walk + Age + Sex + BMI	0.98	0.97	0.99	0.236	1.00	0.99	1.00	<0.001	0.99	0.99	1.00	0.023	1.00	1.00	1.00	0.593

Odds ratios for type 2 diabetes were computed for: low physical activity (LPA) versus moderate/high physical activity, 80 MET minutes increase in total, moderate, vigorous, or walking activity.

Table 5.12 Linear regression analysis showing the association of Low physical activity (LPA) with glycaemic and lipid measures in 1,433 Arab-Emirati (EM2), 344,616 Caucasia (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort.

	HbA1c (mmol/mol)															
	EM2				**CAU**				**SA**				**AC**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
N	**1,433**				**344,616**				**6,029**				**4,201**			
LPA	0.616	-0.150	1.381	0.587	0.633	0.616	0.650	<0.001	1.140	0.945	1.335	<0.001	0.186	-0.060	0.431	0.610
LPA + Age + Sex	1.653	0.881	2.426	0.149	0.701	0.685	0.718	<0.001	1.385	1.196	1.575	<0.001	0.534	0.295	0.774	0.132

LPA + Age + Sex + BMI	1.321	0.556	2.086	0.244	0.312	0.296	0.328	<0.001	1.016	0.827	1.205	<0.001	0.305	0.066	0.544	0.390
	LDL (mmol/L)															
	EM2				**CAU**				**SA**				**AC**			
N	**1,433**				**344,616**				**6,029**				**4,201**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA	0.024	-0.017	0.065	0.697	0.002	-0.001	0.004	0.682	0.003	-0.015	0.020	0.923	0.030	0.009	0.052	0.347
LPA + Age + Sex	0.023	-0.019	0.065	0.710	0.005	0.002	0.007	0.240	-0.012	-0.029	0.005	0.644	0.029	0.007	0.051	0.370
LPA + Age + Sex + BMI	0.022	-0.021	0.064	0.730	-0.004	-0.007	-0.002	0.260	-0.004	-0.021	0.014	0.891	0.025	0.003	0.047	0.450
	HDL (mmol/L)															
	EM2				**CAU**				**SA**				**AC**			
N	**1,433**				**344,616**				**6,029**				**4,201**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA	0.012	-0.004	0.028	0.607	-0.089	-0.090	-0.088	<0.001	-0.072	-0.079	-0.065	<0.001	-0.086	-0.096	-0.076	<0.001
LPA + Age + Sex	-0.047	-0.062	-0.032	0.038	-0.086	-0.087	-0.085	<0.001	-0.064	-0.071	-0.058	<0.001	-0.082	-0.091	-0.073	<0.001
LPA + Age + Sex + BMI	-0.036	-0.051	-0.021	0.106	-0.047	-0.048	-0.046	<0.001	-0.047	-0.054	-0.041	<0.001	-0.066	-0.075	-0.057	<0.001

	TC (mmol/L)															
	EM2				**CAU**				**SA**				**AC**			
N	**1,433**				**344,616**				**6,029**				**4,201**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA	0.023	-0.024	0.069	0.744	-0.049	-0.053	-0.046	<0.001	-0.038	-0.061	-0.015	0.265	-0.023	-0.052	0.005	0.581
LPA + Age + Sex	-0.002	-0.050	0.045	0.974	-0.043	-0.046	-0.039	<0.001	-0.052	-0.074	-0.029	0.126	-0.021	-0.049	0.008	0.623
LPA + Age + Sex + BMI	-0.001	-0.049	0.047	0.989	-0.033	-0.036	-0.030	<0.001	-0.033	-0.056	-0.010	0.335	-0.016	-0.044	0.013	0.717
	Triglycerides (mmol/L)															
	EM2				**CAU**				**SA**				**AC**			
N	**1,433**				**344,616**				**6,029**				**4,201**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA	-0.020	-0.083	0.042	0.826	0.183	0.180	0.186	<0.001	0.137	0.114	0.160	<0.001	0.106	0.085	0.126	<0.001
LPA + Age + Sex	0.074	0.011	0.137	0.427	0.184	0.181	0.187	<0.001	0.127	0.105	0.150	<0.001	0.108	0.089	0.128	<0.001
LPA + Age + Sex + BMI	0.058	-0.005	0.121	0.533	0.094	0.091	0.097	<0.001	0.092	0.070	0.115	0.006	0.092	0.073	0.112	0.002

Abbreviations: HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein.
β coefficients show the unit change in each outcome in individuals with low physical activity (LPA) versus those with moderate or high physical activity.

Table 5.13 Linear regression analysis showing the association of Low physical activity (LPA) with adiposity measures in 1,433 Arab-Emirati (EM2), 344,616 Caucasia (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort.

	BMI (Kg/m^2)															
	EM2				**CAU**				**SA**				**AC**			
N	**1,433**				**344,616**				**6,029**				**4,201**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA	1.32	1.04	1.60	0.002	1.49	1.48	1.51	<0.001	0.79	0.71	0.88	<0.001	0.71	0.57	0.84	0.001
LPA + Age + Sex	1.02	0.74	1.31	0.015	1.50	1.49	1.51	<0.001	0.80	0.72	0.89	<0.001	0.76	0.62	0.89	<0.001
	Waist (cm)															
	EM2				**CAU**				**SA**				**AC**			
N	**1,433**				**344,616**				**6,029**				**4,201**			
	B	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**	**B**	**95% CI**		**P**
LPA	1.47	0.78	2.17	0.153	4.38	4.35	4.42	<0.001	2.74	2.51	2.98	<0.001	2.16	1.84	2.48	<0.001
LPA + Age + Sex	3.21	2.53	3.88	0.001	4.37	4.34	4.41	<0.001	2.82	2.60	3.03	<0.001	2.38	2.06	2.69	<0.001
LPA + Age + Sex + BMI	1.33	1.00	1.66	0.007	1.07	1.05	1.08	<0.001	1.07	0.96	1.19	<0.001	0.86	0.71	1.02	<0.001

	WHR															
	EM2				CAU				SA				AC			
N	1,433				344,616				6,029				4,201			
	B	95% CI		P	B	95% CI		P	B	95% CI		P	B	95% CI		P
LPA	-1.31	-1.71	-0.91	0.029	1.85	1.82	1.88	<0.001	1.35	1.18	1.52	<0.001	0.94	0.74	1.14	0.001
LPA + Age + Sex	0.80	0.48	1.13	0.096	1.84	1.82	1.86	<0.001	1.42	1.29	1.55	<0.001	1.16	0.99	1.33	<0.001
LPA + Age + Sex + BMI	0.51	0.20	0.83	0.275	0.73	0.72	0.75	<0.001	0.89	0.77	1.01	<0.001	0.72	0.57	0.88	0.002
	VAT (L)															
	EM2				CAU				SA				AC			
N	1,433				344,616				6,029				4,201			
	B	95% CI		P	B	95% CI		P	B	95% CI		P	B	95% CI		P
LPA	0.01	-0.09	0.12	0.944	0.56	0.55	0.57	<0.001	0.29	0.26	0.33	<0.001	0.21	0.16	0.26	0.003
LPA + Age + Sex	0.45	0.36	0.54	0.001	0.56	0.56	0.57	<0.001	0.31	0.28	0.34	<0.001	0.26	0.22	0.30	<0.001
LPA + Age + Sex + BMI	0.19	0.14	0.24	0.009	0.08	0.07	0.08	<0.001	0.07	0.06	0.09	0.001	0.04	0.02	0.06	0.231

Abbreviations: BMI; body mass index, WHR; waist to hip ratio, VAT; estimated visceral adipose tissue volume. β coefficients show the unit change in each outcome in individuals with low physical activity (LPA) versus those with moderate or high physical activity.

Table 5.14 T2D risk attributed to low physical activity (LPA) and sedentary time amongst 1,068 Arab-Emirati (EM2), 344,616 Caucasian (CAU), 6,029 South Asian (SA), and 4,201 Afro-Caribbean (AC) males and females. Arab-Emiratis were from the EM2 sample population. Data on CAU, SA, and AC individuals was from the baseline UK Biobank cohort. Physical activity was measured using the Global Physical Activity Questionnaire (GPAQ) in the EM2 sample population, and the International Physical Activity Questionnaire (IPAQ) in the UKBB cohort. Sedentary time was not available for the UKBB cohort.

	Males											
Exposure	**Pe**				**OR (95%CI)**				**PAR (%)**			
	EM2	**CAU**	**SA**	**AC**	**EM2**	**CAU**	**SA**	**AC**	**EM2**	**CAU**	**SA**	**AC**
LPA (low physical activity)												
Model 1	0.59	0.19	0.26	0.22	1.38 (1.2-1.6	1.75 (1.7-1.8)	1.17 (1.1-1.2)	1.01 (0.9-1.1)	16.25	12.47	4.23	0.22
Model 2	0.59	0.19	0.26	0.22	1.36 (1.2-1.5)	1.83 (1.8-1.9)	1.33 (1.2-1.4)	1.08 (1.0-1.2)	16.55	13.62	7.90	1.73
Model 3	0.59	0.19	0.26	0.22	1.17 (1.0-1.3)	1.39 (1.3-1.4)	1.26 (1.2-1.3)	1.02 (0.9-1.1)	8.57	6.90	6.33	0.44
Sedentary time (> 12 hours/day)												
Model 1	0.46				1.36 (1.2-1.5)				12.18			
Model 2	0.46				1.39 (1.2-1.6)				12.91			
Model 3	0.46				1.24 (1.1-1.4)				8.90			

	Females											
Exposure	**Pe**				**OR (95%CI)**				**PAR (%)**			
	EM2	**CAU**	**SA**	**AC**	**EM2**	**CAU**	**SA**	**AC**	**EM2**	**CAU**	**SA**	**AC**
LPA (low physical activity)												
Model 1	0.74	0.18	0.25	0.22	0.95 (0.8-1.1)	1.79 (1.7-1.8)	1.33 (1.2-1.4)	1.08 (1.0-1.2)	-3.89	12.45	7.62	1.73
Model 2	0.74	0.18	0.25	0.22	0.96 (0.8-1.1)	1.89 (1.8-1.9)	1.41 (1.3-1.5)	1.32 (1.2-1.5)	-3.08	13.81	9.30	6.58
Model 3	0.74	0.18	0.25	0.22	0.91 (0.8-1.1)	1.26 (1.2-1.3)	1.14 (1.0-1.3)	1.11 (0.9-1.3)	-7.32	4.47	3.38	2.36
Sedentary time (> 12 hours/day)												
Model 1	0.33				1.37 (1.2-1.5)				8.91			
Model 2	0.33				1.41 (1.3-1.6)				9.60			
Model 3	0.33				1.19 (1.1-1.3)				5.27			

Abbreviations: Pe; proportion exposed, OR; odds ratios for type 2 diabetes; PAR; population attributable risk; BMI, body mass index, WHR; waist-to-hip ratio. Population attributable risk (PAR) for EM2 = Pe (proportion of cases exposed) * (OR-1) / OR. Population attributable risk (PAR) for EM1 and UKBB = Pe (proportion of population exposed) * (OR-1) / 1+(Pe (OR-1)). Model 1: Unadjusted ORs (odds ratios), Model 2: Adjusted ORs for age, Model 3: Adjusted ORs for age and BMI.

Chapter Six: Identifying Genetic Effects Underlying T2D in Arab-Emiratis

So far, the results in my thesis highlighted two key findings; (1) Arab-Emiratis have approximately 12-fold excess risk of T2D compared to Europeans, (2) there is an increased exposure to lifestyle risk factors including obesity and lack of physical activity in the population, however, (3) those factors fail to fully account for the differences in magnitude of T2D risk across populations. Data presented thus far drives our hypothesis that Arab-Emiratis may have an increased genetic susceptibility to T2D, which will be investigated and discussed in this chapter.

All analyses presented in this chapter are for training purposes and the limited power and lack of replication cohort are considered in results interpretations and conclusions. Supplementary figures and tables can be accessed through Zenodo using this link: https://zenodo.org/record/7825396#.ZDfrqC0Rro4.

6.1 OVERVIEW

I conducted a genome-wide association study for T2D in 1,200 Arab-Emirati participants from the EM2 sample population, which was collected in a case-control study design. Genotyping was performed using the Infinium Global Screening Array -24 v3.0 BeadChip, which included 654,027 SNPs. Data on additional (untyped) genetic variants were imputed using the Haplotype Reference Consortium (HRC) panel. In addition, I carried out cross-ancestry comparisons and trans-ethnic meta-analysis using Arab-Emirati-,

European-, and South Asian-based summary statistics. My findings show that previously-reported T2D risk loci cannot fully account for the increased risk of T2D amongst Arab-Emiratis, compared to Europeans and South Asians. These findings raise the possibility that increased risk of T2D amongst Arab-Emiratis may be the result of genetic susceptibility factors that are specific to the Arab-Emirati population. For the first time in Arab-Emiratis, I identified 40 novel loci in association with T2D, which potentially provide additional insight into the genetic mechanisms underlying T2D in the population. Overall, my results reinforce the potential for new discoveries from extending the investigation of T2D genetic susceptibility to populations of non-European ancestry.

6.2 AIMS

My primary aim is to investigate the contribution of known and novel common genetic variants to the risk of T2D among Arab-Emiratis, compared to Europeans and South Asians.

6.2.1 Objectives

1. To carry out a genome-wide association analysis on 1,200 Arab-Emiratis from the EM2 sample population to identify genetic effects underlying T2D in the population. This included:
    a. Genotype calling and preparation of binary files for quality control procedures.
    b. Performing quality control procedures and imputation using the Haplotype Reference Consortium (HRC) panel.
    c. Identifying SNPs that are significantly associated with T2D in Arab-Emiratis at P-value < 5.0e-6.
    d. Carrying out conditional analyses at each novel locus (+/−1Mb region) to identify multiple distinct association signals and map them to potential candidate genes.

e. Deriving an Arab-Emirati-specific polygenic risk score to examine the effects of identified loci on predicting risk of T2D, prediabetes, and secondary quantitative anthropometric and metabolic traits in Arab-Emiratis.

2. To evaluate the heterogeneity of cosmopolitan genetic effects underlying T2D across ethnic groups (Arab-Emiratis, Europeans and South Asians), and identify potentially novel loci through trans-ethnic meta-analysis. This included:
    a. Using summary statistics derived from a meta-analysis in 62,892 T2D cases and 596,424 controls of European ancestry (DIAGRAM, GERA, and UKBB) to generate a European-specific polygenic risk score and test its application in predicting T2D, prediabetes, and secondary quantitative anthropometric and metabolic traits in Arab-Emiratis.
    b. Using European-based summary statistics (62,892 T2D cases and 596,424 controls of European ancestry) and South Asian-based summary statistics (15,043 T2D cases and 22,318 controls of South-Asian ancestry) to evaluate heterogeneity in allelic effects across ancestry groups of known T2D loci and the novel loci identified in Arab-Emiratis.
    c. Perform a trans-ethnic meta-analysis of South Asian- and Arab-Emirati-specific summary statistics to identify potential novel loci with homogeneous effects across the two ethnic groups.

Figures 6.1 and **6.2** summarise the two objectives for this part of my study.

Figure 6.1 Flowchart describing objective 1. 1,199 individuals from the EM2 sample population were genotyped. The EM2 sample population was collected in a case-control study design. 1,169 individuals passed quality control procedures and were included in the analyses. The primary Arab-Emirati GWAS to identify risk variants associated with type 2 diabetes (T2D) was performed in 412 individuals with normal glucose tolerance (NGT) and 578 T2D cases.

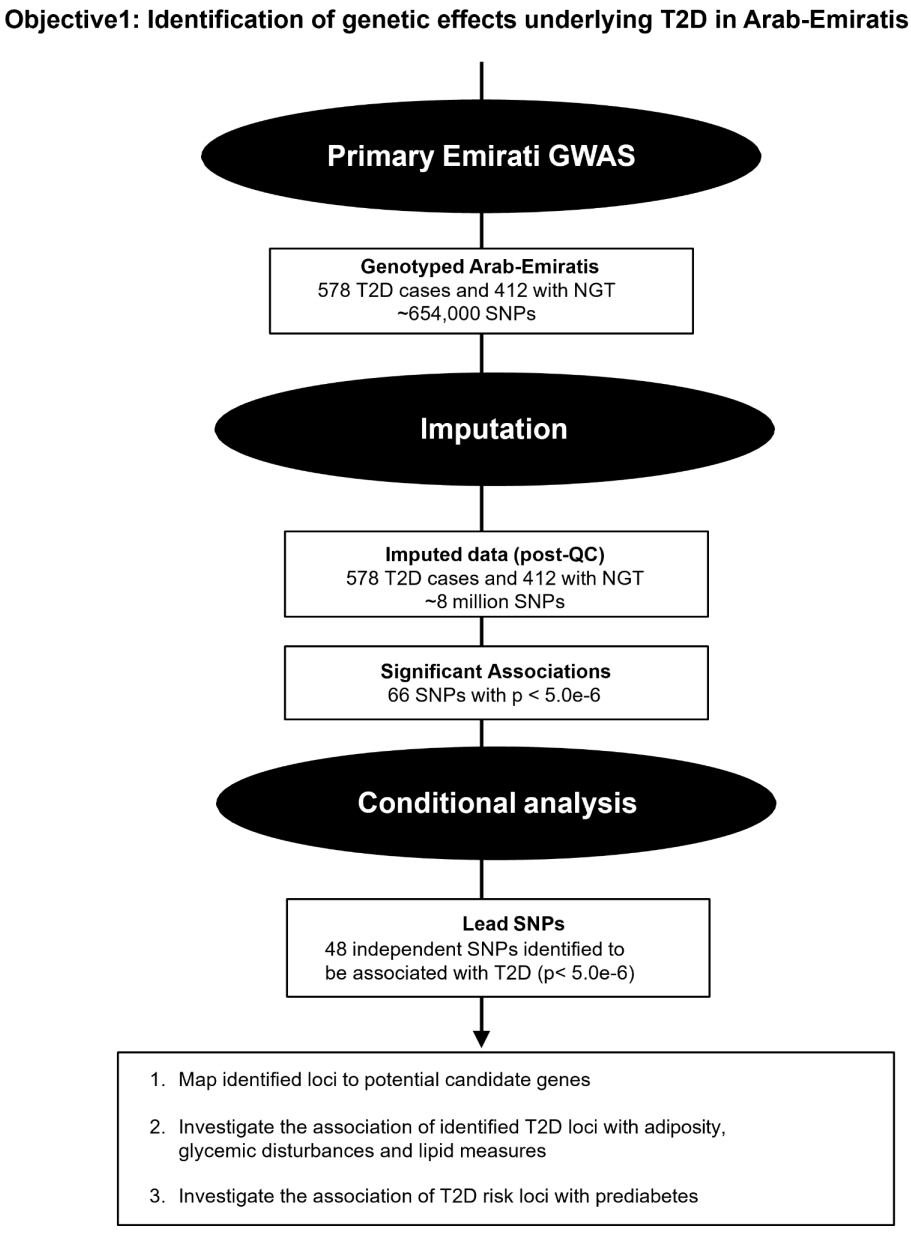
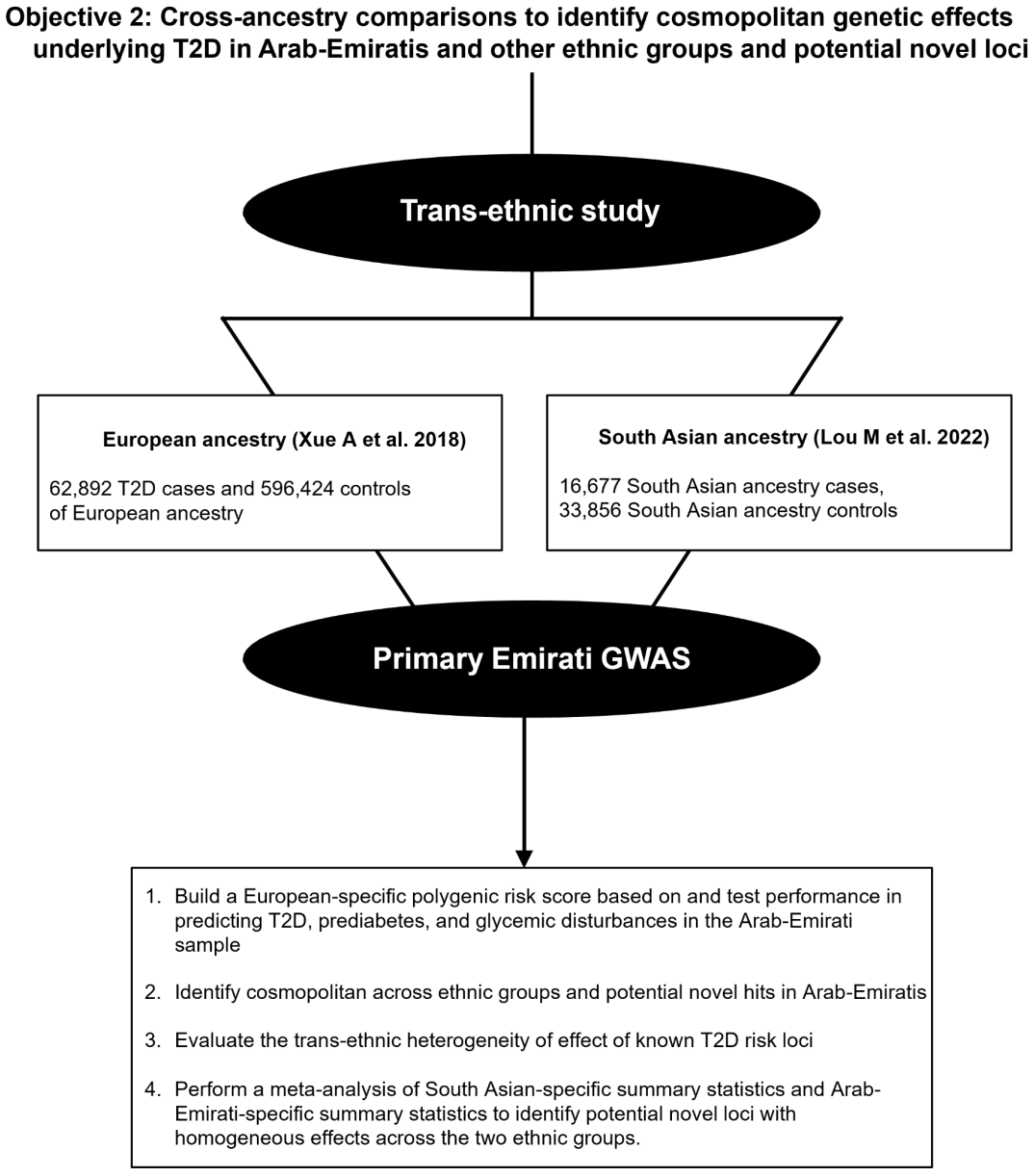


Figure 6.2 Flowchart describing objective 2. Cross-ancestry comparisons were performed using results from (1) the primary Arab-Emirati GWAS performed in 412 controls and 578 type 2 diabetes (T2D) cases, (2) summary statistics derived from a meta-analysis of GWASs in 62,892 T2D cases and 596,424 controls of European ancestry, and (3) summary statistics derived from a meta-analysis of GWAS in 15,043 T2D cases and 22,318 controls of South-Asian ancestry.



6.3 INTRODUCTION

T2D is a leading health problem affecting more than 425 million people worldwide[3]. The burden is especially high in populations of the Arabian Peninsula (AP), including the population of the United Arab Emirates (Arab-Emiratis), with over 991 thousand people living with diabetes in the country, and an age-adjusted prevalence of 16.4% among adults aged 20-79 years[3]. Although data presented in the literature and in this thesis (Chapter 3, 4 and 5) highlight the concurrence of increased obesity and lack of physical activity with increased rates of T2D in the UAE population, Arab-Emiratis have an approximate 12-fold increased risk of T2D that is not fully accounted for by increased obesity or lack of physical activity (Chapter 4 and 5). This has led to the hypothesis that Arab-Emiratis may have an increased genetic susceptibility to T2D.

To date, through genome-wide association studies (GWAS), sequence variants at more than 400 genetic loci have been implicated in the aetiology of T2D[144, 146-151]. These studies were predominantly carried out amongst people of European ancestry. Although, many of these European-derived risk relationships have proven to be cosmopolitan in trans-ethnic meta-analyses investigations[230, 231], population-specific studies have revealed genetic variants associated with T2D that are common in their respective ethnic groups, but rare or absent in others. These included the identification of novel T2D risk loci mapped to *KCNQ1* first in Japanese populations[342], *HNF1A* and *SLC16A11* in Latin American populations[232, 234], *GRB14*, *ST6GAL1*, and *VPS26A* in South Asians[227], as well as *TBC1D4* and *ITGA1* in Greenlandic Inuit[237].

Current estimates indicate that known T2D risk variants account for approximately ~20% of the heritability of the disease, in Europeans at least[149], suggesting that further genetic influences remain to be discovered. Given the distinct geographical location of the UAE, its population substructure is characterised by a continuous spectrum of genotype distribution, with the Middle Eastern component

dominating in the contemporary gene pool of the UAE, with an influence from Central/South Asia, Sub-Saharan Africa and, to a lesser extent, Europe and North Africa regions[239]. In addition, the tribal endogamy and the high consanguinity rates in the UAE population[247, 249] provide the opportunity for preservation of fine-scale genetic structure, and therefore, complex disease risk alleles that are rare or undetectable in other populations, may be amplified by cultural factors to discoverable levels among Arab-Emiratis.

Despite the severity of the health situation in the MENA region, the population remains understudied. Since the start of GWA studies in 2007, the degree to which people of the Arab pan-ethnic group are represented has barely shifted. However, the need for deeper investigation of the genetic basis for T2D and other chronic diseases in these understudied populations is well-recognised[343]. More recent trans-ethnic meta-analyses of GWA studies[344, 345] have repeatedly provided clear evidence of the value of extending the search for complex-trait associations beyond the current focus in Europeans, to populations with different genetic structure[346]. That has however not been the case for Middle Eastern populations. GWA studies in Middle Eastern populations, including the UAE, together with future third generation sequencing efforts will provide further insights into the genetic architecture underlying T2D in these high-risk populations.

6.4 MATERIALS AND METHODS

6.4.1 Study participants

6.4.1.1 Genome-wide association study in Arab-Emiratis

Genome-wide association was carried out amongst 1,200 Arab-Emirati participants from the EM2 sample population. These included 426 controls with normal glucose tolerance (NGT), 589 cases with

T2D, and 185 with prediabetes. Study design, participant recruitment, screening and diagnosis were described in detail in Chapter 2 (section 2.4.3).

6.4.1.2 European GWAS summary statistics

European-specific summary statistics used were derived from a meta-analysis in 62,892 T2D cases and 596,424 controls, performed previously by combining three GWAS datasets of European ancestry[148] : DIAGRAM (Diabetes Genetics Replication And Meta-analysis)[347], GERA (Genetic Epidemiology Research on Ageing)[348], and the full cohort release of the UKBB (UK Biobank)[349]. Data was downloaded from YANGLAB (pub_data | Yang Lab (westlake.edu.cn). Summary statistics from fasting glucose (n = 200,622 non-diabetic individuals) and fasting insulin (n = 151,013 non-diabetic individuals) GWAS meta-analyses were downloaded from the MAGIC website (https://www.magicinvestigators.org/). I used these for further characterisation of identified loci in the primary Emirati GWAS. Note that all summary statistics used were adjusted for age, sex, study-specific covariates and BMI.

6.4.1.3 South Asian summary statistics

South Asian-specific summary statistics used were derived from a meta-analysis performed previously that included 50,533 participants (16,677 T2D cases and 33,856 controls) across sixteen South Asian cohorts (BMI-adjusted model)[345]: BPC (Bangladesh Population Cohort), EpiDREAM (Epidemiologic study of the Screenees for Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication), INDICO (Indian Diabetes Consortium), INTERHEART sub-study-I, INTERHEART sub-study-II, LOLIPOP-IA317 (London Life Sciences Prospective Population Study), LOLIPOP-IA610, LOLIPOP-IAP, LOLIPOP-OmniEE, LOLIPOP-OmniE, PROMIS (1) (Pakistan Risk of Myocardial Infarction Study), PROMIS (2), RHS (Ragama Health Study), SINDI (Singapore Indian Eye Study), UKBB-SA (UK Biobank South Asians). Data was downloaded from GWAS catalogue with the accession number GSST90093110.

6.4.2 Genotyping and genotype calling

One out of the 1,200 blood samples collected was not genotyped due to low DNA quality. Genome-wide genotyping of 1,199 DNA samples was performed by Agiomix Fz LLC (Dubai, UAE), using the Infinium Global Screening Array -24 v3.0 BeadChip, which includes 654,027 SNPs. Genotype calling was performed in GenomeStudio® software 2.0 (Illumina). Final reports were extracted for analysis. A detailed methodology is described in Chapter 2 (section 2.5.4).

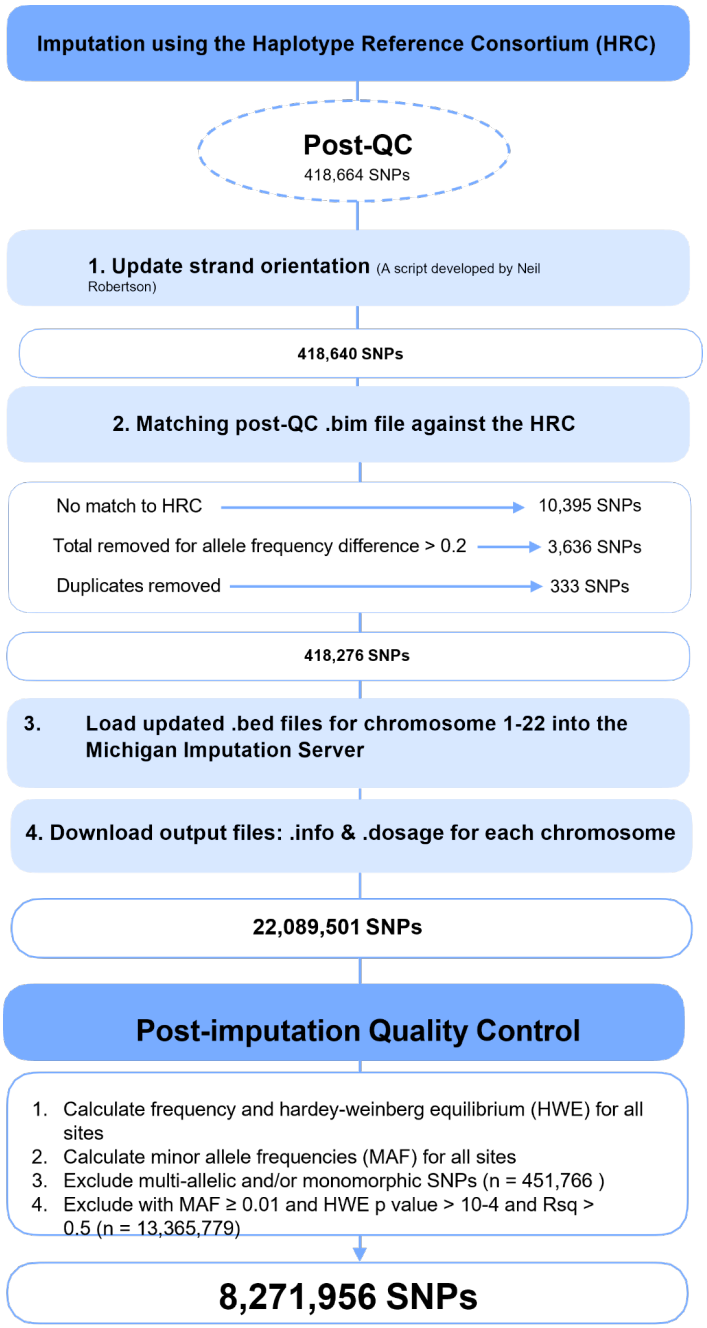
6.4.3 Quality control

Quality control assessment was performed in PLINK (version 1.90) and Genome-wide Complex Trait Analysis (GCTA, version 1.26). SNPs with a call rate < 98% and samples with a call rate < 95% were excluded. Cryptic relatedness was assessed using a set of pruned SNPs in PLINK to perform identical by descent (IBD) analysis. Individuals with evidence of relatedness (pi hat ≥ 0.125) were flagged. Furthermore, a genetic relationship matrix (GRM) was estimated in GCTA using the --make-grm option. Mean heterozygosity was estimated for all individuals, and those with excessive or reduced proportion of heterozygotes were excluded (+/- 3SD of the mean). Sex inconsistencies were initially identified in PLINK, and then verified using mean fluorescence intensities in X and Y chromosome markers as measured by Logratio (LRR). SNPs with minor allele frequencies (MAF) < 1% and/or were deviating from Hardy-Weinberg Equilibrium (HWE) with P-value < 10^{-4}, were excluded.

6.4.4 Imputation

Figure 6.3 summarises the pre- and post-imputation steps that I followed for the primary Arab-Emirati GWAS.

Figure 6.3 Flowchart describing pre- and post-imputation to the HRC panel procedures undertaken for the primary Arab-Emirati GWAS. 1,199 individuals from the EM2 sample population were genotyped. The EM2 sample population was collected in a case-control study design. 1,169 individuals and 418,664 markers passed quality control procedures. In total, 418,640 genotyped and 7,853,680 imputed SNPs were included in the analysis.



6.4.4.1 Pre-imputation checking: aligning data to the HRC panel

To ensure optimal alignment to the reference panel prior to imputation, I used the pre-imputation checking program developed by the McCarthy group (University Oxford, Oxford)[291]. This program compares PLINK marker (.bim) and frequency (.freq) file to the Haplotype Reference Consortium (HRC) panel. It checks all variants in the bim file that match the reference for a variety of potential issues using both location information as well as SNP IDs. Issues checked include chromosome numbers and base pair position, allele frequency, strand alignment, palindromic SNPs, reference allele, and variant naming. Variants not matching to the HRC reference panel or failing one of the checks were listed for removal. I then ran a shell script of PLINK commands that is created to automate the use of these variant lists within PLINK to update or remove variants.

6.4.4.2 Imputation, and post-imputation quality control

After pre-imputation checking, updated files were uploaded into the Michigan Imputation Server for imputation to the HRC panel[350]. Post-imputation, frequency, MAF and HWE for all sites were computed. SNPs that were multiallelic or monomorphic, had MAF < 0.01, HWE P-value < 10^{-4}, and/or imputation had an Rsq (the estimated value of the squared correlation between imputed genotypes and true, unobserved genotypes) < 0.5 were excluded.

6.4.5 Regression analyses

Associations of SNPs with T2D were tested using logistic regression and an additive genetic model. Associations with secondary quantitative anthropometric and metabolic traits were tested using linear regression and an additive genetic model. Two primary models were used; (1) unadjusted, and (2) adjusted for age, sex, and BMI. Analyses were carried out in GCTA, with the --mlma-loco and --grm-bin options used to adjust for the genetic relationship matrix. Associations were tested for T2D (primary

Arab-Emirati GWAS: 412 controls with and 579 cases with T2D), prediabetes (412 controls with NGT and 178 cases with prediabetes), prediabetes and T2D combined (412 controls with NGT and 757 cases with T2D or prediabetes), and for secondary quantitative anthropometric and metabolic traits (1,169 individuals). Quantitative anthropometric and metabolic measures included fasting plasma glucose, HbA1c, fasting insulin, homeostatic model assessment-estimated insulin resistance (HOMA-IR) and β-cell function (HOMA-B), high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides, and total cholesterol (TC). Normality of quantitative measures was assessed using Shapiro tests (**Supplementary Table 4.1**). All quantitative measures were $\log_{10}$ transformed prior to running linear regression analyses. Quantile-Quantile (QQ) plots and genomic control inflation factors were used to assess for inflation and population substructure. Genome-wide significance was set at $P < 5x10^{-6}$ for the primary Arab-Emirati GWAS. Note that a less stringent P value cut-off was used, compared to standard ($5x10^{-8}$), to increase discovery of potential SNP-T2D associations, taking into consideration of the limited power that my GWAS had with a sample size of < 1,000 (after quality control).

6.4.6 Conditional analyses

In an effort to detect multiple distinct association signals at each novel locus (+/- 1Mb region), I performed approximate conditional analyses using GCTA --cojo-cond function in the conditional and joint analysis (COJO) program[351], which performs association analyses of all SNPs included, conditional on a list of given top SNPs). I performed COJO for each chromosome separately, with a P-value threshold of $5x10^{-6}$. COJO analysis calculates linkage disequilibrium (LD) between SNPs, and to avoid multicollinearity, SNPs in high LD ($r2 > 0.9$) with the given SNP are not considered for the COJO analyses. After that, a new top SNP is selected based on the results of the conditional analyses on all given SNPs. The iteration continues until no more SNPs pass the P-value threshold of $5x10^{-6}$.

6.4.7 Regional plots and Arab-Emirati-specific LD matrix

As a measure of LD in the Arab-Emirati sample population, I used squared correlation coefficient between SNP pairs (r2). The r2 parameter was calculated in PLINK using pairwise LD measures for multiple SNPs option. Regional plots were generated for the lead SNPs identified in this study, and r2 was computed for other SNPs in the region indicating the extent of LD each has with the lead SNP. I used a file generated by Github containing a list of all genes and their genomic locations [this was based on data downloaded from the University of California Santa Cruz (UCSC) genome browser database on Human (GR37/Hg19)] to map lead SNP to specific genes[352]. By convention, lead SNPs were mapped according to the nearest gene (+/- 1MB), unless a compelling biological candidate mapped nearby.

6.4.8 Cross-ancestry comparisons

Risk allele frequency and direction of effect of T2D risk loci were compared between Arab-Emiratis, Europeans, and South Asians. These included: (1) lead SNPs identified in the primary GWAS in this study and (2) known T2D risk loci previously reported in Europeans and/or South Asians. A less stringent P-value threshold of $P < 10^{-4}$ was used in cross-ancestry comparisons to maximise the discovery of shared SNP-T2D associations between Arab-Emiratis, European, and South Asians, given the limited power of my primary Arab-Emirati GWAS.

6.4.8.1 Polygenic risk scores

Polygenic risk scores (PRS) were generated in PLINK using the --score and --q-score-range options. The European-specific PRS was derived from summary statistics of a European-based T2D GWAS meta-analysis that identified 139 T2D loci (data described in section 6.4.1.2)[148]. Scores were computed for all individuals in the Arab-Emirati sample (N = 1,169), and their association with T2D, prediabetes, and secondary quantitative anthropometric and metabolic traits in Arab-Emiratis was tested. For training

purposes, PRS was also derived from summary statistics of the current primary Arab-Emirati GWAS, to generate an Arab-Emirati-specific PRS based on the lead SNPs identified. Scores were computed for all individuals and were also tested for their association with T2D, prediabetes, and secondary quantitative anthropometric and metabolic traits in Arab-Emiratis.

6.4.8.2 Trans-ethnic meta-analysis

A fixed-effects inverse variance meta-analysis was used to combine summary statistics from the primary Arab-Emirati GWAS and the South Asian-based meta-analysis (described in section 6.4.1.3). This was performed in PLINK using the --meta-analysis option.

6.4.9 Power calculation

Power calculations were performed in Quanto Version 17.1.0. Post-quality control (QC) and exclusion of individuals with prediabetes, 991 individuals were included in the primary Arab-Emirati GWAS (412 controls with NGT, and 579 T2D cases). Taking an additive disease model, this sample size (cases: controls, 1:0.7) allows us to detect significant associations at P-value < $5x10^{-6}$ for SNPs with MAF ≥ 0.2 and odds ratios (OR) ≥ 1.8 with 76% power, or for SNPs with MAF ≥ 0.2 and OR ≥ 1.9 with 88% power (**Supplementary Table 6.1**).

It is crucial to note that the sample size (n = 1,200) for this part of the study was dependent on the funds available to perform my analyses, and the number of people I could recruit within the duration of my PhD. Thus, I calculated power based on the maximum number of samples that I could recruit and genotype.

6.5 RESULTS

6.5.1 The Primary Arab-Emirati GWAS

6.5.1.1 Quality control and imputing to HRC

Figure 6.4 depicts the quality control procedures followed and the number of individuals and markers excluded at each stage. Five individuals and 10,838 SNPs were excluded due to low sample (< 95%) and marker (< 98%) call rates, respectively. IBD analysis identified 125 individuals with evidence of relatedness (pi hat ≥ 0.125). Those individuals were not excluded, but association results were adjusted for GRM (genetic relationship matrix) and QQ plots and computed genomic inflation factors showed no evidence of population substructure (**Figure 6.5**). Lambda values for the final imputed and GRM-adjusted dataset were 1.01 and 1.13 before and after adjustment for covariates (age, sex, BMI), respectively.

Evidence of excessive or reduced heterozygosity was reported for 14 individuals which were excluded from further analyses. Sex inconsistencies were reported for 11 individuals, which were also excluded from further analyses. In addition to that, genetic sex for 24 individuals was undefined (appeared as SNPSEX = 0 in PLINK's report). All 24 individuals were recorded as females in the phenotype file. As described previously in Chapter 2 (section 2.5.6.1.4), PLINK estimates X chromosome-specific inbreeding coefficient (FX) and classifies a sample as male if FX < 0.2, and female if FX > 0.8. A value of FX between 0.2 and 0.8, or an apparent sex that disagrees with the putative sex, is flagged as a "problem". These individuals (n = 24) had an increased FX, with a mean (SD) of 0.33 (0.10) (**Supplementary Figure 6.2**). In addition, their mean overall inbreeding coefficient (for all autosomes) was significantly higher compared to that of the rest of the sample (n = 1,170) ($p < 0.001$). Furthermore, analysis of their X and Y chromosome variants' intensity files confirmed them as females (**Supplementary Figure 6.3**). Therefore,

those 24 individuals were not excluded from my analysis. Lastly, 224,525 SNPs were excluded due to MAF (n = 221,971 SNPs with MAF < 1%) and/or evidence of deviation from HWE (n = 2,554 SNPs with HWE P-value < 10^{-4}). Post-QC, 1,169 individuals and 418,664 SNPs remained and were included in my analysis. **Table 6.1** summarises characteristics of all individuals that passed QC.

Figure 6.4 Flowchart summarising quality control procedures performed for the study. 1,199 individuals from the EM2 sample population were genotyped using the Infinium Global Screening Array -24 v3.0 BeadChip, which includes 654,027 SNPs. The EM2 sample population was collected in a case-control study design. 1,169 individuals and 418,664 markers passed quality control procedures and were included in the analyses.

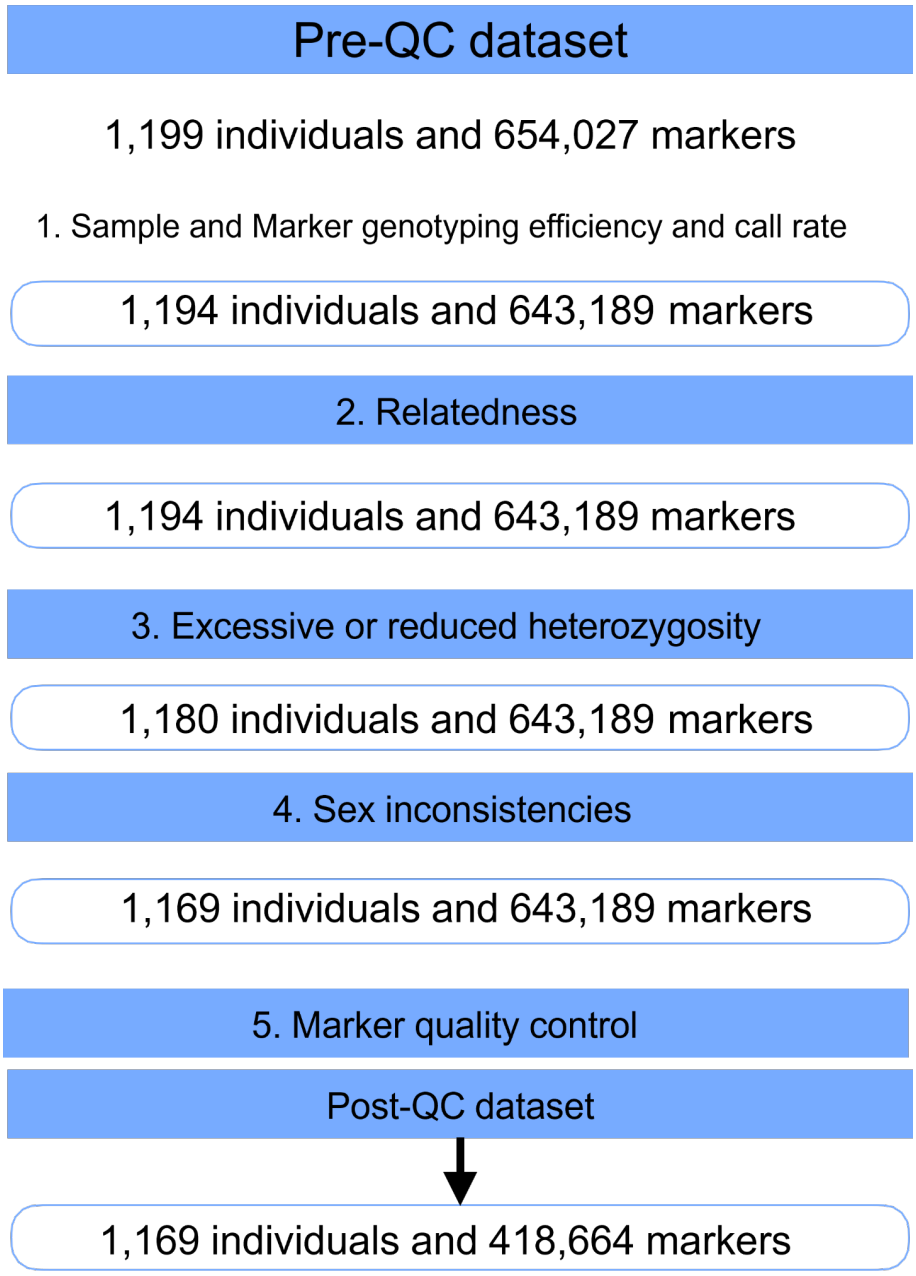


Figure 6.5 Quantile-quantile plots for the primary Arab-Emirati genome-wide association analysis. The primary Arab-Emirati GWAS was performed on 991 Arab-Emirati men and women; 412 with normal glucose tolerance and 578 with type 2 diabetes. QQ plots presented for (A) 418,664 SNPs pre-imputation, (B) 8,271,956 post-imputation SNPs, (C) 8,271,956 post-imputation SNPs, after adjustment for age, sex, and BMI. Genomic inflation factors (lambda) for each QQ plot are shown.

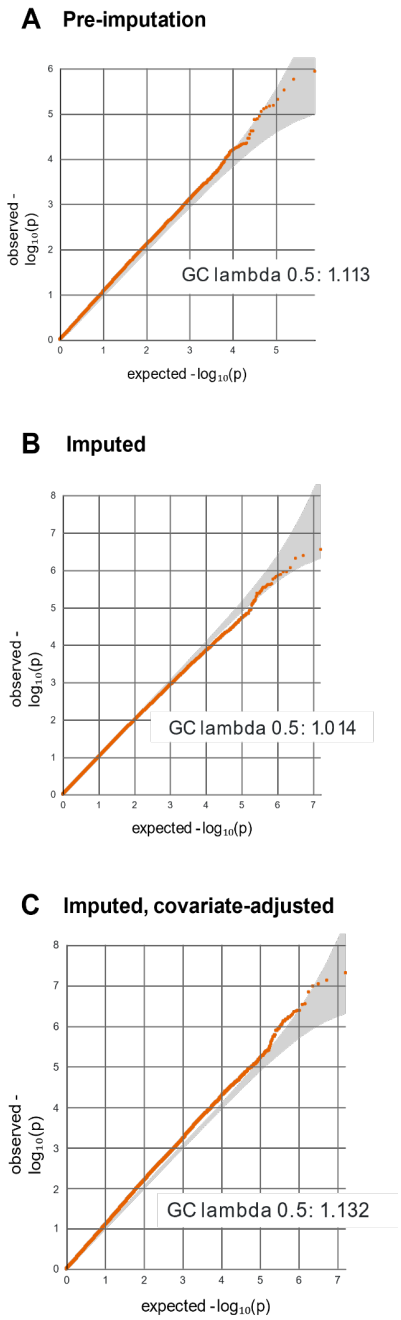


Figure 6.3 describes the steps taken to impute the post-QC dataset (1,169 individuals and 418,664 SNPs) to the HRC reference panel. Out of 418,664 SNPs, 418,276 were successfully matched to the HRC. 21,671,225 SNPs were imputed, and post-imputation QC, 8,271,956 SNPs were included in my analysis. Although the TOPMed (Trans-Omics for Precision Medicine) reference panel has been shown to provide more accurate imputation among non-European populations in comparison to HRC (**Supplementary Figure 6.4**), that improvement was minimal for West Asian and Central/South Asian populations.

6.5.1.2 Genome-wide analysis: identification of T2D risk loci in Arab-Emiratis

My primary analysis included 991 individuals; 412 controls with NGT and 579 cases with T2D (178 individuals with prediabetes were excluded from primary analysis). After the adjustment for age, sex, and BMI, 66 SNPs were found to be significantly associated with T2D ($p < 5.0x10^{-6}$) (**Figure 6.6**). After conditional analysis, 48 SNPs were found to be independent hits representing 48 loci significantly associated with T2D (**Table 6.2, Figure 6.7**). Regional plots of the 48 loci are shown at the end of the supplementary file (**Supplementary Figures 6.13-6.60**). **Figures 6.8-6.13** display regional plots for the 6 loci that were either previously reported to be associated with T2D risk or involved in related pathways (more in the discussion section of this chapter).

Figure 6.6 Manhattan plots for the primary Arab-Emirati genome-wide association analysis. GWAS was performed on 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance (NGT) and 578 cases with type 2 diabetes (T2D). (A) 418,664 SNPs pre-imputation, (B) 8,271,956 post-imputation SNPs, (C) 8,271,956 post-imputation SNPs, after adjustment for age, sex, and BMI.

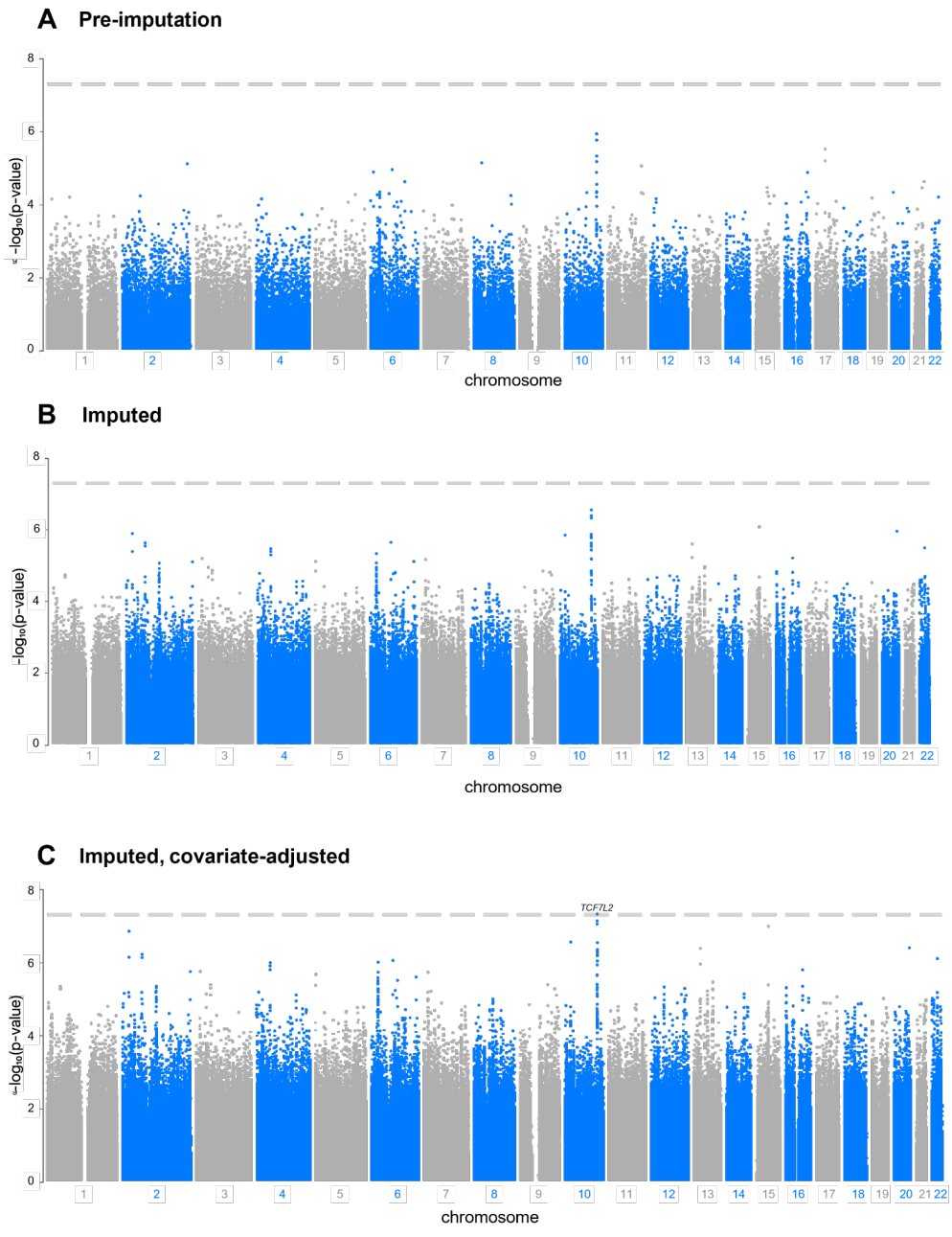
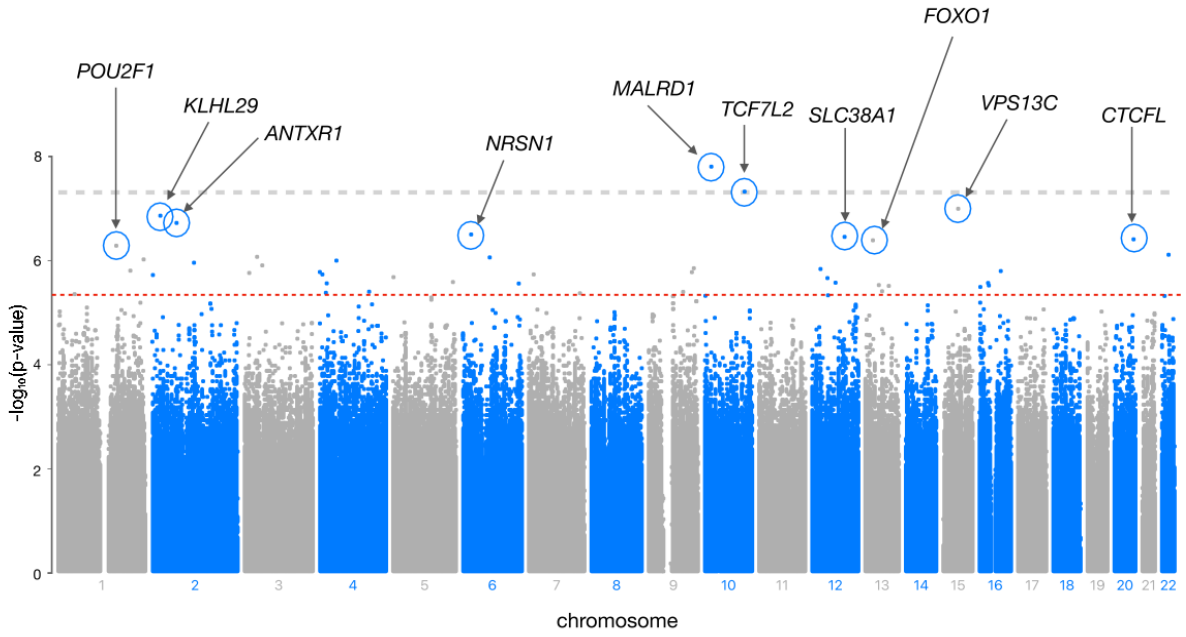


Figure 6.7 Manhattan plot for the primary Arab-Emirati genome-wide association analysis after performing conditional analysis. GWAS was performed on 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance (NGT) and 578 cases with type 2 diabetes (T2D). Genome- wide significance was set at p-value < $5x10^{-6}$, as depicted by the dotted red line. After covariate- adjustment (age, sex, and BMI) and conditional analysis, 48 SNPs were found to be independently associated with T2D. The top 10 SNPs with the most stringent P-values are marked in the figure.



One out of the 48 loci, rs7903146 in *TCF7L2*, a well described T2D risk variant, showed the second highest degree of statistical stringency in my analysis ($p < 5.0x10^{-8}$), after rs10827154, an intron variant in *MALRD1* ($p = 1.6x10^{-8}$), which was not previously reported to be related to T2D. The majority of reported loci (31 out of the 48) were either intronic, 3-Prime UTR, or non-coding transcript variants (**Table 6.2**). These included lead SNP rs6674762, an intronic variant in *KIF26B* (p = 9.7x10-7), and lead SNP rs13016276 in *ANTXR1* (p = 1.93x10-7). The rest of the lead SNPs (n = 17) were within +/-1 MB of potential candidate genes. All identified associations represent common genetic variants with modest effect size (OR for T2D ≤ 1.3), (**Table 6.2**).

6.5.1.3 Association of identified T2D risk loci with glucose and related metabolic traits

After correcting for multiple testing, rs11726869 in *ABLIM2* and rs10903783 in *PFKP* were significantly associated with increased fasting plasma glucose and HbA1c levels ($p < 0.001$). Additionally, rs13333067 in *NAA60* was only associated with decreased fasting plasma glucose levels. Whereas rs6456589 near *VPS13C* and rs10827154 in *MALRD1* were both associated with decreased HbA1c levels only. No significant associations were found with fasting insulin levels, HOMA-IR, or HOMA-B. The lead SNP rs13333067 in *NAA60,* along with rs13016276 in *ANTXR1*, were also significantly associated with increased waist-to-hip ratio (WHR). Whereas rs117089947 in *CNTNAP2* was significantly associated with decreased WHR. One lead SNP, rs13016276 in *ANTXR1* was significantly associated with increased BMI and increased triglyceride levels. No associations with other lipid measures (LDL, HDL, or TC) were observed. All associations, however, became insignificant once adjusted for T2D status (**Supplementary Tables 6.3-6.5**).

6.5.1.4 Testing the association of newly identified T2D lead SNPs with prediabetes

In efforts to identify whether genetic risk is shared between individuals with T2D and those with prediabetes, I firstly carried out a GWA analysis in 590 individuals (412 with controls with NGT and 178 cases with prediabetes). None of the loci that reached genome-wide significance ($p < 5.0x10^{-6}$) in the primary Arab-Emirati GWAS, including rs7903146 in *TCF7L2*, were associated with prediabetes. Secondly, combined analysis including 412 controls with NGT and 757 cases with prediabetes or T2D diluted the association results, with higher P-values overall (data not shown). However, regression analysis including the 48 lead SNPs identified in the primary Arab-Emirati GWAS, I found 2 SNPs, rs13333067 near *NA660* and rs75875192 near *INSIG2*, to be significantly associated with prediabetes ($p < 0.001$) (**Table 6.3**).

Figure 6.8 Regional plot showing the lead SNP identified in the primary Arab-Emirati GWAS in *TEC* gene. Association of SNP with T2D demonstrated is after the adjustment for age, sex and BMI. SNPs are plotted by position on chromosome against association ($-\log_{10}$ p-value) with type 2 diabetes (T2D) from the primary Arab-Emirati GWAS performed in 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance and 578 cases with T2D. P-values demonstrated represent the ones obtained prior to conditional analysis. The lead SNP in Arab-Emiratis is denoted with a purple diamond. The colour coding of all other SNPs indicates LD with the lead SNP (calculated r2 values for the Arab-Emirati sample): red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; cyan, 0.2 ≤ r2 < 0.4; blue r2 < 0.2; grey, r2 unknown. Gene's position is shown below the plot (using data from the UCSC genome browser. genome.ucsc.edu).

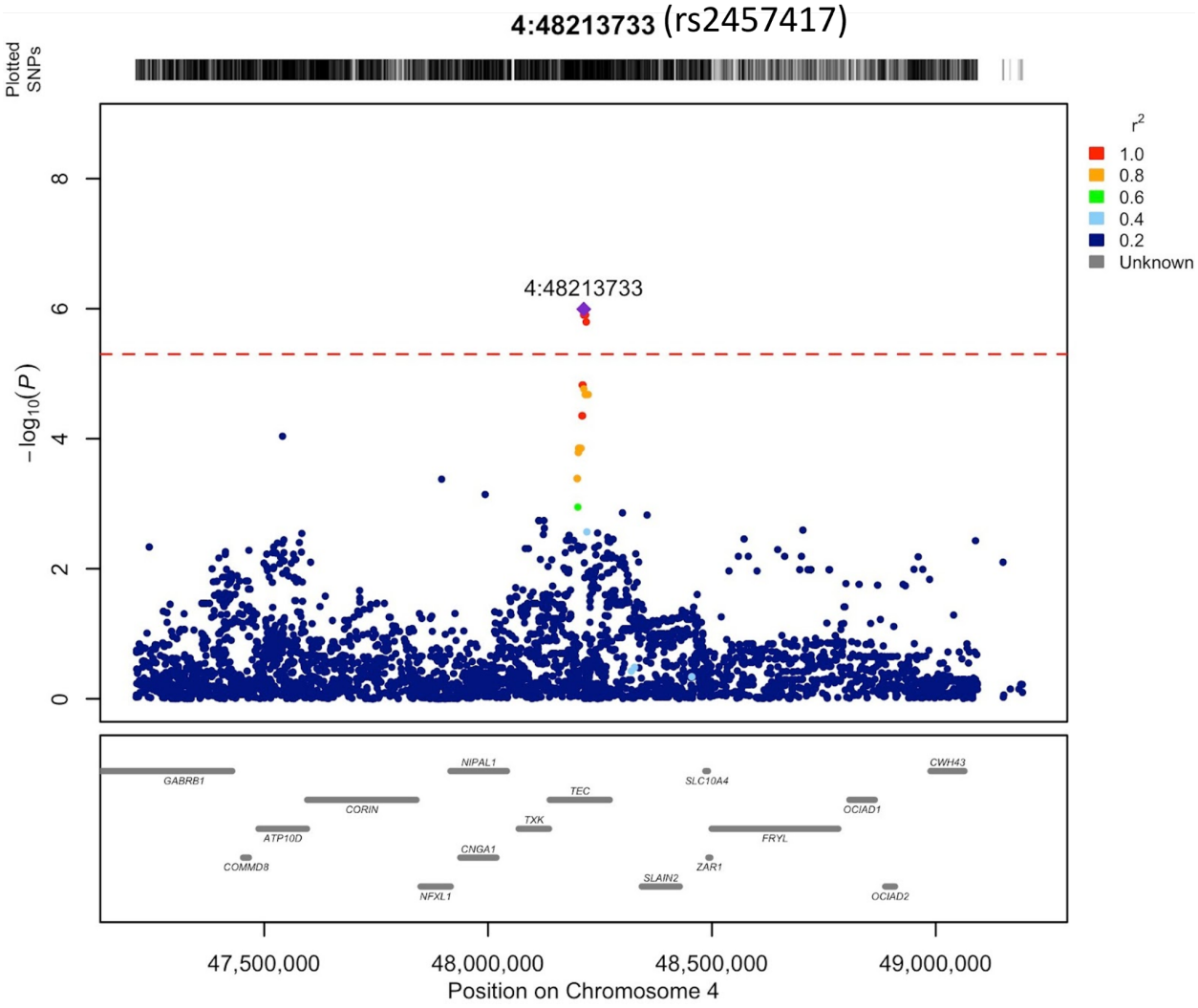


Figure 6.9 Regional plot showing the lead SNP identified in the primary Arab-Emirati GWAS in *IGF2R* gene. Association of SNP with T2D demonstrated is after the adjustment for age, sex and BMI. SNPs are plotted by position on chromosome against association (-$\log_{10}$ p-value) with type 2 diabetes (T2D) from the primary Arab-Emirati GWAS performed in 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance and 578 cases with T2D. P-values demonstrated represent the ones obtained prior to conditional analysis. The lead SNP in Arab-Emiratis is denoted with a purple diamond. The colour coding of all other SNPs indicates LD with the lead SNP (calculated r2 values for the Arab-Emirati sample): red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; cyan, 0.2 ≤ r2 < 0.4; blue r2 < 0.2; grey, r2 unknown. Gene's position is shown below the plot (using data from the UCSC genome browser. genome.ucsc.edu).

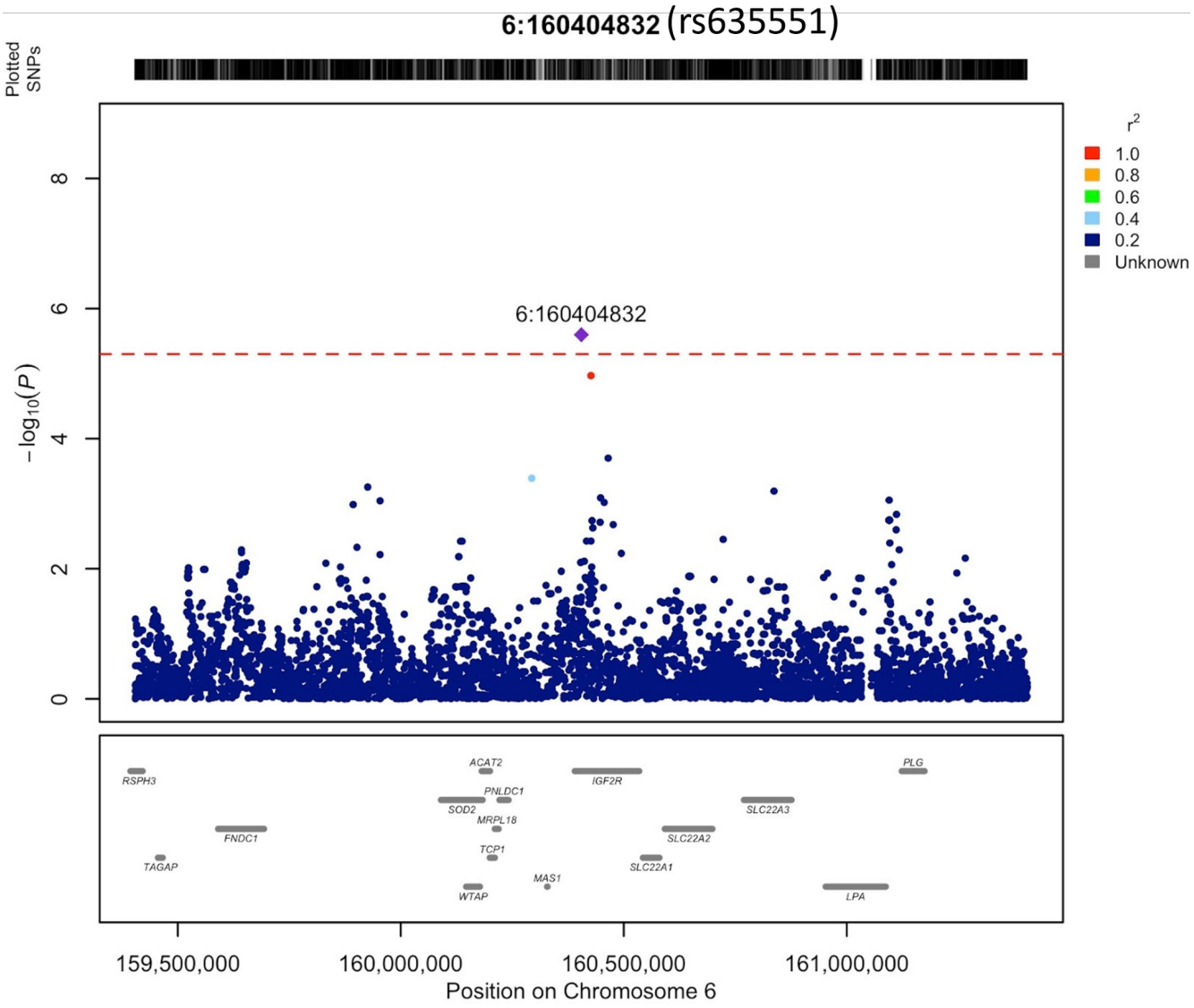


Figure 6.10 Regional plot showing the lead SNP identified in the primary Arab-Emirati GWAS in *CNTNAP2* gene. Association of SNP with T2D demonstrated is after the adjustment for age, sex and BMI. SNPs are plotted by position on chromosome against association ($-\log_{10}$ p-value) with type 2 diabetes (T2D) from the primary Arab-Emirati GWAS performed in 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance and 578 cases with T2D. P-values demonstrated represent the ones obtained prior to conditional analysis. The lead SNP in Arab-Emiratis is denoted with a purple diamond. The colour coding of all other SNPs indicates LD with the lead SNP (calculated r2 values for the Arab-Emirati sample): red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; cyan, 0.2 ≤ r2 < 0.4; blue r2 < 0.2; grey, r2 unknown. Gene's position is shown below the plot (using data from the UCSC genome browser. genome.ucsc.edu).

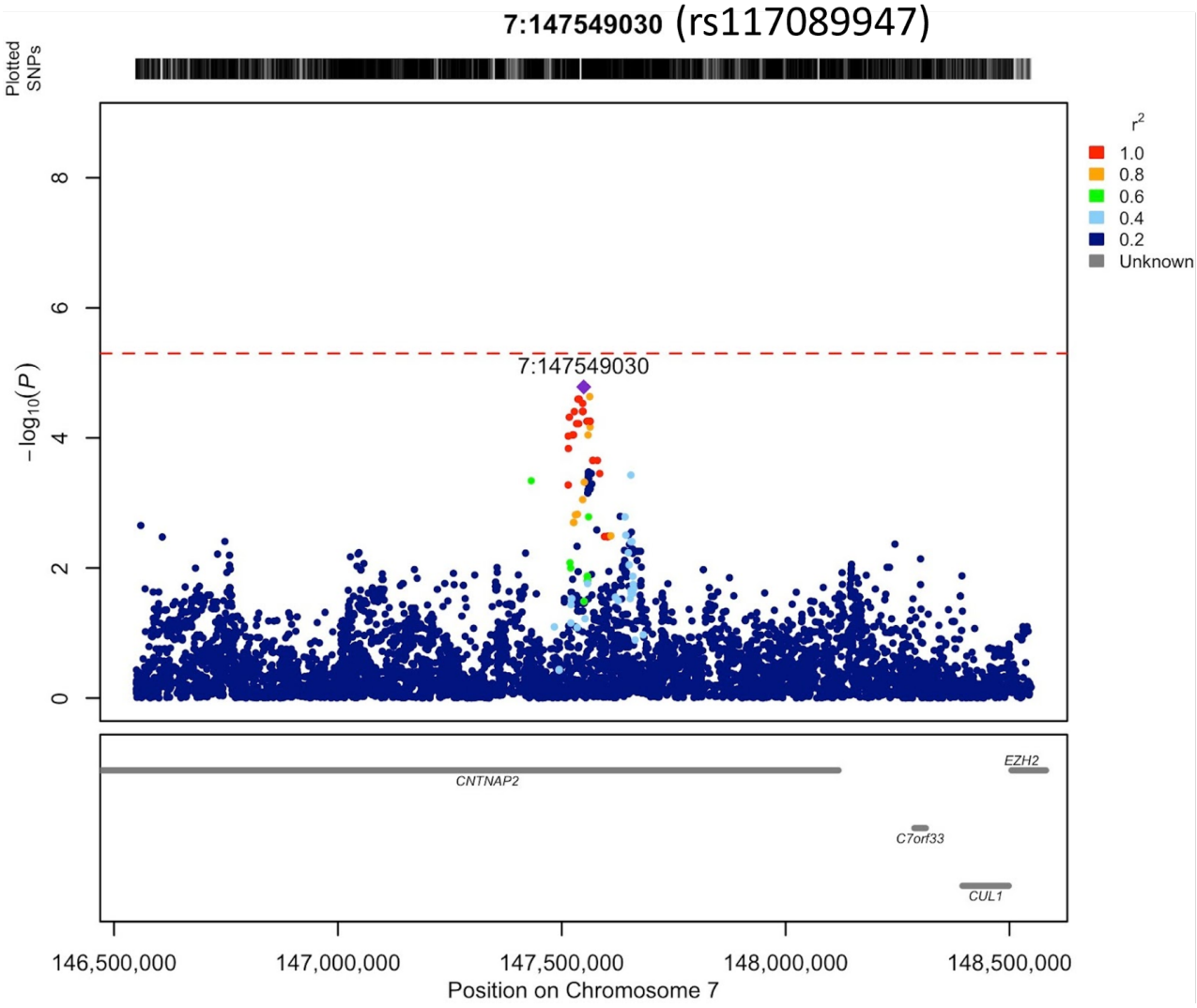


Figure 6.11 Regional plot showing the lead SNP identified in the primary Arab-Emirati GWAS near *DGKB* gene. Association of SNP with T2D demonstrated is after the adjustment for age, sex and BMI. SNPs are plotted by position on chromosome against association (-$\log_{10}$ p-value) with type 2 diabetes (T2D) from the primary Arab-Emirati GWAS performed in 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance and 578 cases with T2D. P-values demonstrated represent the ones obtained prior to conditional analysis. The lead SNP in Arab-Emiratis is denoted with a purple diamond. The colour coding of all other SNPs indicates LD with the lead SNP (calculated r2 values for the Arab-Emirati sample): red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; cyan, 0.2 ≤ r2 < 0.4; blue r2 < 0.2; grey, r2 unknown. Gene's position is shown below the plot (using data from the UCSC genome browser. genome.ucsc.edu).

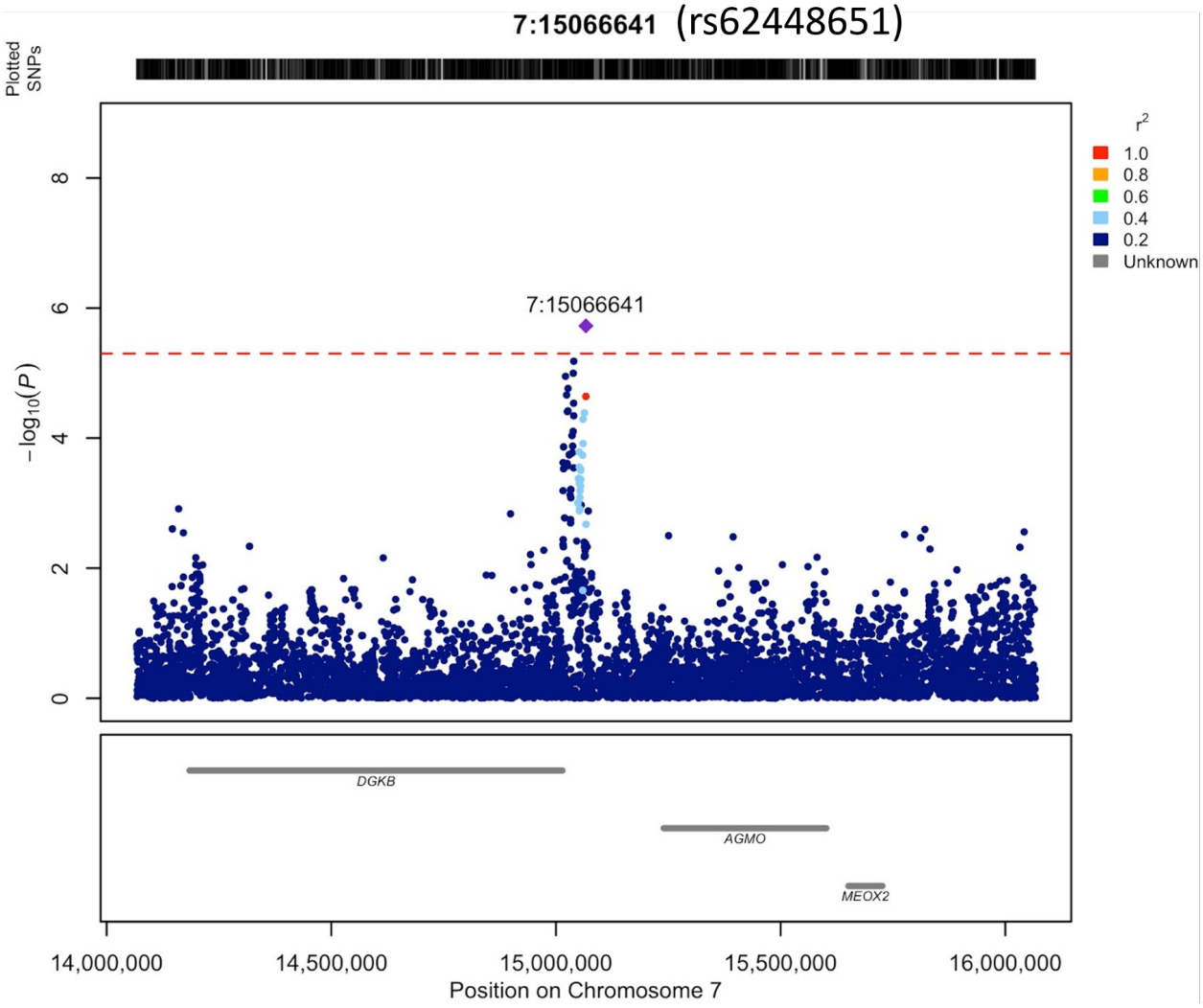


Figure 6.12 Regional plot showing the lead SNP identified in the primary Arab-Emirati GWAS in *MALRD1* gene. Association of SNP with T2D demonstrated is after the adjustment for age, sex and BMI. SNPs are plotted by position on chromosome against association (-$\log_{10}$ p-value) with type 2 diabetes (T2D) from the primary Arab-Emirati GWAS performed in 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance and 578 cases with T2D. P-values demonstrated represent the ones obtained prior to conditional analysis. The lead SNP in Arab-Emiratis is denoted with a purple diamond. The colour coding of all other SNPs indicates LD with the lead SNP (calculated r2 values for the Arab-Emirati sample): red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; cyan, 0.2 ≤ r2 < 0.4; blue r2 < 0.2; grey, r2 unknown. Gene's position is shown below the plot (using data from the UCSC genome browser. genome.ucsc.edu).

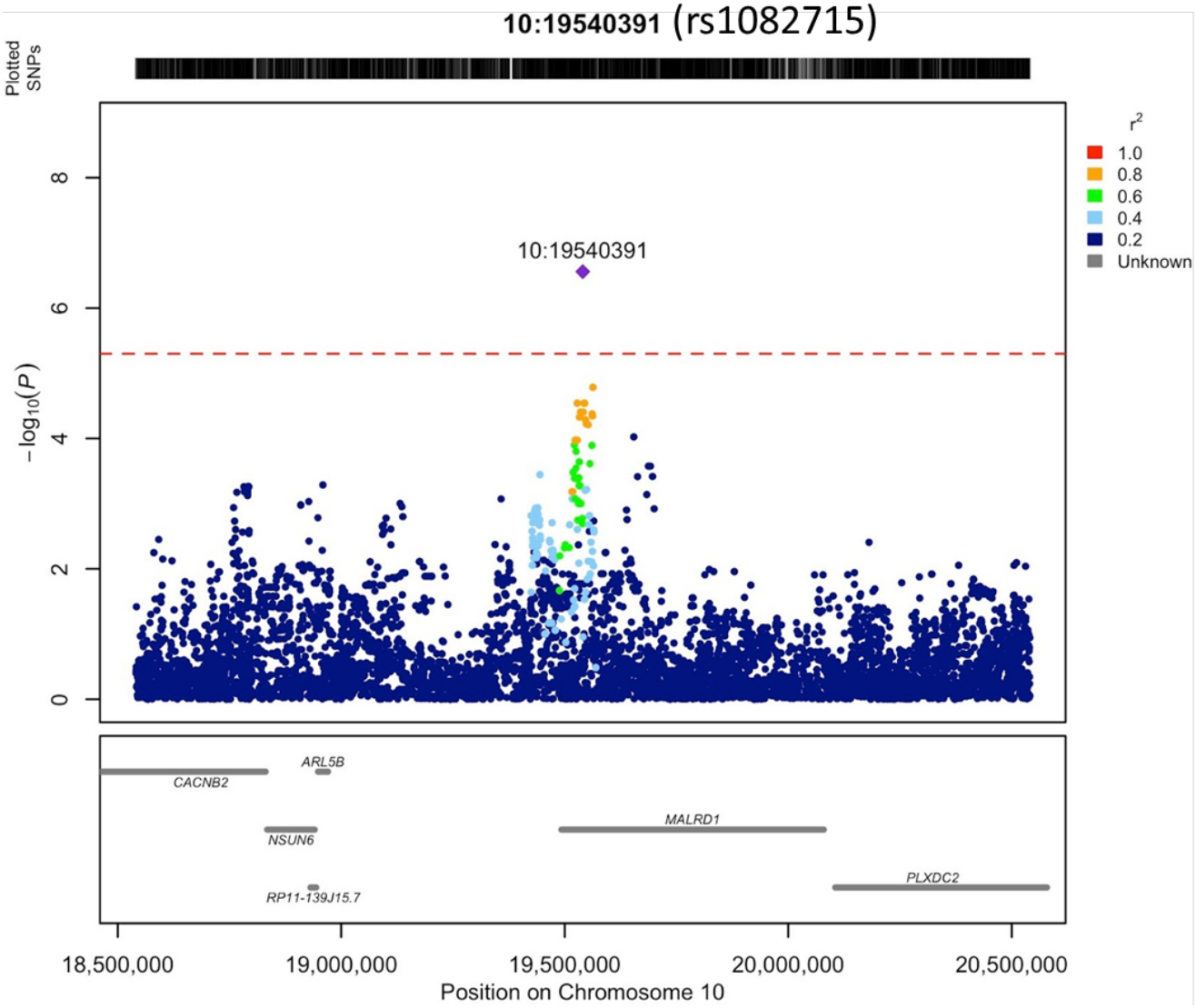
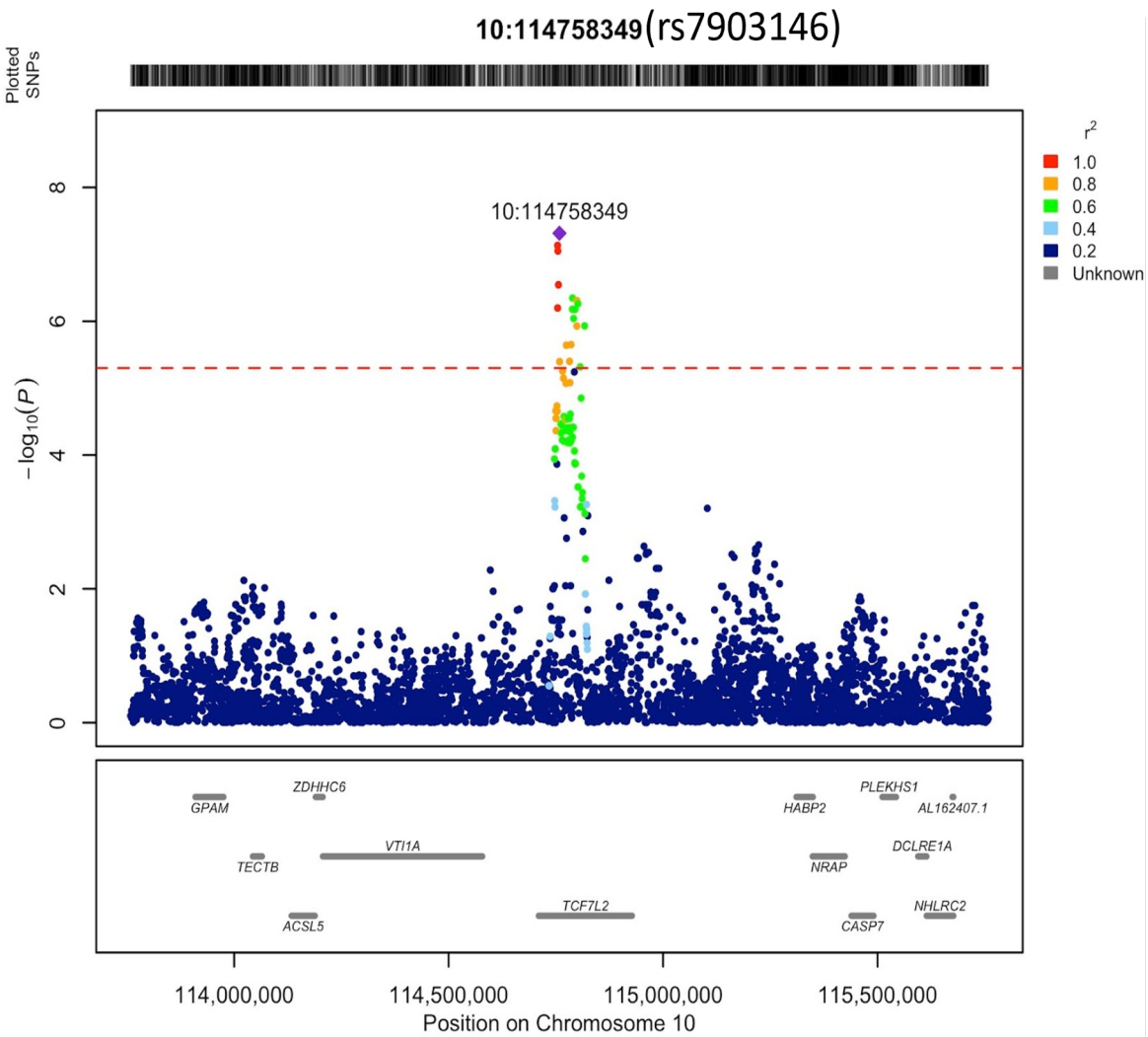


Figure 6.13 Regional plot showing the lead SNP identified in the primary Arab-Emirati GWAS in *TCF7L2* gene. Association of SNP with T2D demonstrated is after the adjustment for age, sex and BMI. SNPs are plotted by position on chromosome against association ($-\log_{10}$ p-value) with type 2 diabetes (T2D) from the primary Arab-Emirati GWAS performed in 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance and 578 cases with T2D. P-values demonstrated represent the ones obtained prior to conditional analysis. The lead SNP in Arab-Emiratis is denoted with a purple diamond. The colour coding of all other SNPs indicates LD with the lead SNP (calculated r2 values for the Arab-Emirati sample): red, r2 ≥ 0.8; gold, 0.6 ≤ r2 < 0.8; cyan, 0.2 ≤ r2 < 0.4; blue r2 < 0.2; grey, r2 unknown. Gene's position is shown below the plot (using data from the UCSC genome browser. genome.ucsc.edu).



6.5.1.5 Arab-Emirati-specific PRS

For training purposes, an Arab-Emirati-specific PRS for T2D was derived based on the 48 loci identified in the primary Arab-Emirati GWAS. Individuals' risk score was found to be significantly associated with increased risk of T2D and prediabetes with ORs of 15.3 (95%CI 13.7-17.1) and 3.9 (3.6-4.3), respectively (**Figure 6.14**). In addition, the risk score was also significantly associated with increased HOMA-IR, fasting glucose, HbA1c, and triglycerides, and decreased HOMA-B and HDL (**Figure 6.14**). In a sub-analysis including only individuals with prediabetes (n = 178), risk scores were not found to be significantly associated with worsened glycaemic or lipid measures (**Supplementary Table 6.6**).

6.5.2 Cross-ancestry comparisons

I evaluated the heterogeneity of allelic effects (i.e., discordance in the direction and/or magnitude of estimated ORs) between ancestry groups at; (1) loci showing genome-wide significance in Europeans and/or South Asian-based summary statistics ($p < 5x10^{-8}$) and a less stringent P-value in the primary Arab-Emirati GWAS ($p < 10^{-4}$), (2) previously reported lead SNPs in Europeans[148] and South Asians[345], (3) potential novel loci identified in the primary Arab-Emirati GWAS. In addition, I further characterised the identified loci in the primary Emirati GWAS by searching fasting glucose and fasting insulin meta-analyses summary statistics (MAGIC) for significant hits ($p < 5x10^{-8}$) within these loci (+/- 1 MB) (**Supplementary Table 6.7**).

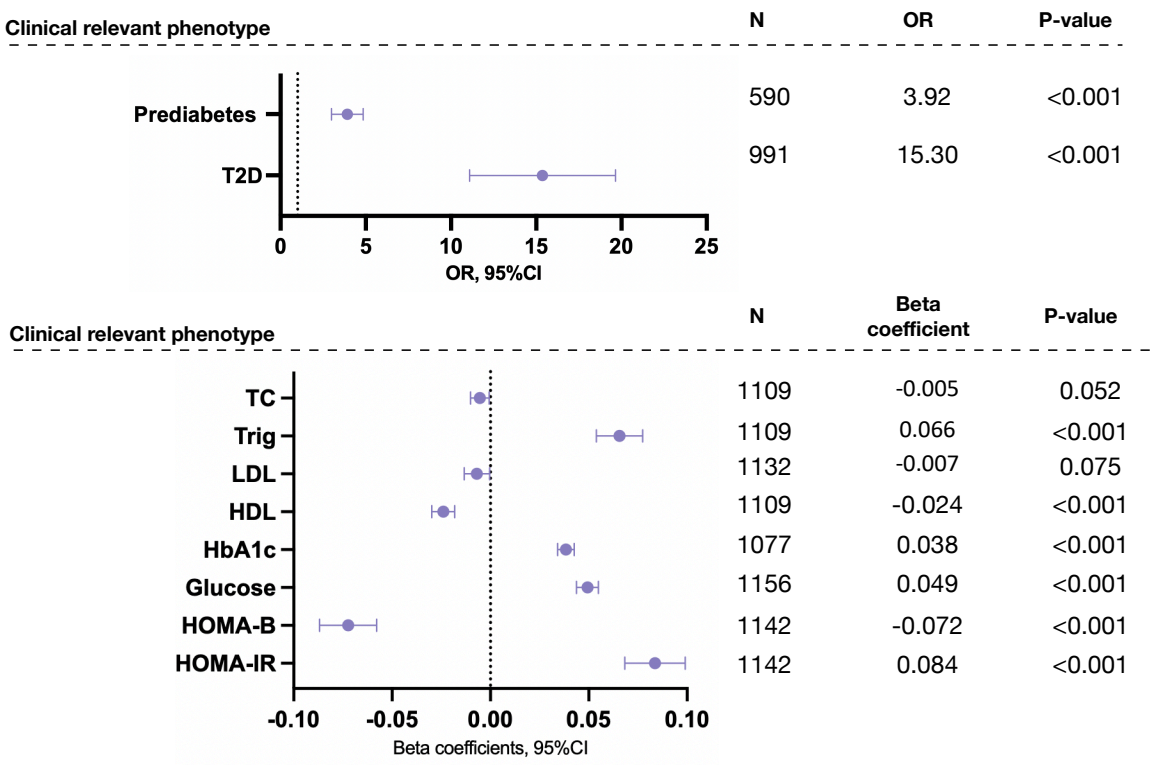
6.5.2.1 Cross-ancestry shared loci for T2D or glycaemic traits

The lead SNP rs7903146 in *TCF7L2*, had a lower effect size in Arab-Emiratis (OR = 1.12) than in either Europeans (OR = 1.36) or South Asians (OR = 1.30) (**Table 6.4** and **6.5**). With a p-value threshold of 5x10-[8], this lead SNP was the only one shared between Arab-Emiratis and the other ethnic groups. With a less

stringent P-value threshold ($p < 10^{-4}$), in addition to *TCF7L2*, SNPs in or near *DGKB*, *HNF1B*, *EYA2*, *PPARG*, and *CDKAL1* were found to be shared across all three ethnic groups (**Tables 6.6 and 6.7**). All SNPs showed concordance in direction of effect across ethnic groups. Overall, SNPs with a predisposing effect, had higher risk allele frequencies (RAF) among Arab-Emiratis than in Europeans and South Asians, whereas SNPs with a protective effect had lower RAF among Arab-Emiratis compared to the other ethnic groups (**Tables 6.6 and 6.7**).

Several SNPs near *DGKB* were previously reported in Europeans and South Asians (**Tables 6.6 and 6.7**). The lead SNP identified in South Asians (rs7786873) is moderately correlated with the lead SNP identified in the primary Arab-Emirati GWAS (rs62448651), but that is not the case for the European lead SNP (rs2191348) (**Supplementary Figure 6.5**). Like *DGKB*, several SNPs within/near *IGF2R*, *SOCS2* and *VPS13C* were also previously reported to be significantly associated with T2D in Europeans, also with different lead SNPs from those identified in Arab-Emiratis (**Supplementary Table 6.4**). In addition, SNPs within/near *KIF26B*, *RAD54L2*, *VPS13C*, and *NAA60* were also previously reported to be significantly associated with fasting glucose and/or fasting insulin in Europeans (**Supplementary Table 6.7**), but these SNPs were different from the ones identified to be significantly associated with T2D in Arab-Emiratis in this study.

Figure 6.14 Forest plots showing estimates from regression analyses testing the association of Arab-Emirati-specific polygenic risk scores against relevant phenotypes in Arab-Emiratis. Arab-Emirati-specific polygenic risk score (PRS) was derived based on the 48 lead SNPs identified in the primary Arab-Emirati GWAS to be associated with T2D (after adjustment for age, sex and BMI). GWAS was performed on 991 Arab-Emirati men and women; 412 controls with normal glucose tolerance (NGT) and 578 cases with type 2 diabetes (T2D). Estimates for type 2 diabetes (T2D) and prediabetes are shown as odds ratios (OR) and confidence intervals (error bars) for every 1 SD increase in PRS. Linear regression β coefficient estimates along with confidence intervals are shown for homeostatic model assessment-estimated insulin resistance (HOMA-IR) and β-cell function (HOMA-B), high-density lipoprotein [HDL (mmol/L)], low-density lipoprotein [LDL (mmol/L)], triglycerides [Trig (mmol/L)], and total cholesterol [TC (mmol/L)].



Meta-analysis of the primary Arab-Emirati GWAS data with South Asian-based data highlighted

additional SNPs (that did not reach genome-wide significance in the South Asian only results) in or near previously reported genes related to T2D (**Table 6.8**). The majority of these SNPs were moderately or strongly correlated (r2 = 0.4-0.8) with previously reported lead SNPs (**Supplementary Figures 6.9-6.12**). On the other hand, some SNPs, including rs555875344, rs532518450 and rs59306735, in *UBE2E2* (**Supplementary Figure 6.6**), *IGF2BP2* (**Supplementary Figure 6.7**), and *PPARG* (**Supplementary Figure 6.8**), respectively, were not in LD (r2 < 0.2) with previously reported lead SNP.

Supplementary tables 6.8-6.10 summarise the lead SNPs and the genes they map to, reported previously/recently in Europeans and South Asian populations, but did not replicate in Arab-Emiratis (except for *TCF7L2*, *HNF1B* and *EYA2* as highlighted above*)*. Despite that, most of those loci still showed concordance in direction of effect across ethnic groups.

6.5.2.2 Cross-ancestry heterogeneity of effect of the T2D risk variants identified in Arab-Emiratis

Out of the 48 lead SNPs identified in the primary Arab-Emirati GWAS, 25 were studied in the European-based meta-analysis[148], and 45 in the South Asian-based meta-analysis[345].

Out of the 25 SNPs studied in Europeans, 13 (50%) showed concordance in direction of effect between Europeans and Arab-Emiratis, with stronger effects in Arab-Emiratis (**Table 6.4**). In South Asians, on the other hand, 27 out of the 45 lead SNPs (59%) showed concordance in direction of effect with lower effect sizes compared to Arab-Emiratis (**Table 6.5**). In the trans-ethnic meta-analysis of South Asian and Arab-Emirati data, all 48 lead SNPs showed more stringent P-values, including rs531674840 in *POU2F1* and rs62448651 near *DGKB* ($p < 5.0x10^{-6}$), in comparison to summary statistics from South Asian only meta-analysis (**Table 6.9**).

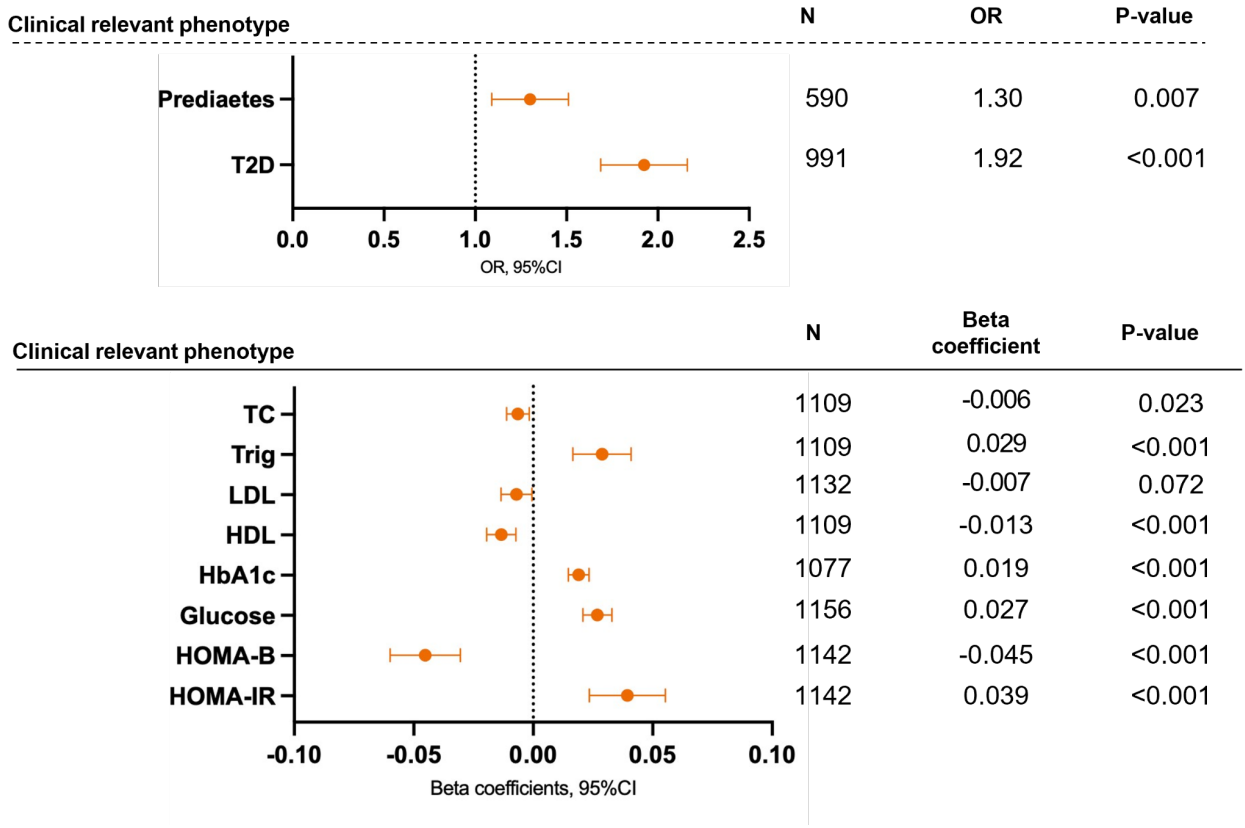
6.5.2.3 Evaluating the efficacy of European-specific PRS in predicting T2D risk in Arab-Emiratis

I generated a European-specific PRS based on 132 loci reported previously to be associated with T2D risk in Europeans[148]. The generated risk score was associated with an approximately 2-fold increased risk of T2D in Arab-Emiratis (OR: 1.9 (95%CI 1.8-2.0)) (**Figure 6.15**). It was also significantly associated with increased HOMA-IR, fasting plasma glucose, HbA1c, triglycerides, and total cholesterol levels, and decreased HOMA-B and HDL (**Figure 6.15**).

6.5.4 Sensitivity analyses

Given that the effects of some risk variants on T2D could be mediated by BMI, I investigated whether there are additional T2D-associated hits that were missed in my primary GWAS analyses, due to adjustment for BMI. However, in the unadjusted model (only adjusted for age and sex), I found 38 SNPs to be significantly associated with T2D in Arab-Emiratis ($p < 5x10^{-6}$), none of which were additional hits that were not discovered in my primary analyses (**Supplementary Table 6.11**).

Figure 6.15 Forest plots showing estimates from regression analyses testing the association of European-specific polygenic risk score against relevant phenotypes in Arab-Emiratis. European-specific polygenic risk score (PRS) was derived based on the 139 lead SNPs identified in the primary European GWAS meta-analyses. Estimates for type 2 diabetes (T2D) and prediabetes are shown as odds ratios (OR) and confidence intervals (error bars) for every 1 SD increase in PRS. Linear regression β coefficient estimates along with confidence intervals are shown for homeostatic model assessment-estimated insulin resistance (HOMA-IR) and β-cell function (HOMA-B), high-density lipoprotein [HDL (mmol/L)], low-density lipoprotein [LDL (mmol/L)], triglycerides [Trig (mmol/L)], and total cholesterol [TC (mmol/L)].



6.6 DISCUSSION

I carried out a genome-wide association analysis to identify known and novel loci underlying T2D amongst Arab-Emiratis, a population with an increased susceptibility to T2D. Firstly, my results support the role of some known common genetic variants in the pathogenesis of T2D amongst Arab-Emiratis. Those risk loci had modest but homogeneous effects across Arab-Emirati, Europeans and South Asians, and cannot account for the increased risk of T2D in the Arab-Emirati population. Secondly, I have identified, for the first time, 40 potentially novel loci, with rs10827154 mapping to *MALRD1* showing the strongest association. These SNPs provide some insight into the genetic mechanisms underlying T2D in this high-risk population and demonstrate the potential for new discovery from genetic association studies in populations of non-European ancestry.

Known, previously reported loci for T2D and/or glycaemic traits, including *TCF7L2, HNF1B*, *DGKB*, *PPARG*, *CDKAL1*, *IGF2R*, *KIF26B*, *RAD54L2*, *SOCS2*, *VPS13C*, and *NAA60* were replicated in my study, highlighting their role in T2D susceptibility among Arab-Emiratis. Shared variants within these loci had modest and homogeneous effects across Europeans, South Asians, and Arab-Emiratis, with smaller effect sizes in Arab-Emiratis. This indicates that these variants alone are unlikely to explain the increased T2D risk in Arab-Emiratis. Possible reasons for the failure to replicate many of the associations of previously identified T2D risk loci in this population may relate to limitations of sample size and power. Alternatively, differential exposure to environmental risk factors in different populations may alter the impact of susceptibility loci on T2D risk, i.e., via gene-environmental interactions[353]. Whether our observations of lower effects of known T2D risk loci reflect differences in study design, true heterogeneity of effect, or gene-environmental interactions remains to be determined.

TCF7L2 (transcription factor 7-like 2) is the most potent locus for T2D and is the first to have been robustly reported by genomic linkage studies[154] and GWAS[354]. Its association with T2D is one of the strongest genetic discoveries of complex diseases, and it has been replicated in multiple ethnic groups[355]. TCF7L2 is a member of the T-cell factor/lymphoid enhancer binding factor family (TCF/LEF), a group of transcription factors binding to DNA through high-mobility group (HMG) domain. The *TCF7L2* gene forms a basic part of the Wnt signalling pathway, which has been related to various common diseases, including different models of cancers[355]. TCF7L2 and Wnt signalling has been recognised for its fundamental developmental and metabolic roles in adipose tissue, including promoting adipocyte hypertrophy and peripheral and hepatic insulin resistance[355].

HNF1B encodes a liver-enriched homeodomain-containing transcription factor, which plays critical roles in the regulation of glucose transport and metabolism, glucose stimulated insulin secretion, and also in the embryonic development and later function of the pancreas, kidneys, and liver[356] . Mutations in *HNF1B* are known to cause MODY type 5, characterised by defective pancreatic β-cell function and impaired insulin secretion, and other distinct clinical features including renal disease[356]. Similarly, DGKB (diacylglycerol kinase β) also plays a role in insulin secretion from pancreatic β-cells[357]. *PPARG* encodes the peroxisome proliferator-activator receptor γ is a nuclear hormone receptor preferentially expressed in adipose tissue and has been repeatedly associated with a decreased risk of T2D[358]. Lastly, The *CDKAL1* gene encodes cyclin-dependent kinase 5 regulatory subunit-associated protein 1-like 1 that also plays a critical role in glucose-dependent regulation of insulin secretion, and variants within this gene have been strongly associated with an increased risk of T2D and obesity in dozens of studies and in diverse populations[359].

In addition to known T2D susceptibility loci, I have identified 40 novel loci to potentially be associated with risk of T2D in Arab-Emiratis. At 10p12.31, rs10827154 in *MALRD1*, a gene encoding a conserved protein that features multiple MAM (meprin-A5-protein tyrosine phosphatase mu) and LDLR A2 (low density lipoprotein receptor A2) domains, showed the strongest association to T2D in my study (stronger than rs7903146 in *TCF7L2*). MALRD1 is a newly identified protein that modulates levels of fibroblast growth factor 19 (FGF19), a protein involved in the regulation of hepatic metabolic processes. Alterations in the homeostasis of these processes can result in inflammation and development of diseases, such as cholestatic liver diseases, dyslipidaemia, and diabetes[360].

The new loci also included intronic variants mapping to *CNTNAP2*, and *TEC* genes. Previous reports identify multiple variants within these genes that are associated with phenotypes related to obesity and T2D.

CNTNAP2 gene encodes the contactin-associated protein-like 2, a member of the neurexin family that functions in the nervous system as cell adhesion molecules. Variants within this gene were associated with phenotypes including BMI[361], HDL and LDL levels[362], visceral adipose tissue[363], and waist circumference[364]. *TEC* encodes protein tyrosine kinase, which is involved in intracellular signalling mechanisms of cytokine receptors, lymphocyte surface antigens, and integrin molecules. Variants within this gene were previously associated with systolic blood pressure[365] and inflammatory bowel diseases[366].

My results showed that European-specific PRS can be valuable in predicting risk of T2D in Arab-Emiratis. Whether the Arab-Emirati-specific PRS can improve the efficacy and the discriminatory value in identifying Arab-Emiratis at increased T2D risk further will need further testing in a separate replication cohort from the population.

Assessing the impact of lead SNPs at T2D susceptibility loci on related quantitative phenotypes, such as fasting glycaemic traits, anthropometric measures and lipid profiles, can provide insight into relevant biology underlying the disease. In line with previous studies[363, 364], variants in *CNTNAP2* were associated with waist and WHR in Arab-Emiratis. In addition to *CNTNAP2,* lead SNPs in *ANTXR1* and *CDH8* were associated with increased WHR, but no biology relating these genes (and/or variants in these genes) to central adiposity has not previously been reported, so that their role is unclear. The lead SNP in *ABLIM2* was associated with increased fasting glucose levels, HbA1c, and BMI. *ABLIM2* encodes actin binding LIM protein family, member 2, predicted to enable actin filament binding activity, but no relationship of this gene or its variants with glucose metabolism or T2D has been described to date. Variants within *NAA60* and *VPS13C* were previously reported to be associated with changes in fasting blood glucose [367], and in line with these findings, I found rs13333067 in *NAA60* and rs6456589 near *VPS13C* to be associated with decreased fasting plasma glucose levels in my study. It must be noted that associations with quantitative traits observed here were dependent on T2D. Therefore, future investigations on the relationship of identified novel loci with adiposity measures, glycaemic or lipid traits among individuals with no diabetes would give valuable insight on the mechanisms underlying the measured genetic susceptibility.

My trans-ethnic meta-analysis of Arab-Emirati and South Asian data indicated that some of the novel common SNPs identified here, could be amenable to discovery through trans-ethnic meta-analysis in larger samples. That was the case for loci in or near *POU2F1* (POU Class 2 Homeobox 1), *RNF150* (Ring Finger Protein 150), and *DYRK2* (Dual Specificity Tyrosine Phosphorylation Regulated Kinase 2) in particular, characterised with overall stronger association with T2D in the trans-ethnic meta-analysis compared to the South Asian-specific results. Furthermore, despite the small sample size of Arab-Emirati sample, we identified additional SNPs in or near *UBE2E2* (Ubiquitin Conjugating Enzyme E2 E2), *IGF2BP2*

(Insulin Like Growth Factor 2 MRNA Binding Protein 2), and *PPARG* (Peroxisome Proliferator Activated Receptor Gamma) that are independent of (i.e., not in LD with) previously reported SNPs. This highlights the value of trans-ethnic meta-analysis in the identification of susceptibility variants that might be at more appreciable allele frequencies in certain ethnic groups, and/or can only be discovered in studies with larger sample sizes.

Studies investigating the genetic basis of T2D in the Middle East, and the UAE in particular, are limited, and summary statistics for published studies are not publicly-available. Thus, replication and confirmation of our potential novel loci in Middle Eastern populations was not possible. Previously, Al Safar and colleagues[268] studied an extended family of 66 T2D cases and 112 controls and identified novel loci at *GABRA4* and *PRKD1* with T2D, but these were not replicated in my study of Arab-Emiratis. In another study[269] of 464 Arab-Emirati T2D cases and 415 controls free from diabetes, reported additional novel loci at *KCNK3* and *RARB*, but these were also not replicated in my study.

Numerous epidemiological and clinical trials have shown that onset of T2D can be delayed and may be preventable through early lifestyle modification and/or pharmacological intervention, preventing the decline of β-cell function[368]. This highlights the importance of early detection of individuals at risk, including those with prediabetes. My results here indicate that the identified T2D risk loci are unlikely to play a role in the progression of prediabetes to T2D. The Arab-Emirati-specific PRS based on the lead SNPs identified in the primary Arab-Emirati GWAS was not associated with worsened glycaemic or lipid measures in people with prediabetes. Overall, however, Arab-Emirati- and European-specific PRS were both moderately associated with prediabetes (OR of 1.3). These results should be interpreted with caution, given the small sample of individuals with prediabetes (n = 178).

In conclusion, my findings do not show that previously reported T2D risk loci account for the increased risk of T2D amongst Arab-Emiratis, compared to Europeans and South Asians. Through the Arab-Emirati specific GWAS, I have identified 40 novel loci that can potentially provide insight into the genetic mechanisms underlying T2D in the population. Overall, my results reinforce the importance of extending the investigation of T2D genetic susceptibility to populations of non-European ancestry. Although this study is underpowered and results could be highlighting false discoveries, the replication of several known T2D or glycaemic traits loci in this study, gives us some confidence on the novel observed associations. Nevertheless, testing the replicability of those novel loci in another pan-Arab cohort (e.g., The Emirati Genome Project), is highly required.

Table 6.1 Characteristics of the 1,169 Arab-Emirati men and women who were genotyped and passed quality control procedures. The 1,169 individuals are part of the EM2 sample population, which was collected in a case-control study design. They included 412 controls with normal glucose tolerance (NGT), 178 with prediabetes, and 579 cases with type 2 diabetes (T2D).

	NGT				**Prediabetes**					**T2D**				
N	**412**				**178**					**579**				
	Mean SD		**Median IQR**		**Mean SD**		**Median IQR**		**P**	**Mean SD**		**Median IQR**		**P**
Age (years)	37.14	5.83	36.82	33.3-40.6	39.29	6.22	38.63	35.2-42.7	<0.001	37.69	5.55	38.35	34.3-41.9	0.009
SBP (mmHg)	113.5	14.2	113	104.0-121.0	117.89	16.00	117	107.0-125.0	<0.001	122.3	14.00	121.00	113.0-132.0	<0.001
DBP (mmHg)	69.13	10.1	68	61.0-76.0	73.03	10.6	72.00	66.0-79.0	<0.001	74.96	9.99	75.00	68.0-82.0	<0.001
Adiposity														
BMI (Kg/m^2)	29.26	6.06	28.38	24.8-32.7	31.05	5.89	30.85	26.6-34.8	<0.001	32.83	6.31	32.03	28.5-36.1	<0.001
Weight (Kg)	79.13	17.9	77.6	66.2-89.6	85.06	17.6	83	73.8-96.2	<0.001	90.75	19	88.1	77.0-101.8	<0.001
Waist (cm)	91.33	13.6	92	81.0-101.0	98.24	14	97	89.0-106.0	<0.001	103.5	13.9	102	93.0-150.0	<0.001
Hip (cm)	105	12.5	104	96.0-112.0	108.56	12.4	107.5	100.0-115.0	<0.001	110.4	13.3	108	101.0-118.0	<0.001

WHR	0.87	0.08	0.87	0.81-0.94	0.91	0.08	0.91	0.85-0.97	<0.001	0.94	0.08	0.94	0.89-0.99	<0.001
VAT (L)	2.88	1.95	2.86	1.4-4.1	3.86	2.09	3.5	2.4-5.3	<0.001	4.67	2.14	4.42	3.0-5.9	<0.001
Glycaemic profile														
Glucose (mmol/L)	5.27	0.33	5.26	5.1-5.5	5.81	0.53	5.75	5.5-6.0	<0.001	8.11	2.93	7.3	6.2-9.0	<0.001
HbA1c (mmol/mol)	33.73	3.93	34.43	31.2-36.6	37.27	4.64	37.71	34.4-39.9	<0.001	54.06	17	49.73	42.1-61.8	<0.001
Insulin (pmol/L)	88.85	57.7	76.07	53.9-104.9	118.29	73.8	105.1	68.7-143.3	<0.001	133.8	92.2	111.6	77.2-165.3	<0.001
HOMA-IR	3.03	2.09	2.52	1.8-3.6	4.47	2.99	3.8	2.5-5.5	<0.001	7.2	6.47	5.4	3.5-8.6	<0.001
Lipid profile														
TC (mmol/L)	4.68	0.87	4.56	4.1-5.2	4.86	0.78	4.87	4.3-5.4	<0.001	4.48	1.12	4.38	3.7-5.1	<0.001
LDL (mmol/L)	3.09	0.81	3.01	2.6-3.6	3.30	0.75	3.34	2.7-3.9	0.005	2.97	0.99	2.87	2.2-3.5	<0.001
HDL (mmol/L)	1.42	0.36	1.39	1.2-1.6	1.32	0.33	1.26	1.1-1.5	<0.001	1.16	0.31	1.12	1.0-1.3	<0.001
Trig (mmol/L)	1.09	0.86	0.89	0.7-1.2	1.19	0.55	1.09	0.8-1.4	<0.001	1.77	1.63	1.41	1.1-2.0	<0.001

NGT; normal glucose tolerance, T2D; type 2 diabetes, BMI; body mass index, HbA1c; glycated haemoglobin, TC; total cholesterol, LDL; low-density lipoprotein, HDL; high-density lipoprotein, SBP; systolic blood pressure, DBP; diastolic blood pressure, HOMA-IR; homeostatic model assessment of insulin resistance. P: p

values for Wilcoxon rank-sum test performed comparing parameters in prediabetes and T2D with NGT.

Table 6.2 Lead SNPs associated with T2D at P < 5.0e-6 in the primary Arab-Emirati GWA analysis. The primary Arab-Emirati GWAS was performed on 991 Arab-Emirati men and women; 412 with normal glucose tolerance and 578 with type 2 diabetes. Data shown is after adjustment for covariates (age, sex, and BMI).

			Location									
SNP	**Chr**	**Position**	**Nearest gene**	**Description**	**A1**	**A2**	**RAF**	**B**	**SE**	**OR**	**P**	**N**
rs12119289	1	48008522	*FOXD2*	within 1 MB	A	G	0.327	-0.100	0.022	0.905	4.49E-06	991
rs6674762	1	245625216	*KIF26B*	Intron variant	A	G	0.010	-0.462	0.094	0.630	9.68E-07	991
rs17026195	1	207553808	*CD55*	within 1 MB	G	A	0.012	-0.441	0.092	0.643	1.59E-06	991
rs2861451	1	167312409	*POU2F1*	Intron variant	T	G	0.065	0.194	0.039	1.214	5.28E-07	991
rs4487114	2	23050710	*KLHL29*	within 1 MB	G	C	0.313	0.120	0.023	1.128	1.40E-07	991
rs13016276	2	69391521	*ANTXR1*	Intron variant	G	A	0.453	-0.105	0.020	0.900	1.93E-07	991
rs75875192	2	119217818	*LOC105373579 (INSIG2)*	within 1 MB	G	T	0.050	0.212	0.044	1.236	1.12E-06	991
rs60745626	2	1689209	*PXDN*	Intron variant	T	A	0.278	0.104	0.022	1.110	1.93E-06	991
rs150527739	3	14432140	*LOC105376959 (SLC6A6)*	Intron variant	A	G	0.023	-0.323	0.067	0.724	1.76E-06	991
rs144006677	3	51681325	*RAD54L2*	Intron variant	A	G	0.016	-0.375	0.077	0.688	1.26E-06	991

rs3934473	3	36828646	*DCLK3*	within 1 MB	A	C	0.045	-0.236	0.048	0.790	8.64E-07	991
rs2457417	4	48213733	*TEC*	Intron variant	A	G	0.149	0.145	0.030	1.155	1.02E-06	991
rs2129645	4	141792539	*RNF150*	Intron variant	A	T	0.187	-0.120	0.026	0.887	4.03E-06	991
rs139588925	4	18131933	*LCORL*	Non-coding transcript variant	G	A	0.026	-0.294	0.064	0.745	4.23E-06	991
rs511804	4	20545268	*SLIT2*	Intron variant	G	C	0.146	-0.129	0.027	0.879	2.79E-06	991
rs11726869	4	8026645	*ABLIM2*	Intron variant	G	T	0.221	-0.113	0.024	0.893	1.88E-06	991
rs6811804	4	774617	*LOC100129917 (CPLXL)*	Non-coding transcript variant	C	T	0.231	-0.110	0.023	0.895	1.69E-06	991
rs77547415	5	2560883	*IRX2*	within 1 MB	T	C	0.023	-0.334	0.070	0.716	2.12E-06	991
rs17676033	5	172743694	*STC2*	3-Prime UTR variant	T	C	0.060	0.195	0.042	1.215	2.63E-06	991
rs6919765	6	77261591	*LINC02540 (IMPG1)*	within 1 MB	G	G	0.245	-0.122	0.025	0.885	8.86E-07	991
rs6456589	6	24092339	*NRSN1*	within 1 MB	C	A	0.377	0.108	0.021	1.114	3.21E-07	991
rs635551	6	160404832	*IGF2R*	Intron variant	C	T	0.113	0.150	0.032	1.162	2.81E-06	991
rs62448651	7	15066641	*DGKB*	within 1 MB	A	C	0.441	0.105	0.022	1.111	1.88E-06	991
rs117089947	7	147549030	*CNTNAP2*	Intron variant	C	T	0.060	-0.184	0.040	0.832	4.32E-06	991
rs111777155	9	99820736	*CTSV*	within 1 MB	G	T	0.017	-0.376	0.082	0.686	4.06E-06	991

rs117077124	9	124942205	*MORN5*	Intron variant	A	G	0.047	0.233	0.049	1.262	1.70E-06	991
rs11792868	9	130480322	*TTC16*	Intron variant	C	A	0.092	-0.160	0.033	0.852	1.43E-06	991
rs7903146	10	114758349	*TCF7L2*	Intron variant	T	C	0.423	0.116	0.021	1.123	4.83E-08	991
rs10827154	10	19540391	*MALRD1*	Intron variant	A	G	0.216	-0.138	0.024	0.871	1.60E-08	991
rs10903783	10	2540623	*LINC02645 (PFKP)*	Intron variant	T	C	0.440	0.089	0.019	1.093	4.84E-06	991
rs17096794	12	46604091	*SLC38A1*	Intron variant	G	C	0.112	0.147	0.032	1.159	4.73E-06	991
rs7953156	12	44929687	*NELL2*	Intron variant	A	G	0.279	0.103	0.022	1.108	2.21E-06	991
rs11107089	12	93941505	*SOCS2*	within 1 MB	A	A	0.027	-0.316	0.062	0.729	3.57E-07	991
rs11047700	12	25061338	*BCAT1*	Intron variant	T	A	0.103	-0.159	0.033	0.853	1.48E-06	991
rs17104138	12	68005294	*DYRK2*	within 1 MB	G	G	0.066	-0.183	0.039	0.832	2.71E-06	991
rs2324604	13	40971758	*LINC00598 (FOXO1)*	Intron variant	C	T	0.363	0.107	0.021	1.113	4.16E-07	991
rs9594071	13	86512325	*SLITRK6*	within 1 MB	A	G	0.456	-0.093	0.020	0.911	3.12E-06	991
rs1031302	13	58038376	*LOC105370217 (PCDH17)*	Intron variant	A	A	0.230	0.108	0.023	1.114	3.01E-06	991
rs9540874	13	67290606	*PCDH9*	Intron variant	T	G	0.141	-0.135	0.029	0.874	3.95E-06	991
rs16944568	15	62044136	*VPS13C*	within 1 MB	G	T	0.163	-0.150	0.028	0.861	1.03E-07	991
rs142668827	15	55114632	*UNC13C*	within 1 MB	C	A	0.100	-0.147	0.033	0.863	9.71E-06	991

rs117027386	16	61969102	*CDH8*	Intron variant	A	C	0.021	-0.344	0.072	0.709	1.62E-06	991
rs13333067	16	3495383	*NAA60*	Intron variant	T	C	0.071	0.173	0.037	1.189	3.27E-06	991
rs55870929	16	25647503	*HS3ST4*	within 1 MB	T	C	0.058	-0.198	0.042	0.821	2.74E-06	991
rs727836	16	27832033	*GSG1L*	Intron variant	T	G	0.148	-0.129	0.028	0.879	3.06E-06	991
rs913548	20	56053525	*CTCFL*	within 1 MB	A	T	0.314	0.112	0.022	1.118	3.98E-07	991
rs77316551	22	35285864	*LINC02885 (ISX)*	Intron variant	T	C	0.045	-0.247	0.050	0.781	7.87E-07	991
rs5759835	22	23816142	*FAM230I*	Intron variant	T	C	0.077	-0.173	0.038	0.841	4.87E-06	991

RAF: risk allele frequency; A1: reference allele, A2: alternative allele.

Table 6.3 The association of the 48 lead SNPs identified in the primary Arab-Emirati GWAS with prediabetes under an additive genetic model. Associations nominally significant at P < 0.001 after adjusting for age, BMI, and sex are marked in green.

						Prediabetes	
N						**590**	
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**OR**	**P**
rs12119289	1	48008522	*FOXD2*	A	G	0.8327	1.69E-01
rs6674762	1	245625216	*KIF26B*	A	G	0.8091	6.53E-01
rs17026195	1	207553808	*CD55*	G	A	0.8153	6.76E-01
rs2861451	1	167312409	*POU2F1*	T	G	1.59	1.08E-01
rs4487114	2	23050710	*KLHL29*	G	C	1.238	1.46E-01
rs13016276	2	69391521	*ANTXR1*	G	A	0.8094	1.06E-01
rs75875192	2	119217818	*INSIG2*	G	T	2.811	8.49E-04
rs60745626	2	1689209	*PXDN*	T	A	1.504	4.33E-03
rs150527739	3	14432140	*SLC6A6*	A	G	0.524	1.16E-01
rs144006677	3	51681325	*RAD54L2*	A	G	0.6929	3.87E-01
rs3934473	3	36828646	*DCLK3*	A	C	0.6529	1.60E-01
rs2457417	4	48213733	*TEC*	A	G	1.176	4.25E-01
rs2129645	4	141792539	*RNF150*	A	T	0.6545	9.86E-03
rs139588925	4	18131933	*LCORL*	G	A	0.652	2.40E-01
rs511804	4	20545268	*SLIT2*	G	C	0.8064	2.27E-01
rs11726869	4	8026645	*ABLIM2*	G	T	0.8037	1.53E-01
rs6811804	4	774617	*CPLXL*	C	T	0.7581	6.76E-02
rs77547415	5	2560883	*IRX2*	T	C	0.5402	1.34E-01
rs17676033	5	172743694	*STC2*	T	C	1.621	1.17E-01
rs6919765	6	77261591	*IMPG1*	G	G	0.6	1.55E-03
rs6456589	6	24092339	*NRSN1*	C	A	1.41	1.27E-02

rs635551	6	160404832	*IGF2R*	C	T	1.171	4.70E-01
rs62448651	7	15066641	*DGKB*	A	C	1.251	1.13E-01
rs117089947	7	147549030	*CNTNAP2*	C	T	1.115	6.07E-01
rs111777155	9	99820736	*CTSV*	G	T	0.6681	3.65E-01
rs117077124	9	124942205	*MORN5*	A	G	2.576	6.96E-03
rs11792868	9	130480322	*TTC16*	C	A	0.9421	7.62E-01
rs7903146	10	114758349	*TCF7L2*	T	C	1.126	3.74E-01
rs10827154	10	19540391	*MALRD1*	A	G	0.7842	1.16E-01
rs10903783	10	2540623	*PFKP*	T	C	1.347	2.09E-02
rs17096794	12	46604091	*SLC38A1*	G	C	1.363	1.65E-01
rs7953156	12	44929687	*NELL2*	A	G	1.425	1.28E-02
rs11107089	12	93941505	*SOCS2*	A	A	0.4673	6.61E-02
rs11047700	12	25061338	*BCAT1*	T	A	0.5991	2.21E-02
rs17104138	12	68005294	*DYRK2*	G	G	1.007	9.75E-01
rs2324604	13	40971758	*FOXO1*	C	T	1.421	1.01E-02
rs9594071	13	86512325	*SLITRK6*	A	G	0.906	4.43E-01
rs1031302	13	58038376	*PCDH17*	A	A	1.341	6.29E-02
rs9540874	13	67290606	*PCDH9*	T	G	0.636	1.96E-02
rs16944568	15	62044136	*VPS13C*	G	T	0.7781	1.37E-01
rs142668827	15	55114632	*UNC13C*	C	A	0.6196	2.33E-02
rs117027386	16	61969102	*CDH8*	A	C	0.26	1.36E-02
rs13333067	16	3495383	*NAA60*	T	C	2.441	3.84E-04
rs55870929	16	25647503	*HS3ST4*	T	C	0.6362	9.40E-02
rs727836	16	27832033	*GSG1L*	T	G	0.748	9.63E-02
rs913548	20	56053525	*CTCFL*	A	T	1.306	6.42E-02
rs77316551	22	35285864	*ISX*	T	C	0.9212	7.44E-01
rs5759835	22	23816142	*FAM230I*	T	C	0.8493	4.64E-01

Chr: chromosome, A1: reference allele, A2: alternative allele.

Table 6.4 Comparing GWAS results of 25 out of the 48 lead SNPs identified in the primary Arab-Emirati GWAS with European-based GWAS results. Summary statistics for the European-based GWAS was accessed via DIAGRAM (Data was published by Xue et al. 2018). Associations with the same direction of effect are marked in green.

				RAF				EU					EM					
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**EU**	**EM**	**B**	**SE**	**OR**	**P**	**N**	**B**	**SE**	**OR**	**P**	**N**	**DOE**
rs7903146	10	114758349	*TCF7L2*	T	C	0.292	0.423	0.306	0.008	1.36	0.00E+00	597475	0.116	0.021	1.12	4.83E-08	991	1
rs10827154	10	19540391	*MALRD1*	A	G	0.199	0.216	-0.007	0.010	0.99	4.61E-01	578400	-0.138	0.024	0.87	1.60E-08	991	1
rs10903783	10	2540623	*PFKP*	C	T	0.443	0.560	-0.003	0.008	1.00	7.54E-01	572064	0.089	0.019	1.09	4.84E-06	991	-1
rs7953156	12	44929687	*NELL2*	A	G	0.408	0.279	-0.002	0.008	1.00	7.77E-01	575843	0.103	0.022	1.11	2.21E-06	991	-1
rs17096794	12	46604091	*SLC38A1*	G	A	0.108	0.112	0.012	0.013	1.01	3.38E-01	573704	0.147	0.032	1.16	4.73E-06	991	1
rs17104138	12	68005294	*DYRK2*	G	A	0.050	0.066	0.010	0.018	1.01	5.83E-01	571415	-0.183	0.039	0.83	2.71E-06	991	-1
rs2324604	13	40971758	*FOXO1*	C	T	0.315	0.363	0.023	0.009	1.02	8.01E-03	565808	0.107	0.021	1.11	4.16E-07	991	1
rs1031302	13	58038376	*PCDH17*	A	G	0.336	0.230	-0.004	0.008	1.00	6.15E-01	563345	0.108	0.023	1.11	3.01E-06	991	-1
rs9540874	13	67290606	*PCDH9*	T	A	0.113	0.141	-0.008	0.012	0.99	5.16E-01	568692	-0.135	0.029	0.87	3.95E-06	991	1
rs9594071	13	86512325	*SLITRK6*	A	G	0.445	0.456	0.003	0.008	1.00	7.36E-01	563585	-0.093	0.020	0.91	3.12E-06	991	-1

rs16944568	15	62044136	*VPS13C*	G	A	0.114	0.163	-0.020	0.012	0.98	1.05E-01	571188	-0.150	0.028	0.86	1.03E-07	991	1
rs55870929	16	25647503	*HS3ST4*	T	C	0.020	0.058	0.008	0.028	1.01	7.80E-01	573274	-0.198	0.042	0.82	2.74E-06	991	-1
rs727836	16	27832033	*GSG1L*	T	C	0.151	0.148	0.019	0.011	1.02	8.09E-02	569701	-0.129	0.028	0.88	3.06E-06	991	-1
rs75875192	2	119217818	*INSIG2*	G	A	0.055	0.050	0.008	0.017	1.01	6.29E-01	572058	0.212	0.044	1.24	1.12E-06	991	1
rs4487114	2	23050710	*KLHL29*	G	A	0.386	0.313	-0.003	0.008	1.00	6.98E-01	569281	0.120	0.023	1.13	1.40E-07	991	-1
rs13016276	2	69391521	*ANTXR1*	G	T	0.341	0.453	0.013	0.008	1.01	1.11E-01	577833	-0.105	0.020	0.90	1.93E-07	991	-1
rs913548	20	56053525	*CTCFL*	A	T	0.323	0.314	0.010	0.008	1.01	2.50E-01	566555	0.112	0.022	1.12	3.98E-07	991	1
rs5759835	22	23816142	*FAM230I*	T	C	0.115	0.077	0.005	0.012	1.01	6.77E-01	567420	-0.173	0.038	0.84	4.87E-06	991	-1
rs2129645	4	141792539	*RNF150*	A	T	0.119	0.187	-0.009	0.012	0.99	4.53E-01	571381	-0.120	0.026	0.89	4.03E-06	991	1
rs511804	4	20545268	*SLIT2*	G	C	0.340	0.146	-0.003	0.008	1.00	7.40E-01	564516	-0.129	0.027	0.88	2.79E-06	991	1
rs2457417	4	48213733	*TEC*	A	T	0.079	0.149	-0.013	0.015	0.99	3.87E-01	571914	0.145	0.030	1.16	1.02E-06	991	-1
rs17676033	5	172743694	*STC2*	T	C	0.028	0.060	0.053	0.024	1.05	2.41E-02	576846	0.195	0.042	1.22	2.63E-06	991	1
rs6456589	6	24092339	*NRSN1*	C	G	0.287	0.377	0.008	0.009	1.01	3.49E-01	571242	0.108	0.021	1.11	3.21E-07	991	1
rs117089947	7	147549030	*CNTNAP2*	C	T	0.051	0.060	0.006	0.018	1.01	7.37E-01	572991	-0.184	0.040	0.83	4.32E-06	991	-1

rs117077124	9	124942205	*MORN5*	A	G	0.023	0.047	0.024	0.026	1.02	3.44E-01	573545	0.233	0.049	1.26	1.70E-06	991	1

Chr: chromosome, RAF: risk allele frequency; A1: reference allele, A2: alternative allele, DOE: Direction of effect comparing European GWAS results and the original Arab-Emirati GWAS results.

Table 6.5 Comparing GWAS results of 45 out of the 48 sentinel SNPs identified in the original Arab-Emirati GWAS with South Asian-based GWAS results. Summary statistics of the South Asian-based GWAS was accessed via the GWAS catalogue (accession number: GCST90093110). Associations with the same direction of effect are marked in green.

				RAF				SA					EM					
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**SA**	**EM**	**B**	**SE**	**OR**	**P**	**N**	**B**	**SE**	**OR**	**P**	**N**	**DOE**
rs2861451	1	167312409	*POU2F1*	T	G	0.010	0.065	0.266	0.092	1.30	3.68E-03	50533	0.194	0.039	1.21	5.28E-07	991	1
rs12119289	1	48008522	*FOXD2*	A	G	0.303	0.327	-0.012	0.017	0.99	5.05E-01	50533	-0.100	0.022	0.91	4.49E-06	991	1
rs7903146	10	114758349	*TCF7L2*	T	C	0.309	0.423	0.264	0.016	1.30	7.94E-59	50533	0.116	0.021	1.12	4.83E-08	991	1
rs10827154	10	19540391	*MALRD1*	A	G	0.161	0.216	0.024	0.021	1.02	2.40E-01	50533	-0.138	0.024	0.87	1.60E-08	991	-1
rs10903783	10	2540623	*PFKP*	T	C	0.411	0.440	-0.022	0.016	0.98	1.59E-01	50533	0.089	0.019	1.09	4.84E-06	991	-1
rs11047700	12	25061338	*BCAT1*	T	C	0.105	0.103	0.008	0.025	1.01	7.46E-01	50533	-0.159	0.033	0.85	1.48E-06	991	-1
rs7953156	12	44929687	*NELL2*	A	G	0.355	0.279	-0.002	0.016	1.00	8.83E-01	50533	0.103	0.022	1.11	2.21E-06	991	-1

rs17096794	12	46604091	*SLC38A1*	A	G	0.867	0.888	0.032	0.023	1.03	1.67E-01	50533	0.147	0.032	1.16	4.73E-06	991	1
rs17104138	12	68005294	*DYRK2*	A	G	0.890	0.934	0.056	0.025	1.06	2.46E-02	50533	-0.183	0.039	0.83	2.71E-06	991	-1
rs11107089	12	93941505	*SOCS2*	A	G	0.019	0.027	-0.041	0.066	0.96	5.33E-01	50533	-0.316	0.062	0.73	3.57E-07	991	1
rs2324604	13	40971758	*FOXO1*	T	C	0.628	0.637	-0.003	0.016	1.00	8.59E-01	50533	0.107	0.021	1.11	4.16E-07	991	-1
rs1031302	13	58038376	*PCDH17*	A	G	0.303	0.230	-0.004	0.017	1.00	8.17E-01	50533	0.108	0.023	1.11	3.01E-06	991	-1
rs9540874	13	67290606	*PCDH9*	A	T	0.880	0.859	-0.051	0.023	0.95	2.86E-02	50533	-0.135	0.029	0.87	3.95E-06	991	1
rs9594071	13	86512325	*SLITRK6*	A	G	0.334	0.456	-0.007	0.016	0.99	6.51E-01	50533	-0.093	0.020	0.91	3.12E-06	991	1
rs142668827	15	55114632	*UNC13C*	T	C	0.983	0.900	-0.027	0.062	0.97	6.66E-01	50533	-0.147	0.033	0.86	9.71E-06	991	1
rs16944568	15	62044136	*VPS13C*	A	G	0.837	0.837	-0.028	0.021	0.97	1.75E-01	50533	-0.150	0.028	0.86	1.03E-07	991	1
rs55870929	16	25647503	*HS3ST4*	T	C	0.007	0.058	0.090	0.097	1.09	3.49E-01	50533	-0.198	0.042	0.82	2.74E-06	991	-1
rs727836	16	27832033	*GSG1L*	T	C	0.098	0.148	-0.021	0.026	0.98	4.01E-01	50533	-0.129	0.028	0.88	3.06E-06	991	1
rs13333067	16	3495383	*NAA60*	T	C	0.067	0.071	0.022	0.031	1.02	4.71E-01	50533	0.173	0.037	1.19	3.27E-06	991	1
rs117027386	16	61969102	*CDH8*	A	G	0.027	0.021	-0.023	0.051	0.98	6.48E-01	50533	-0.344	0.072	0.71	1.62E-06	991	1
rs75875192	2	119217818	*INSIG2*	A	G	0.953	0.950	0.051	0.037	1.05	1.67E-01	50533	0.212	0.044	1.24	1.12E-06	991	1

rs60745626	2	1689209	*PXDN*	T	C	0.238	0.278	0.021	0.020	1.02	2.80E-01	50533	0.104	0.022	1.11	1.93E-06	991	1
rs4487114	2	23050710	*KLHL29*	A	G	0.526	0.687	-0.003	0.015	1.00	8.39E-01	50533	0.120	0.023	1.13	1.40E-07	991	-1
rs13016276	2	69391521	*ANTXR1*	T	G	0.536	0.547	-0.016	0.015	0.98	2.99E-01	50533	-0.105	0.020	0.90	1.93E-07	991	1
rs913548	20	56053525	*CTCFL*	A	T	0.426	0.314	0.002	0.015	1.00	8.85E-01	50533	0.112	0.022	1.12	3.98E-07	991	1
rs5759835	22	23816142	*FAM230I*	T	C	0.133	0.077	-0.033	0.024	0.97	1.74E-01	50533	-0.173	0.038	0.84	4.87E-06	991	1
rs77316551	22	35285864	*ISX*	T	C	0.015	0.045	-0.025	0.068	0.98	7.11E-01	50533	-0.247	0.050	0.78	7.87E-07	991	1
rs150527739	3	14432140	*SLC6A6*	A	G	0.089	0.023	-0.020	0.028	0.98	4.72E-01	50533	-0.323	0.067	0.72	1.76E-06	991	1
rs3934473	3	36828646	*DCLK3*	A	G	0.013	0.045	-0.059	0.080	0.94	4.63E-01	50533	-0.236	0.048	0.79	8.64E-07	991	1
rs144006677	3	51681325	*RAD54L2*	A	C	0.033	0.016	-0.008	0.051	0.99	8.77E-01	50533	-0.375	0.077	0.69	1.26E-06	991	1
rs2129645	4	141792539	*RNF150*	A	T	0.108	0.187	-0.035	0.025	0.97	1.66E-01	50533	-0.120	0.026	0.89	4.03E-06	991	1
rs139588925	4	18131933	*LCORL*	A	G	0.973	0.974	-0.040	0.048	0.96	4.06E-01	50533	-0.294	0.064	0.75	4.23E-06	991	1
rs511804	4	20545268	*SLIT2*	C	G	0.813	0.854	-0.020	0.020	0.98	3.19E-01	50533	-0.129	0.027	0.88	2.79E-06	991	1
rs2457417	4	48213733	*TEC*	A	T	0.118	0.149	-0.002	0.024	1.00	9.47E-01	50533	0.145	0.030	1.16	1.02E-06	991	-1
rs6811804	4	774617	*CPLXL*	C	G	0.194	0.231	0.000	0.019	1.00	9.87E-01	50533	-0.110	0.023	0.90	1.69E-06	991	-1

rs11726869	4	8026645	*ABLIM2*	T	G	0.824	0.779	0.015	0.020	1.02	4.57E-01	50533	-0.113	0.024	0.89	1.88E-06	991	-1
rs17676033	5	172743694	*STC2*	T	C	0.042	0.060	-0.020	0.041	0.98	6.15E-01	50533	0.195	0.042	1.22	2.63E-06	991	-1
rs635551	6	160404832	*IGF2R*	T	C	0.887	0.887	-0.018	0.025	0.98	4.60E-01	50533	0.150	0.032	1.16	2.81E-06	991	-1
rs6456589	6	24092339	*NRSN1*	C	G	0.430	0.377	0.000	0.016	1.00	9.79E-01	50533	0.108	0.021	1.11	3.21E-07	991	-1
rs6919765	6	77261591	*IMPG1*	A	G	0.755	0.755	0.005	0.020	1.00	8.21E-01	50533	-0.122	0.025	0.89	8.86E-07	991	-1
rs117089947	7	147549030	*CNTNAP2*	T	C	0.965	0.940	-0.047	0.042	0.95	2.64E-01	50533	-0.184	0.040	0.83	4.32E-06	991	1
rs62448651	7	15066641	*DGKB*	A	C	0.472	0.441	0.044	0.017	1.05	1.10E-02	50533	0.105	0.022	1.11	1.88E-06	991	1
rs117077124	9	124942205	*MORN5*	A	G	0.060	0.047	0.050	0.033	1.05	1.29E-01	50533	0.233	0.049	1.26	1.70E-06	991	1
rs11792868	9	130480322	*TTC16*	A	C	0.879	0.908	0.004	0.024	1.00	8.60E-01	50533	-0.160	0.033	0.85	1.43E-06	991	-1
rs111777155	9	99820736	*CTSV*	T	G	0.958	0.983	0.067	0.040	1.07	9.37E-02	50533	-0.376	0.082	0.69	4.06E-06	991	-1

Chr: chromosome, RAF: risk allele frequency; A1: reference allele, A2: alternative allele, DOE: Direction of effect comparing South Asian GWAS results and the original Arab-Emirati GWAS results.

Table 6.6 SNPs identified to be associated with T2D with P < 1.0E-4 in both the European-based GWAS and the original Arab-Emirati GWAS. Summary statistics for the European-based GWAS was accessed via DIAGRAM (Data was published by Xue et al. 2018). Associations with the same direction of effect are marked in green.

				RAF				EU					EM					
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**EU**	**EM**	**B**	**SE**	**OR**	**P**	**N**	**B**	**SE**	**OR**	**P**	**N**	**DOE**
rs4074720	10	114748497	*TCF7L2*	T	C	0.467	0.581	0.173	0.008	1.189	4.38E-109	597475	-0.082	0.021	1.086	8.08E-05	991	1
rs4074718	10	114748617	*TCF7L2*	A	G	0.467	0.581	0.174	0.007	1.190	4.96E-127	597475	-0.082	0.021	1.086	8.08E-05	991	1
rs35011184	10	114749734	*TCF7L2*	A	G	0.227	0.357	0.280	0.009	1.323	2.98E-221	597475	0.093	0.022	1.098	2.18E-05	991	1
rs61875118	10	114750236	*TCF7L2*	C	T	0.228	0.358	0.272	0.009	1.312	7.82E-202	597475	0.092	0.022	1.097	2.83E-05	991	1
rs61875119	10	114750380	*TCF7L2*	A	G	0.227	0.353	0.272	0.009	1.313	2.17E-202	597475	0.091	0.022	1.095	4.32E-05	991	1
rs72826075	10	114750500	*TCF7L2*	G	A	0.227	0.358	0.280	0.009	1.323	9.26E-222	597475	0.092	0.022	1.097	2.83E-05	991	1
rs17747324	10	114752503	*TCF7L2*	C	T	0.227	0.358	0.282	0.008	1.326	1.35E-253	597475	0.094	0.022	1.099	1.85E-05	991	1
rs61875120	10	114753259	*TCF7L2*	C	T	0.220	0.357	0.280	0.009	1.324	2.22E-211	597475	0.093	0.022	1.098	2.22E-05	991	1
rs34872471	10	114754071	*TCF7L2*	C	T	0.291	0.432	0.292	0.008	1.340	5.17E-270	597475	0.114	0.021	1.120	7.36E-08	991	1
rs7901695	10	114754088	*TCF7L2*	C	T	0.309	0.448	0.279	0.008	1.322	1.29E-301	597475	0.105	0.021	1.111	6.34E-07	991	1
rs35198068	10	114754784	*TCF7L2*	C	T	0.291	0.431	0.292	0.008	1.339	5.90E-270	597475	0.113	0.021	1.120	8.89E-08	991	1
rs4506565	10	114756041	*TCF7L2*	T	A	0.313	0.446	0.280	0.008	1.323	4.46E-305	597475	0.109	0.021	1.115	2.85E-07	991	1

rs7903146	10	114758349	*TCF7L2*	T	C	0.292	0.423	0.306	0.008	1.358	0.00E+00	597475	0.116	0.021	1.123	4.83E-08	991	1
rs4267006	10	114758779	*TCF7L2*	T	G	0.221	0.342	0.287	0.009	1.333	1.56E-230	597475	0.103	0.022	1.109	4.02E-06	991	1
rs10885402	10	114761697	*TCF7L2*	C	A	0.458	0.575	0.176	0.007	1.192	3.51E-131	597475	-0.086	0.021	1.090	3.46E-05	991	1
rs6585197	10	114762064	*TCF7L2*	T	G	0.456	0.582	0.169	0.008	1.184	9.52E-107	597475	-0.087	0.021	1.091	3.42E-05	991	1
rs6585198	10	114762237	*TCF7L2*	A	G	0.458	0.581	0.175	0.007	1.191	1.96E-129	597475	-0.085	0.021	1.089	4.69E-05	991	1
rs10787471	10	114764661	*TCF7L2*	A	G	0.443	0.579	0.159	0.008	1.172	7.80E-91	597475	-0.084	0.021	1.088	5.95E-05	991	1
rs6585199	10	114765171	*TCF7L2*	A	G	0.457	0.582	0.171	0.008	1.187	7.75E-106	597475	-0.087	0.021	1.091	3.66E-05	991	1
rs4575195	10	114765747	*TCF7L2*	A	C	0.313	0.469	0.268	0.008	1.308	6.49E-241	597475	0.096	0.021	1.101	5.56E-06	991	1
rs4132670	10	114767771	*TCF7L2*	A	G	0.313	0.469	0.274	0.008	1.315	8.65E-293	597475	0.095	0.021	1.100	7.12E-06	991	1
rs6585200	10	114768609	*TCF7L2*	G	A	0.458	0.583	0.174	0.007	1.190	2.08E-128	597475	-0.088	0.021	1.092	2.66E-05	991	1
rs6585201	10	114768783	*TCF7L2*	A	G	0.456	0.570	0.174	0.008	1.190	6.51E-111	597475	-0.084	0.021	1.088	6.22E-05	991	1
rs55899248	10	114773608	*TCF7L2*	G	A	0.221	0.347	0.284	0.009	1.329	4.48E-227	597475	0.100	0.022	1.105	8.47E-06	991	1
rs12244851	10	114773926	*TCF7L2*	T	C	0.320	0.430	0.255	0.008	1.291	6.92E-213	597475	0.090	0.022	1.094	2.94E-05	991	1
rs7904519	10	114773927	*TCF7L2*	G	A	0.467	0.585	0.167	0.007	1.181	1.35E-114	597475	-0.086	0.021	1.090	4.32E-05	991	1
rs36090025	10	114774433	*TCF7L2*	C	A	0.299	0.407	0.275	0.008	1.317	6.08E-242	597475	0.102	0.022	1.108	2.28E-06	991	1
rs7918599	10	114777396	*TCF7L2*	T	C	0.466	0.584	0.165	0.008	1.180	8.73E-99	597475	-0.088	0.021	1.092	2.85E-05	991	1

rs10885405	10	114777670	*TCF7L2*	T	C	0.467	0.576	0.167	0.007	1.182	1.09E-119	597475	-0.084	0.021	1.088	6.53E-05	991	1
rs10885406	10	114777724	*TCF7L2*	G	A	0.468	0.583	0.167	0.007	1.182	9.05E-120	597475	-0.087	0.021	1.090	3.93E-05	991	1
rs11196190	10	114778252	*TCF7L2*	G	A	0.466	0.583	0.165	0.008	1.179	1.58E-98	597475	-0.087	0.021	1.090	3.93E-05	991	1
rs7899529	10	114779538	*TCF7L2*	A	G	0.466	0.581	0.166	0.008	1.180	1.86E-99	597475	-0.085	0.021	1.088	6.00E-05	991	1
rs11196191	10	114780633	*TCF7L2*	C	A	0.466	0.584	0.159	0.008	1.172	2.09E-94	597475	-0.088	0.021	1.092	2.85E-05	991	1
rs10787472	10	114781297	*TCF7L2*	C	A	0.468	0.583	0.168	0.007	1.183	2.17E-124	597475	-0.087	0.021	1.090	3.93E-05	991	1
rs10787473	10	114781400	*TCF7L2*	A	C	0.466	0.581	0.160	0.008	1.174	1.49E-96	597475	-0.085	0.021	1.088	6.00E-05	991	1
rs12258200	10	114781698	*TCF7L2*	C	T	0.466	0.584	0.161	0.008	1.174	2.46E-97	597475	-0.088	0.021	1.092	2.85E-05	991	1
rs55853916	10	114781950	*TCF7L2*	C	G	0.221	0.345	0.284	0.009	1.329	2.33E-227	597475	0.103	0.022	1.109	3.98E-06	991	1
rs55972445	10	114782790	*TCF7L2*	T	G	0.221	0.347	0.284	0.009	1.329	7.23E-227	597475	0.100	0.022	1.105	8.27E-06	991	1
rs6585202	10	114782803	*TCF7L2*	C	T	0.466	0.576	0.161	0.008	1.175	2.48E-97	597475	-0.084	0.021	1.088	6.53E-05	991	1
rs7921525	10	114783333	*TCF7L2*	T	C	0.464	0.577	0.160	0.008	1.173	3.79E-92	597475	-0.086	0.021	1.089	4.52E-05	991	1
rs6585203	10	114783403	*TCF7L2*	G	C	0.466	0.582	0.165	0.008	1.180	5.81E-99	597475	-0.086	0.021	1.090	3.94E-05	991	1
rs6585204	10	114783586	*TCF7L2*	G	C	0.466	0.575	0.165	0.008	1.180	9.92E-99	597475	-0.084	0.021	1.087	6.53E-05	991	1
rs11196193	10	114783838	*TCF7L2*	C	T	0.466	0.584	0.165	0.008	1.180	1.01E-98	597475	-0.089	0.021	1.093	2.45E-05	991	1
rs4309084	10	114785278	*TCF7L2*	G	A	0.466	0.582	0.165	0.008	1.179	5.22E-98	597475	-0.086	0.021	1.090	3.94E-05	991	1

rs7074440	10	114785424	*TCF7L2*	A	G	0.300	0.406	0.282	0.008	1.325	3.03E-261	597475	0.103	0.022	1.108	2.23E-06	991	1
rs4128598	10	114785939	*TCF7L2*	G	A	0.466	0.580	0.164	0.008	1.179	8.28E-98	597475	-0.085	0.021	1.088	6.02E-05	991	1
rs4128597	10	114786154	*TCF7L2*	G	A	0.466	0.582	0.164	0.008	1.178	2.01E-97	597475	-0.086	0.021	1.090	3.94E-05	991	1
rs7924080	10	114787012	*TCF7L2*	C	T	0.469	0.582	0.167	0.007	1.181	2.45E-118	597475	-0.086	0.021	1.090	3.88E-05	991	1
rs7907610	10	114787090	*TCF7L2*	G	A	0.467	0.582	0.160	0.008	1.173	1.45E-95	597475	-0.086	0.021	1.090	3.88E-05	991	1
rs56299331	10	114788436	*TCF7L2*	T	C	0.201	0.319	0.272	0.009	1.313	1.61E-194	597475	0.113	0.023	1.120	6.62E-07	991	1
rs12243326	10	114788815	*TCF7L2*	C	T	0.282	0.394	0.262	0.008	1.300	9.44E-218	597475	0.109	0.022	1.115	4.50E-07	991	1
rs7077039	10	114789077	*TCF7L2*	C	T	0.462	0.573	0.171	0.007	1.186	3.36E-127	597475	-0.085	0.021	1.088	5.42E-05	991	1
rs10885408	10	114791388	*TCF7L2*	G	A	0.462	0.580	0.166	0.008	1.181	1.21E-99	597475	-0.086	0.021	1.090	3.86E-05	991	1
rs12260037	10	114791490	*TCF7L2*	T	C	0.283	0.386	0.260	0.008	1.296	2.61E-209	597475	0.106	0.022	1.112	9.09E-07	991	1
rs7086857	10	114793237	*TCF7L2*	G	A	0.461	0.584	0.168	0.008	1.183	9.70E-102	597475	-0.082	0.021	1.086	8.70E-05	991	1
rs12359102	10	114793451	*TCF7L2*	A	G	0.461	0.584	0.164	0.008	1.178	1.98E-100	597475	-0.082	0.021	1.086	8.70E-05	991	1
rs4367880	10	114795256	*TCF7L2*	C	G	0.201	0.319	0.266	0.009	1.304	1.03E-179	597475	0.113	0.023	1.120	6.62E-07	991	1
rs61872784	10	114798893	*TCF7L2*	A	T	0.281	0.375	0.257	0.008	1.293	4.13E-211	597475	0.106	0.022	1.111	1.18E-06	991	1
rs56087297	10	114799172	*TCF7L2*	G	A	0.281	0.365	0.267	0.008	1.306	8.95E-227	597475	0.110	0.022	1.116	4.93E-07	991	1
rs72826094	10	114801488	*TCF7L2*	T	A	0.201	0.318	0.271	0.009	1.311	6.03E-191	597475	0.114	0.023	1.121	5.46E-07	991	1

rs61872786	10	114806697	*TCF7L2*	A	G	0.207	0.330	0.262	0.009	1.300	8.59E-184	597475	0.103	0.023	1.108	4.81E-06	991	1
rs12255372	10	114808902	*TCF7L2*	T	G	0.287	0.401	0.264	0.008	1.303	7.97E-262	597475	0.093	0.022	1.098	1.41E-05	991	1
rs11196211	10	114817009	*TCF7L2*	C	A	0.284	0.382	0.255	0.008	1.290	5.64E-202	597475	0.105	0.022	1.111	1.17E-06	991	1
rs11651755	17	36099840	*HNF1B*	C	T	0.485	0.585	0.074	0.008	1.077	8.98E-22	597475	-0.084	0.021	1.087	9.00E-05	991	1
rs6066137	20	45590658	*EYA2*	T	C	0.261	0.192	-0.052	0.009	0.950	1.14E-08	597475	-0.103	0.026	0.903	7.50E-05	991	1
rs1211644	20	45592842	*EYA2*	C	T	0.262	0.233	-0.051	0.009	0.950	1.30E-08	597475	-0.103	0.024	0.902	2.44E-05	991	1
rs1206773	20	45593210	*EYA2*	A	G	0.259	0.215	-0.048	0.008	0.954	1.03E-08	597475	-0.101	0.025	0.904	4.75E-05	991	1
rs6066138	20	45594711	*EYA2*	A	G	0.278	0.205	-0.049	0.008	0.952	1.93E-09	597475	-0.100	0.025	0.905	8.29E-05	991	1
rs6063046	20	45596378	*EYA2*	G	A	0.281	0.205	-0.052	0.009	0.949	2.67E-09	597475	-0.100	0.025	0.905	8.29E-05	991	1
rs6063048	20	45598564	*EYA2*	A	G	0.279	0.204	-0.052	0.009	0.949	3.29E-09	597475	-0.100	0.025	0.905	9.32E-05	991	1
rs59306735	3	12062003	*PPARG*	C	T	0.065	0.061	-0.123	0.016	0.884	6.92E-14	597475	-0.183	0.047	0.833	9.10E-05	991	1
rs13216165	6	20717121	*CDKAL1*	G	A	0.430	0.306	-0.038	0.007	0.963	2.66E-07	597475	-0.095	0.023	0.909	2.22E-05	991	1
rs6928012	6	20728513	*CDKAL1*	C	T	0.400	0.464	0.057	0.007	1.059	3.90E-15	597475	0.085	0.021	1.089	5.82E-05	991	1
rs7800343	7	15021223	*DGKB*	A	G	0.381	0.260	0.035	0.007	1.036	1.63E-06	597475	0.106	0.024	1.112	1.12E-05	991	1
rs2358157	7	15023729	*DGKB*	G	A	0.405	0.261	0.038	0.008	1.039	1.28E-06	597475	0.101	0.024	1.106	2.17E-05	991	1
rs12375103	7	15025396	*DGKB*	A	T	0.363	0.234	0.033	0.008	1.034	3.68E-05	597475	0.100	0.024	1.105	3.91E-05	991	1

rs12670968	7	15026787	*DGKB*	A	G	0.364	0.238	0.034	0.008	1.034	2.22E-05	597475	0.100	0.024	1.105	3.80E-05	991	1
rs6461145	7	15026872	*DGKB*	G	A	0.403	0.259	0.040	0.007	1.040	4.18E-08	597475	0.103	0.024	1.108	1.72E-05	991	1
rs62448613	7	15035755	*DGKB*	T	C	0.347	0.225	0.033	0.008	1.034	3.77E-05	597475	0.097	0.025	1.102	9.16E-05	991	1
rs17168579	7	15038077	*DGKB*	G	A	0.348	0.226	0.033	0.008	1.034	2.04E-05	597475	0.098	0.025	1.103	7.85E-05	991	1
rs11971503	7	15038452	*DGKB*	T	C	0.349	0.269	0.035	0.008	1.035	6.26E-06	597475	0.104	0.024	1.109	1.00E-05	991	1
rs62448614	7	15039249	*DGKB*	A	G	0.348	0.309	0.034	0.008	1.035	3.29E-05	597475	0.094	0.023	1.099	2.91E-05	991	1
rs1002061	7	15039499	*DGKB*	T	A	0.480	0.444	-0.058	0.007	0.943	4.00E-16	597475	-0.085	0.021	0.919	4.53E-05	991	1
rs1005256	7	15039605	*DGKB*	C	A	0.401	0.379	0.043	0.008	1.044	1.09E-08	597475	0.097	0.022	1.102	6.56E-06	991	1
rs10231021	7	15060429	*DGKB*	T	A	0.474	0.540	-0.065	0.008	0.937	1.82E-16	597475	0.086	0.021	0.918	5.12E-05	991	1
rs80005921	7	15063379	*DGKB*	G	A	0.434	0.311	0.053	0.008	1.055	2.26E-11	597475	0.094	0.023	1.099	4.09E-05	991	1

Chr: chromosome, RAF: risk allele frequency; A1: reference allele, A2: alternative allele, DOE: Direction of effect comparing European GWAS results and the original Arab-Emirati GWAS results.

Table 6.7 SNPs identified to be associated with T2D with P < 1.0E-4 in both the South Asian-based GWAS and the original Arab-Emirati GWAS. Summary statistics of the South Asian-based GWAS was accessed via the GWAS catalogue (accession number: GCST90093110). Associations with the same direction of effect are marked in green.

				RAF				SA					EM					
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**SA**	**EM**	**B**	**SE**	**OR**	**P**	**N**	**B**	**SE**	**OR**	**P**	**N**	**DOE**
rs4074720	10	114748497	*TCF7L2*	T	C	0.408	0.581	0.179	0.016	1.196	5.36E-30	50533	-0.082	0.021	1.086	8.08E-05	991	1
rs4074718	10	114748617	*TCF7L2*	A	G	0.408	0.581	0.179	0.016	1.196	4.52E-30	50533	-0.082	0.021	1.086	8.08E-05	991	1
rs35011184	10	114749734	*TCF7L2*	A	G	0.256	0.357	0.237	0.018	1.267	2.08E-41	50533	0.093	0.022	1.098	2.18E-05	991	1
rs61875118	10	114750236	*TCF7L2*	T	C	0.744	0.642	-0.236	0.018	0.790	2.98E-41	50533	0.092	0.022	0.912	2.83E-05	991	1
rs61875119	10	114750380	*TCF7L2*	A	G	0.236	0.353	0.229	0.018	1.258	9.01E-37	50533	0.091	0.022	1.095	4.32E-05	991	1
rs72826075	10	114750500	*TCF7L2*	A	G	0.744	0.642	-0.237	0.018	0.789	2.04E-41	50533	0.092	0.022	0.912	2.83E-05	991	1
rs17747324	10	114752503	*TCF7L2*	T	C	0.745	0.642	-0.235	0.018	0.790	4.04E-41	50533	0.094	0.022	0.910	1.85E-05	991	1
rs61875120	10	114753259	*TCF7L2*	T	C	0.745	0.643	-0.235	0.018	0.791	3.31E-41	50533	0.093	0.022	0.911	2.22E-05	991	1
rs34872471	10	114754071	*TCF7L2*	T	C	0.691	0.568	-0.263	0.016	0.769	5.20E-58	50533	0.114	0.021	0.892	7.36E-08	991	1
rs7901695	10	114754088	*TCF7L2*	T	C	0.692	0.552	-0.252	0.016	0.778	1.61E-53	50533	0.105	0.021	0.900	6.34E-07	991	1
rs35198068	10	114754784	*TCF7L2*	T	C	0.691	0.569	-0.263	0.016	0.768	2.22E-58	50533	0.113	0.021	0.893	8.89E-08	991	1
rs4506565	10	114756041	*TCF7L2*	A	T	0.687	0.554	-0.257	0.016	0.773	4.31E-56	50533	0.109	0.021	0.897	2.85E-07	991	1

rs7903146	10	114758349	*TCF7L2*	T	C	0.309	0.423	0.264	0.016	1.302	7.94E-59	50533	0.116	0.021	1.123	4.83E-08	991	1
rs4267006	10	114758779	*TCF7L2*	T	G	0.235	0.342	0.228	0.018	1.256	3.45E-37	50533	0.103	0.022	1.109	4.02E-06	991	1
rs10885402	10	114761697	*TCF7L2*	A	C	0.587	0.425	-0.179	0.016	0.836	5.73E-31	50533	-0.086	0.021	0.917	3.46E-05	991	1
rs6585197	10	114762064	*TCF7L2*	T	G	0.413	0.582	0.179	0.016	1.195	7.87E-31	50533	-0.087	0.021	1.091	3.42E-05	991	1
rs6585198	10	114762237	*TCF7L2*	A	G	0.413	0.581	0.179	0.016	1.196	6.50E-31	50533	-0.085	0.021	1.089	4.69E-05	991	1
rs10787471	10	114764661	*TCF7L2*	A	G	0.402	0.579	0.180	0.016	1.197	9.58E-31	50533	-0.084	0.021	1.088	5.95E-05	991	1
rs6585199	10	114765171	*TCF7L2*	A	G	0.413	0.582	0.179	0.015	1.196	2.87E-31	50533	-0.087	0.021	1.091	3.66E-05	991	1
rs4575195	10	114765747	*TCF7L2*	A	C	0.294	0.469	0.246	0.017	1.279	1.65E-49	50533	0.096	0.021	1.101	5.56E-06	991	1
rs4132670	10	114767771	*TCF7L2*	A	G	0.294	0.469	0.246	0.017	1.278	2.64E-49	50533	0.095	0.021	1.100	7.12E-06	991	1
rs6585200	10	114768609	*TCF7L2*	A	G	0.583	0.417	-0.177	0.015	0.838	1.96E-30	50533	-0.088	0.021	0.916	2.66E-05	991	1
rs6585201	10	114768783	*TCF7L2*	A	G	0.376	0.570	0.177	0.016	1.194	1.74E-29	50533	-0.084	0.021	1.088	6.22E-05	991	1
rs55899248	10	114773608	*TCF7L2*	A	G	0.764	0.653	-0.225	0.018	0.799	1.83E-36	50533	0.100	0.022	0.905	8.47E-06	991	1
rs12244851	10	114773926	*TCF7L2*	T	C	0.308	0.430	0.254	0.016	1.289	4.05E-54	50533	0.090	0.022	1.094	2.94E-05	991	1
rs7904519	10	114773927	*TCF7L2*	A	G	0.588	0.415	-0.177	0.015	0.838	1.70E-30	50533	-0.086	0.021	0.918	4.32E-05	991	1
rs36090025	10	114774433	*TCF7L2*	A	C	0.713	0.593	-0.250	0.017	0.778	1.28E-50	50533	0.102	0.022	0.903	2.28E-06	991	1
rs7918599	10	114777396	*TCF7L2*	T	C	0.412	0.584	0.176	0.015	1.192	3.14E-30	50533	-0.088	0.021	1.092	2.85E-05	991	1

rs10885405	10	114777670	*TCF7L2*	T	C	0.391	0.576	0.163	0.016	1.177	1.71E-25	50533	-0.084	0.021	1.088	6.53E-05	991	1
rs10885406	10	114777724	*TCF7L2*	A	G	0.588	0.417	-0.176	0.015	0.838	2.82E-30	50533	-0.087	0.021	0.917	3.93E-05	991	1
rs11196190	10	114778252	*TCF7L2*	A	G	0.588	0.417	-0.176	0.015	0.838	2.85E-30	50533	-0.087	0.021	0.917	3.93E-05	991	1
rs7899529	10	114779538	*TCF7L2*	A	G	0.408	0.581	0.176	0.015	1.193	3.75E-30	50533	-0.085	0.021	1.088	6.00E-05	991	1
rs11196191	10	114780633	*TCF7L2*	A	C	0.588	0.416	-0.176	0.015	0.839	4.78E-30	50533	-0.088	0.021	0.916	2.85E-05	991	1
rs10787472	10	114781297	*TCF7L2*	A	C	0.588	0.417	-0.176	0.015	0.839	3.61E-30	50533	-0.087	0.021	0.917	3.93E-05	991	1
rs10787473	10	114781400	*TCF7L2*	A	C	0.408	0.581	0.176	0.015	1.192	4.95E-30	50533	-0.085	0.021	1.088	6.00E-05	991	1
rs12258200	10	114781698	*TCF7L2*	T	C	0.588	0.416	-0.176	0.015	0.839	5.07E-30	50533	-0.088	0.021	0.916	2.85E-05	991	1
rs55853916	10	114781950	*TCF7L2*	C	G	0.235	0.345	0.225	0.018	1.252	2.46E-36	50533	0.103	0.022	1.109	3.98E-06	991	1
rs55972445	10	114782790	*TCF7L2*	T	G	0.236	0.347	0.224	0.018	1.252	3.45E-36	50533	0.100	0.022	1.105	8.27E-06	991	1
rs6585202	10	114782803	*TCF7L2*	T	C	0.609	0.424	-0.162	0.016	0.851	3.13E-25	50533	-0.084	0.021	0.919	6.53E-05	991	1
rs7921525	10	114783333	*TCF7L2*	T	C	0.407	0.577	0.176	0.016	1.192	7.91E-30	50533	-0.086	0.021	1.089	4.52E-05	991	1
rs6585203	10	114783403	*TCF7L2*	C	G	0.589	0.418	-0.176	0.015	0.839	5.21E-30	50533	-0.086	0.021	0.917	3.94E-05	991	1
rs6585204	10	114783586	*TCF7L2*	C	G	0.610	0.425	-0.162	0.016	0.851	4.24E-25	50533	-0.084	0.021	0.920	6.53E-05	991	1
rs11196193	10	114783838	*TCF7L2*	T	C	0.589	0.416	-0.175	0.016	0.839	9.22E-30	50533	-0.089	0.021	0.915	2.45E-05	991	1
rs4309084	10	114785278	*TCF7L2*	A	G	0.589	0.418	-0.175	0.016	0.840	1.72E-29	50533	-0.086	0.021	0.917	3.94E-05	991	1

rs7074440	10	114785424	*TCF7L2*	A	G	0.286	0.406	0.251	0.017	1.285	3.36E-50	50533	0.103	0.022	1.108	2.23E-06	991	1
rs4128598	10	114785939	*TCF7L2*	A	G	0.592	0.420	-0.173	0.016	0.841	5.79E-29	50533	-0.085	0.021	0.919	6.02E-05	991	1
rs4128597	10	114786154	*TCF7L2*	A	G	0.589	0.418	-0.172	0.016	0.842	1.33E-28	50533	-0.086	0.021	0.917	3.94E-05	991	1
rs7924080	10	114787012	*TCF7L2*	T	C	0.588	0.418	-0.172	0.016	0.842	1.35E-28	50533	-0.086	0.021	0.917	3.88E-05	991	1
rs7907610	10	114787090	*TCF7L2*	A	G	0.588	0.418	-0.172	0.016	0.842	1.48E-28	50533	-0.086	0.021	0.917	3.88E-05	991	1
rs56299331	10	114788436	*TCF7L2*	T	C	0.208	0.319	0.211	0.019	1.234	1.90E-29	50533	0.113	0.023	1.120	6.62E-07	991	1
rs12243326	10	114788815	*TCF7L2*	T	C	0.741	0.606	-0.237	0.017	0.789	6.84E-43	50533	0.109	0.022	0.897	4.50E-07	991	1
rs7077039	10	114789077	*TCF7L2*	T	C	0.614	0.427	-0.156	0.016	0.855	2.52E-23	50533	-0.085	0.021	0.919	5.42E-05	991	1
rs10885408	10	114791388	*TCF7L2*	A	G	0.597	0.420	-0.168	0.016	0.845	3.09E-27	50533	-0.086	0.021	0.917	3.86E-05	991	1
rs12260037	10	114791490	*TCF7L2*	T	C	0.259	0.386	0.238	0.017	1.268	3.37E-43	50533	0.106	0.022	1.112	9.09E-07	991	1
rs7086857	10	114793237	*TCF7L2*	A	G	0.594	0.416	-0.168	0.016	0.845	1.85E-27	50533	-0.082	0.021	0.921	8.70E-05	991	1
rs12359102	10	114793451	*TCF7L2*	A	G	0.406	0.584	0.169	0.016	1.184	1.63E-27	50533	-0.082	0.021	1.086	8.70E-05	991	1
rs4367880	10	114795256	*TCF7L2*	C	G	0.208	0.319	0.210	0.019	1.233	3.34E-29	50533	0.113	0.023	1.120	6.62E-07	991	1
rs61872784	10	114798893	*TCF7L2*	A	T	0.257	0.375	0.237	0.017	1.267	1.47E-42	50533	0.106	0.022	1.111	1.18E-06	991	1
rs56087297	10	114799172	*TCF7L2*	A	G	0.743	0.635	-0.237	0.017	0.789	1.69E-42	50533	0.110	0.022	0.896	4.93E-07	991	1
rs72826094	10	114801488	*TCF7L2*	A	T	0.791	0.682	-0.207	0.019	0.813	1.92E-28	50533	0.114	0.023	0.892	5.46E-07	991	1

rs61872786	10	114806697	*TCF7L2*	A	G	0.209	0.330	0.207	0.019	1.230	1.28E-28	50533	0.103	0.023	1.108	4.81E-06	991	1
rs12255372	10	114808902	*TCF7L2*	T	G	0.258	0.401	0.236	0.017	1.266	5.79E-43	50533	0.093	0.022	1.098	1.41E-05	991	1
rs11196211	10	114817009	*TCF7L2*	A	C	0.743	0.618	-0.238	0.018	0.789	1.39E-41	50533	0.105	0.022	0.900	1.17E-06	991	1
rs7405776	17	36093022	*HNF1B*	A	G	0.318	0.437	0.075	0.017	1.078	1.00E-05	50533	0.093	0.021	1.097	1.31E-05	991	1
rs2005705	17	36096300	*HNF1B*	A	G	0.344	0.526	0.076	0.017	1.079	4.09E-06	50533	-0.086	0.021	1.090	4.86E-05	991	1
rs4430796	17	36098040	*HNF1B*	A	G	0.638	0.461	-0.082	0.016	0.922	2.73E-07	50533	-0.090	0.021	0.914	2.43E-05	991	1
rs7216163	17	36099840	*HNF1B*	T	C	0.583	0.415	-0.084	0.016	0.919	1.01E-07	50533	-0.084	0.021	0.920	9.00E-05	991	1
rs59306735	3	12062003	*PPARG*	T	C	0.944	0.939	0.132	0.034	1.142	8.24E-05	50533	-0.183	0.047	1.200	9.10E-05	991	1
rs180674385	6	20728513	*CDKAL1*	T	C	0.498	0.536	-0.063	0.015	0.939	3.34E-05	50533	0.085	0.021	0.918	5.82E-05	991	1
rs142841100	7	15060429	*DGKB*	A	T	0.579	0.460	0.082	0.016	1.086	1.03E-07	50533	0.086	0.021	1.089	5.12E-05	991	1
rs80005921	7	15063379	*DGKB*	A	G	0.535	0.689	-0.070	0.016	0.933	7.17E-06	50533	0.094	0.023	0.910	4.09E-05	991	1

Chr: chromosome, RAF: risk allele frequency; A1: reference allele, A2: alternative allele, DOE: Direction of effect comparing South Asian GWAS results and the original Arab-Emirati GWAS results.

Table 6.8 Results of SNPs that reached genome-wide significance with P < 5.0e-8 in the meta-analysis performed using the primary Arab-Emirati GWAS and the South Asian-based GWAS summary statistics but did not reach genome-wide significance in the South Asian-only meta-analysis. Summary statistics of the South Asian-based GWAS was accessed via the GWAS catalogue (accession number: GCST90093110).

						Meta-analysis		SA		EM	
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**OR**	**P**	**OR**	**P**	**OR**	**P**
rs556912302	10	12311465	*CDC123*	A	G	1.070	4.83E-08	1.073	5.64E-06	1.065	2.37E-03
rs10830963	11	92708710	*MTNR1B*	C	G	0.926	3.12E-09	0.920	2.50E-07	0.936	2.93E-03
rs2005705	17	36096300	*HNF1B*	A	G	1.083	9.39E-10	1.079	4.09E-06	1.090	4.86E-05
rs7216163	17	36099840	*HNF1B*	T	C	0.919	3.76E-11	0.919	1.01E-07	0.920	9.00E-05
rs11657964	17	36100767	*HNF1B*	A	G	1.077	1.17E-08	1.080	3.23E-06	1.072	9.40E-04
rs7501939	17	36101156	*HNF1B*	T	C	1.075	1.91E-08	1.080	3.06E-06	1.068	1.72E-03
rs8064454	17	36101586	*HNF1B*	A	C	1.088	4.87E-11	1.091	7.02E-08	1.084	1.63E-04
rs11651052	17	36102381	*HNF1B*	A	G	1.084	4.43E-10	1.085	4.90E-07	1.082	2.17E-04
rs11658063	17	36103872	*HNF1B*	C	G	1.076	2.10E-08	1.083	2.02E-06	1.065	2.40E-03
rs59306735	3	12062003	*PPARG*	T	C	1.161	4.15E-08	1.142	8.24E-05	1.200	9.10E-05
rs6781687	3	12216260	*PPARG*	A	G	1.162	3.35E-08	1.151	2.13E-05	1.185	3.71E-04
rs74280642	3	12262934	*PPARG*	A	G	0.841	4.68E-08	0.846	2.46E-06	0.823	5.34E-03
rs532518450	3	185404193	*IGF2BP2*	C	G	1.108	8.97E-09	1.100	1.38E-06	1.147	1.12E-03
rs62289125	3	185468884	*IGF2BP2*	T	C	1.096	7.36E-11	1.164	4.95E-01	1.054	7.99E-02
rs12636352	3	185469678	*IGF2BP2*	A	G	0.925	4.85E-09	0.924	1.23E-06	0.928	1.03E-03
rs6800985	3	185470700	*IGF2BP2*	A	T	0.925	5.16E-09	0.923	1.02E-06	0.928	1.28E-03
rs11925694	3	185475706	*IGF2BP2*	A	G	0.924	3.66E-09	0.920	3.92E-07	0.932	2.40E-03

rs6763887	3	185482618	*IGF2BP2*	T	C	0.926	6.15E-09	0.926	2.00E-06	0.925	7.52E-04
rs544704091	3	185484257	*IGF2BP2*	A	G	1.080	8.63E-09	1.083	1.27E-06	1.075	1.79E-03
rs555875344	3	23236109	*UBE2E2*	T	C	0.914	3.96E-08	0.907	5.93E-08	0.949	1.85E-01
rs6787641	3	23642413	*UBE2E2*	C	G	1.121	1.21E-08	1.127	1.31E-07	1.100	2.63E-02
rs9465851	6	20640537	*CDKAL1*	A	G	0.930	8.14E-09	0.925	3.21E-07	0.941	6.55E-03
rs542490126	6	20641900	*CDKAL1*	A	G	0.926	8.16E-10	0.924	2.29E-07	0.930	9.52E-04
rs6925328	6	20642785	*CDKAL1*	A	G	1.080	9.64E-10	1.083	2.22E-07	1.074	1.12E-03
rs2328531	6	20643752	*CDKAL1*	A	G	1.081	7.10E-10	1.083	1.91E-07	1.075	9.48E-04
rs1012635	6	20675295	*CDKAL1*	A	G	1.078	1.90E-09	1.085	1.15E-07	1.063	3.49E-03
rs118154473	6	20716253	*CDKAL1*	A	G	1.077	2.67E-08	1.094	6.86E-08	1.047	3.62E-02
rs530052774	6	20724250	*CDKAL1*	A	G	1.094	7.12E-09	1.108	6.03E-08	1.065	1.86E-02
rs150983983	6	20725097	*CDKAL1*	A	T	1.092	1.37E-08	1.107	7.47E-08	1.061	2.63E-02
rs180674385	6	20728513	*CDKAL1*	T	C	0.932	1.13E-08	0.939	3.34E-05	0.918	5.82E-05
rs367557649	7	15016108	*DGKB*	T	C	1.078	4.32E-08	1.070	3.51E-05	1.095	2.39E-04
rs17168570	7	15017023	*DGKB*	T	C	1.072	3.45E-08	1.067	2.82E-05	1.079	2.96E-04
rs111328900	7	15023729	*DGKB*	A	G	0.930	3.10E-08	0.942	1.26E-04	0.904	2.17E-05
rs17168594	7	15048776	*DGKB*	T	C	0.931	1.07E-08	0.931	2.77E-06	0.931	9.94E-04
rs6976381	7	15048814	*DGKB*	T	C	0.931	1.07E-08	0.931	2.89E-06	0.931	9.94E-04
rs6962498	7	15050305	*DGKB*	C	G	0.928	1.97E-09	0.929	1.16E-06	0.925	4.18E-04
rs12666046	7	15050381	*DGKB*	A	G	0.928	1.97E-09	0.929	1.14E-06	0.925	4.18E-04
rs4725953	7	15052860	*DGKB*	T	C	0.929	5.11E-09	0.929	1.11E-06	0.931	1.22E-03
rs73466540	7	15053307	*DGKB*	A	G	1.078	2.41E-09	1.079	7.49E-07	1.076	8.18E-04
rs10280166	7	15053485	*DGKB*	T	C	1.076	3.73E-09	1.076	1.52E-06	1.077	6.48E-04

rs10264723	7	15054370	*DGKB*	T	C	1.078	1.27E-09	1.078	1.04E-06	1.079	2.98E-04
rs58028761	7	15054908	*DGKB*	T	C	1.077	2.85E-09	1.077	1.37E-06	1.075	5.46E-04
rs6980437	7	15054916	*DGKB*	T	G	0.926	9.48E-10	0.928	7.70E-07	0.923	3.14E-04
rs2052129	7	15054935	*DGKB*	A	C	0.928	2.19E-09	0.928	1.33E-06	0.929	4.31E-04
rs2215382	7	15055010	*DGKB*	C	G	1.077	2.12E-09	1.077	1.29E-06	1.077	4.31E-04
rs12538043	7	15058998	*DGKB*	C	G	1.073	3.28E-08	1.086	2.28E-07	1.050	2.08E-02
rs11514706	7	15059272	*DGKB*	A	C	0.923	1.77E-10	0.924	2.33E-07	0.920	1.83E-04
rs7786873	7	15060410	*DGKB*	A	G	0.920	3.69E-11	0.921	6.75E-08	0.918	1.21E-04
rs142841100	7	15060429	*DGKB*	A	T	1.087	2.36E-11	1.086	1.03E-07	1.089	5.12E-05
rs2215383	7	15062983	*DGKB*	T	C	0.932	2.18E-08	0.925	7.16E-07	0.945	6.65E-03
rs80005921	7	15063379	*DGKB*	A	G	0.926	1.71E-09	0.933	7.17E-06	0.910	4.09E-05
rs567154098	7	15063430	*DGKB*	A	T	0.930	1.03E-08	0.924	5.00E-07	0.942	4.68E-03
rs75118469	7	15063569	*DGKB*	T	C	1.073	1.73E-08	1.082	5.34E-07	1.058	6.65E-03
rs4725982	7	15063833	*DGKB*	T	G	0.929	4.86E-09	0.920	1.32E-07	0.945	6.44E-03
rs10950550	7	15063923	*DGKB*	T	G	1.075	1.04E-08	1.082	5.11E-07	1.062	4.71E-03
rs147969035	7	15063926	*DGKB*	T	C	1.073	1.78E-08	1.082	5.73E-07	1.058	6.44E-03
rs10228796	7	15064190	*DGKB*	C	G	0.932	1.83E-08	0.925	5.84E-07	0.945	6.44E-03
rs10258074	7	15064216	*DGKB*	A	T	0.932	1.83E-08	0.925	5.85E-07	0.945	6.44E-03
rs369192064	7	15064896	*DGKB*	T	G	0.931	1.13E-08	0.924	4.54E-07	0.944	5.34E-03
rs2072411	7	15064984	*DGKB*	A	G	1.074	1.05E-08	1.082	4.89E-07	1.060	4.73E-03
rs75648145	7	15065003	*DGKB*	T	C	0.931	1.37E-08	0.925	6.69E-07	0.943	4.73E-03
rs6967891	7	15065074	*DGKB*	T	C	0.932	1.73E-08	0.925	6.59E-07	0.944	5.34E-03
rs576438427	7	15065467	*DGKB*	T	C	1.072	3.17E-08	1.079	1.58E-06	1.060	4.83E-03

rs576771830	7	15065564	*DGKB*	A	T	1.072	2.80E-08	1.079	1.56E-06	1.061	4.19E-03
rs1558318	7	15065612	*DGKB*	A	T	0.932	2.72E-08	0.926	1.60E-06	0.943	4.19E-03

Chr: chromosome, A1: reference allele, A2: alternative allele.

Table 6.9 Results of 40 of the 48 SNPs after performing a meta-analysis of the original Arab-Emirati GWAS summary statistics with South-Asian based GWAS summary statistics. Summary statistics of the South Asian-based GWAS was accessed via the GWAS catalogue (accession number: GCST90093110). SNPs that had a p-value < 10^{-5} in the meta-analysis are marked in green.

						Meta-analysis		SA		EM	
SNP	**Chr**	**Position**	**Nearest gene**	**A1**	**A2**	**OR**	**P**	**OR**	**P**	**OR**	**P**
rs12119289	1	48008522	*FOXD2*	A	G	0.955	7.24E-04	0.988	5.05E-01	1.39	4.49E-06
rs2861451	1	167312409	*POU2F1*	T	G	1.208	2.82E-06	1.305	3.68E-03	1.07	5.28E-07
rs60745626	2	1689209	*PXDN*	T	C	1.051	3.26E-04	1.021	2.80E-01	1.32	1.93E-06
rs4487114	2	23050710	*KLHL29*	A	G	0.967	2.04E-03	0.997	8.39E-01	1.37	1.40E-07
rs75875192	2	119217818	*INSIG2*	A	G	0.959	8.93E-02	1.052	1.67E-01	1.05	1.12E-06
rs150527739	3	14432140	*SLC6A6*	A	G	0.933	1.19E-02	0.980	4.72E-01	1.02	1.76E-06
rs144006677	3	51681325	*RAD54L2*	A	C	0.894	1.04E-02	0.992	8.77E-01	1.02	1.26E-06
rs11726869	4	8026645	*ABLIM2*	T	G	1.057	6.15E-04	1.015	4.57E-01	1.25	1.88E-06
rs139588925	4	18131933	*LCORL*	A	G	1.074	7.34E-02	0.961	4.06E-01	1.03	4.23E-06
rs511804	4	20545268	*SLIT2*	C	G	1.025	9.02E-02	0.980	3.19E-01	1.16	2.79E-06
rs2129645	4	141792539	*RNF150*	A	T	0.929	4.87E-05	0.966	1.66E-01	1.21	4.03E-06
rs17676033	5	172743694	*STC2*	T	C	1.077	8.76E-03	0.980	6.15E-01	1.06	2.63E-06
rs6456589	6	24092339	*NRSN1*	C	G	1.033	4.74E-03	1.000	9.79E-01	1.46	3.21E-07
rs6919765	6	77261591	*IMPG1*	A	G	1.051	1.10E-03	1.005	8.21E-01	1.28	8.86E-07

rs635551	6	160404832	*IGF2R*	T	C	0.931	6.82E-04	0.982	4.60E-01	1.12	2.81E-06
rs62448651	7	15066641	*DGKB*	A	C	1.070	7.53E-07	1.045	1.10E-02	1.55	1.88E-06
rs117089947	7	147549030	*CNTNAP2*	T	C	1.068	2.90E-02	0.954	2.64E-01	1.06	4.32E-06
rs111777155	9	99820736	*CTSV*	T	G	1.134	4.29E-04	1.069	9.37E-02	1.02	4.06E-06
rs117077124	9	124942205	*MORN5*	A	G	1.109	1.60E-04	1.052	1.29E-01	1.05	1.70E-06
rs11792868	9	130480322	*TTC16*	A	C	1.056	7.51E-03	1.004	8.60E-01	1.10	1.43E-06
rs10903783	10	2540623	*PFKP*	T	C	1.018	1.81E-01	0.978	1.59E-01	1.55	4.84E-06
rs10827154	10	19540391	*MALRD1*	A	G	0.966	1.98E-02	1.024	2.40E-01	1.24	1.60E-08
rs7903146	10	114758349	*TCF7L2*	T	C	1.232	8.11E-59	1.302	7.94E-59	1.53	4.83E-08
rs11047700	12	25061338	*BCAT1*	T	C	0.959	3.60E-02	1.008	7.46E-01	1.11	1.48E-06
rs7953156	12	44929687	*NELL2*	A	G	1.031	1.92E-02	0.998	8.83E-01	1.32	2.21E-06
rs17096794	12	46604091	*SLC38A1*	A	G	0.977	1.25E-01	1.032	1.67E-01	1.12	4.73E-06
rs17104138	12	68005294	*DYRK2*	A	G	1.091	4.59E-05	1.058	2.46E-02	1.07	2.71E-06
rs11107089	12	93941505	*SOCS2*	A	G	0.848	4.82E-04	0.960	5.33E-01	1.03	3.57E-07
rs2324604	13	40971758	*FOXO1*	T	C	0.960	1.49E-03	0.997	8.59E-01	1.44	4.16E-07
rs1031302	13	58038376	*PCDH17*	A	G	1.030	3.04E-02	0.996	8.17E-01	1.26	3.01E-06
rs9540874	13	67290606	*PCDH9*	A	T	1.020	2.94E-01	0.950	2.86E-02	1.15	3.95E-06
rs142668827	15	55114632	*UNC13C*	T	C	1.111	5.12E-04	0.974	6.66E-01	1.10	9.71E-06
rs16944568	15	62044136	*VPS13C*	A	G	1.035	3.98E-02	0.972	1.75E-01	1.18	1.03E-07
rs13333067	16	3495383	*NAA60*	T	C	1.084	8.93E-04	1.022	4.71E-01	1.07	3.27E-06
rs55870929	16	25647503	*HS3ST4*	T	C	0.875	8.11E-04	1.095	3.49E-01	1.06	2.74E-06
rs727836	16	27832033	*GSG1L*	T	C	0.936	5.15E-04	0.979	4.01E-01	1.16	3.06E-06
rs117027386	16	61969102	*CDH8*	A	G	0.877	1.61E-03	0.977	6.48E-01	1.02	1.62E-06

rs913548	20	56053525	*CTCFL*	A	T	1.039	2.53E-03	1.002	8.85E-01	1.37	3.98E-07
rs5759835	22	23816142	*FAM230I*	T	C	0.932	6.05E-04	0.968	1.74E-01	1.08	4.87E-06
rs77316551	22	35285864	*ISX*	T	C	0.845	2.73E-05	0.975	7.11E-01	1.05	7.87E-07

Chr: chromosome, RAF: risk allele frequency; A1: reference allele, A2: alternative allele, OR: fixed-effect OR estimate, P: fixed-effects meta-analysis p-value.

Chapter Seven: Discussion

In this final chapter, I discuss key findings from my PhD, strengths and limitations, as well as address how this study helps reduce the existing research gap. I then derive suggestions for future work that can potentially solve the main research problem: the ethnic disparity in T2D susceptibility.

7.1 The role of variation in adiposity in type 2 diabetes susceptibility

My secondary analyses of global datasets highlighted the differential impact of general obesity on T2D burden across populations at a global level. The primary case that formed the basis of this study's hypothesis relates to the comparison of obesity and diabetes rates and trends in the UAE over the last 3 to 4 decades, to that observed in the UK. Both countries have experienced an increasing upward trend of obesity prevalence rates between 1980 and 2014. Yet, the repercussions of this on T2D burden has been witnessed in the UAE, but not to the same extent in the UK. That has driven us to consider the possibility of other risk factors making a bigger contribution towards the increased risk of T2D in Arab-Emiratis, and other populations of the Gulf region.

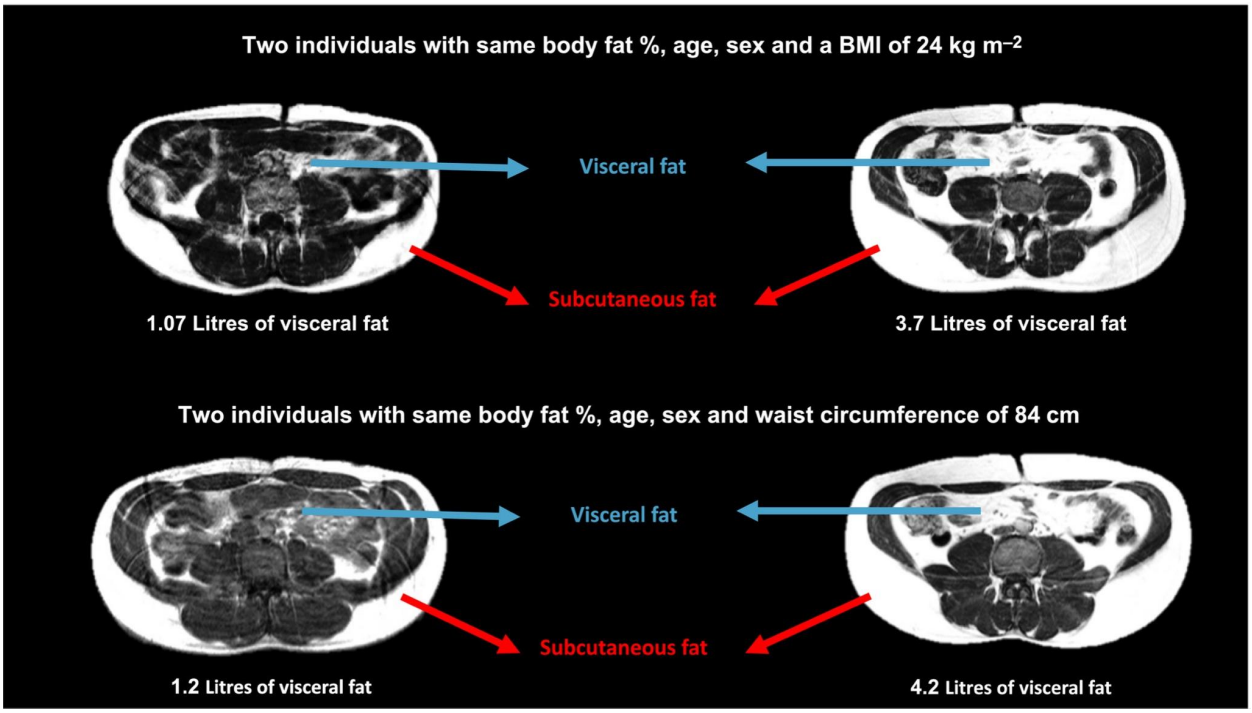
In the epidemiological studies I conducted during my PhD, I wanted to first validate this hypothesis by quantifying the excess risk of T2D in a representative sample of Arab-Emirati men and women and estimating the contribution of obesity towards this increased risk as compared to individuals in the UK of European descent. Consistent with publicly-available estimates, I found that overall T2D prevalence rates among Arab-Emiratis were ~3-4 fold higher than Europeans[205]. After considering differences in age and BMI between the two populations, Arab-Emiratis still have a striking 12 to 14-fold excess risk of T2D. I also demonstrate that BMI is an important risk factor for T2D in Arab-Emiratis as in Europeans, but

with weaker effect sizes; the association of obesity (defined as BMI ≥ 30 Kg/m^2) with T2D was 1.5- and 2.9-fold higher among European men and women compared to Arab-Emiratis, respectively. My own work also shows that the difference in prevalence of T2D was greatest among Arab-Emirati individuals with normal BMI, whereas prevalence of T2D was similar in magnitude (between Arab-Emiratis and Europeans) in individuals with Class I/II/III obesity.

Besides the drawbacks of a cross-sectional study design, failure of simple clinical measures of obesity, like BMI, to account for the excess risk of T2D among Arab-Emiratis compared to Europeans may relate to its inability to fully capture the complex distribution and spectrum of biological effects associated with adipose tissue. BMI cannot distinguish between fat and lean mass and is not indicative of body fatness; approximately 50% of overweight and 30% of obese individuals, based on BMI, are free from any obvious sign of metabolic and cardiovascular complications[369]. It is also unable to differentiate between different locations of stored fat within the body.

Like BMI, waist circumference and WHR measures have some limitations that primarily include their inability to distinguish between visceral and abdominal subcutaneous fat depots[101]. **Figure 7.1** was adapted from Yaghootkar H. et al. [101] and it shows transverse slices of four individuals from abdominal MRI measures. It illustrates how people with the same BMI or waist circumference, age, and gender can have very different levels of subcutaneous and visceral fat.

Figure 7.1 Transverse slices of 4 participants from abdominal MRI measures. Intensity-corrected fat image using overlay colours illustrates how two people with same BMI or waist circumference, same age and gender can have very different levels of subcutaneous and visceral fat and therefore, different disease risk. Adapted from Yaghootkar H (2020)[101].



Therefore, it is believed that accurate phenotyping of anthropometry and body fat distribution using medical imaging for precise body fat mapping and quantification is required to determine their potential contribution to ethnic differences in T2D risk[101]. Despite that, central adiposity as measured by WHR in this study, appears to highlight some possible reasons behind the increased risk of T2D in the Arab-Emirati population. Central adiposity rates were higher among Arab-Emiratis compared to Europeans, and at any given BMI, the association of central adiposity with T2D was also stronger in Arab-Emirati men and women. These findings could be indicative of numerous adiposity mechanisms/measures that increase disease susceptibility in this population. That is likely, particularly as Arab-Emiratis were also found to be characterised by increased fasting glucose, insulin, and insulin resistance, when compared

to other ethnic groups (Indian Asians and Europeans from the LOLIPOP study)[283]. In addition, the association of waist and anthropometry-derived VAT with insulin resistance (HOMA-IR) in Arab-Emirati men and women was stronger than that observed for BMI. That was not the case for South Asians and Europeans in the LOLIPOP study[283].

Individuals of African ancestry are known to have a greater risk of T2D, and higher prevalence of obesity and morbid obesity, but exhibit significantly lower levels of visceral and ectopic fat compared to Europeans[216]. This phenomenon is referred to as the African ectopic fat paradox[216] . Consistent with this phenomenon, in this study, men of African ancestry had similar BMI and obesity rates to Arab-Emiratis (higher than Europeans and South Asians). Yet, they had the lowest rates of central adiposity, and lowest waist circumference and WHR, when compared to all other ethnic groups studied. In contrast, women of African ancestry had the highest rates of central adiposity compared to all other ethnic groups (even Arab-Emiratis), had similar BMI to Arab-Emiratis, and similar waist and WHR to both Arab-Emiratis and South Asians.

Sex differences in adiposity measures were observed across all ethnic groups studied, and not only in individuals of African descent. In line with these observations, substantial differences in the metabolic responses to dietary challenges, weight gain, weight loss, as well as pharmacological interventions have been reported between males and females, although mechanisms underlying them remain largely unknown[370]. While males accumulate significantly more visceral fat, females predominantly store excess fat in their subcutaneous adipose tissue[370, 371], however such differences diminish as postmenopausal women have increased visceral fat accrual[372]. These observations highlight the role of gonadal steroids in this phenomenon, but given that clear differences are also observed even before the onset of puberty, sex hormones cannot account for all dissimilarities[370]. Other mechanisms involved include; (1)

increased expression of various genes involved in mitochondrial function, including uncoupling protein (UCP)-1, which can explain differences in metabolic rate per kilogram adipose tissue between sexes[373], (2) the protective effect of oestrogens through diminishing inflammatory signalling and improving insulin action, particularly in post-menopausal women[374], (3) differential gene expression related to insulin signalling and lipid synthesis, which are elevated in females regardless of dietary intake[375]. Further understanding these differences and mechanisms involved will undoubtedly lead to better treatment for both sexes, with some having varied efficaciousness in males and females.

The stronger association of central adiposity (as measured by WHR) with T2D observed in Arab-Emiratis in this study could relate to the 'adipose tissue expandability' hypothesis[101]. Under this hypothesis, individuals who are capable of sufficiently expanding their subcutaneous adipose tissue, will be protected to some extent against the adverse effects of increased adiposity, independent of considerable increase of their adipose tissue[376]. Effective subcutaneous fat storage is characterised by the enlargement of pre-existing adipocytes (hypertrophy) or by the generation of new small adipocytes (hyperplasia), or both[102]. Recent evidence from large-scale GWASs have provided a number of 'favourable adiposity' genetic variants, that increase overall adiposity (higher body fat % and BMI) yet protect against obesity-related metabolic complications[172, 218, 219]. These variants were found to also be associated with increased subcutaneous fat storage efficiency[218]. Compared to Europeans, South Asians were found to have fewer of these protective alleles[101], and these findings may partly explain the excess prevalence of T2D in South Asians. Expanding these investigations to populations like Arab-Emiratis would yield useful insights into the observed increased T2D risk.

Increased adipocyte volume is associated with higher fasting insulin, insulin resistance, and risk of T2D[103, 104]. In excessive nutrient states, adipocytes become enlarged and undergo endoplasmic reticulum

stress, which leads to an unfolded protein response[377]. The UPR consequently activates stress and inflammatory pathways and the production of cytokines and chemokines that interfere with the insulin signalling pathway[377]. The correlation between simple measures such as WHR/waist/BMI with adipocyte hypertrophy is unclear, and it is unlikely that they can capture this specific phenotype.

Ectopic fat accumulation in the liver has been associated with hepatic insulin resistance as well as with peripheral insulin resistance in skeletal muscle and adipose tissue[378, 379]. Non-alcoholic fatty liver diseases (NAFLD), which refers to a spectrum of conditions caused by the build-up of fat in the liver, has been closely associated with insulin resistance and T2D. Recent efforts have estimated NAFLD rates in the UAE population to be more than 25%[380]. This high burden in a relatively younger population, compared to Europeans with similar estimates[381], could potentially imply larger increases in NAFLD burden in the coming few decades[380]. Hepatic fat is not strongly correlated with BMI or WHR. In a study that included 83 severely obese individuals with no history of liver disease, T2D, or hepatoxic exposure, they found only moderate correlation of WHR with hepatic fat (R-squared = 0.4)[382]. Furthermore, hepatic triglyceride content was found to be the best predictor of liver insulin sensitivity, independent of body fat percentage and BMI[379].

Ethnic disparities in ectopic fat accumulation can partly explain differences in T2D risk across populations. South Asian populations for example, exhibit both greater levels of VAT and intrahepatic lipid compared to Europeans[211, 216] . Interestingly, African populations, despite being at greater risk of T2D, and having higher prevalence of obesity and morbid obesity, exhibit significantly lower levels of visceral and ectopic fat compared to Europeans[216]. In addition, the degree of ectopic pancreatic fat deposition varies between ethnic groups. This has been demonstrated in a study investigating ethnic differences in pancreatic triglyceride levels and its relationship with β-cell dysfunction, among

individuals of African-American, European, and Hispanic ancestry with mild obesity and no history of T2D[383]. They found increased levels of pancreatic triglycerides among Hispanics, an ethnic group with marked insulin resistance and high levels of obesity and T2D. To date, no studies have characterised ectopic fat (e.g., liver or pancreatic) and its contribution to insulin resistance and T2D burden in the Arab-Emirati population. Future studies using state of the art imaging modalities, to characterise hepatic and pancreatic fat among Arab-Emiratis will provide valuable insights into the contribution of ectopic fat deposition to T2D in this population.

7.1.1 Future work and implications for public health

Through studying simple clinical measures of adiposity, findings here highlighted the potential of fully characterising body fat distribution in the Arab-Emirati population in yielding useful information that could better explain their increased T2D risk. One key characteristic that would be important to investigate in this population is subcutaneous adipose tissue expandability. This includes investigating the cellular characteristics of adipose tissue in Arab-Emiratis as compared to other ethnic groups for example. Although a limited number of observational studies investigated this across ethnicities, they highlighted key aspects that may explain the increased metabolic risk among South Asians[212]. Furthermore, the discovery of favourable adiposity variants was originally made in populations of European descent and therefore findings may not be transferable to other ethnic groups. The allelic architecture of favourable adiposity in Arab-Emiratis, as well as in Africans, and Asians may differ, and demands future trans-ethnic investigations.

The limitations of BMI as a primary indicator of obesity have been well known for decades. Even though newer, more specific measures of adiposity-related risk have been developed[384], several guidelines[385, 386] continue to recommend BMI as the optimal tool to measure obesity-related morbidity and mortality

risk. That is also the case for the UAE[387]. On the other hand, despite the clear and convincing evidence of the major contribution of visceral and ectopic fat to metabolic risk well above BMI, their assessment in clinical practice remains a challenge. Therefore, the development of simple, clinically-applicable methods to identify individuals with excess visceral adiposity and ectopic fat, such as the one used in this study by Lin et al. [280] (for visceral adipose tissue volume estimation), would help in going beyond tracking and targeting BMI at the population level. Given the different age structure of the Arab-Emirati population compared to Europeans, and the stronger association of WHR-measured central adiposity with risk of T2D observed here, the use of equations derived from European-based studies to estimate abdominal VAT volume may not be optimal. Therefore, it would be crucial to build simple clinically-applicable tools that can accurately estimate adiposity measures, such as VAT volume, in this population. These tools can be used in general clinical practice as well as in public health programs. In addition, further investigations of the association of estimated adiposity measures and clinical outcomes will be required, as well as establishing the proper cut-off values for metabolic risk for the population.

7.2 The relationship of physical inactivity and sedentary time with T2D burden in the UAE

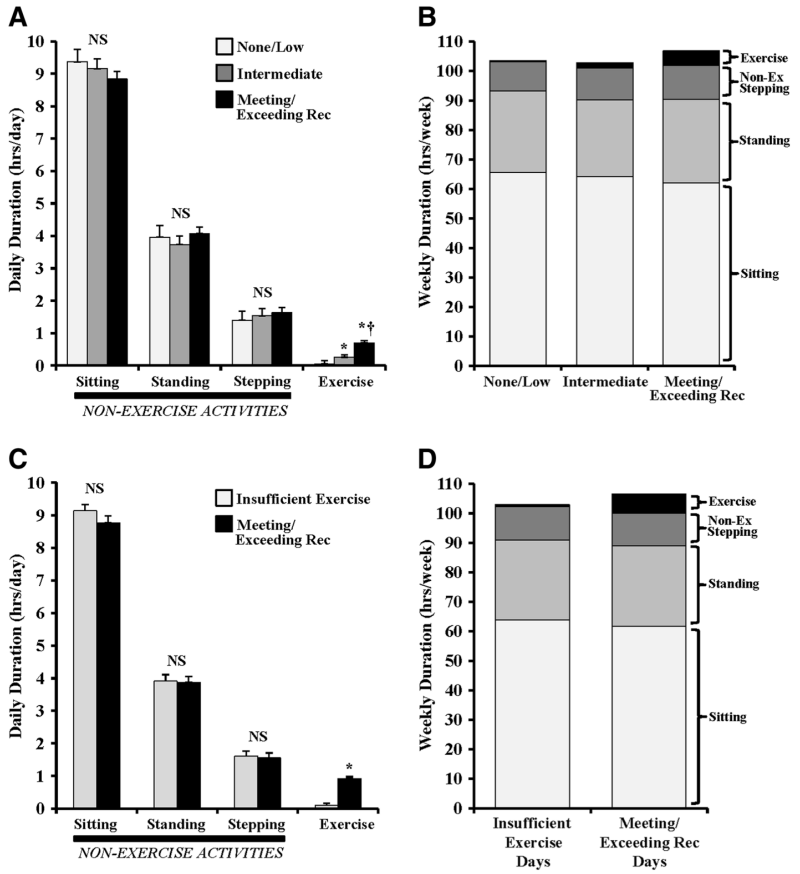
Consistent with previous work, my results highlight the low levels of physical activity among Arab-Emiratis[262] ; despite being a relatively younger population their activity levels are over 2-fold lower compared to Europeans. Physical activity has been repeatedly established as a key determinant of T2D among European[82, 126, 127]. Here, I replicate the protective role of physical activity among Europeans. Although a weak protective role of leisure time physical activity was observed among Arab-Emiratis, the association did not reach statistical significance. The opposite was also true for increased sedentary time and overall low physical activity levels. In addition, the risk of T2D that is attributable to lack of physical activity was relatively low across all ethnic groups, particularly among Arab-Emirati women. My sensitivity analyses showed that even after adjustment for differences in physical activity levels,

individuals of South Asian or African descent still had ~ 6-fold and ~3-fold excess of T2D compared to Europeans, respectively. These analyses could not be performed in Arab-Emiratis, as physical activity was not measured in the representative sample (EM1). Consistent with these results are findings, Sehmi et al.[283] showed that Indian Asians were 57% less likely than Europeans to participate in the minimum amount of leisure time physical activity. In addition, even amongst the most physically active, they found that Indian Asians have a striking ~6-fold excess risk of T2D.

Although there was a lack of direct association of physical inactivity with T2D, my work demonstrates a significant association of increased sedentary time and low physical activity levels with insulin resistance, decreased HDL and increased triglycerides among Arab-Emiratis, particularly in women. There is a growing area of physical activity research highlighting the physiological consequences of increased sedentary time on risk of T2D, insulin resistance, lipid metabolism, and glucose, independent of moderate-vigorous physical activity levels[323]. Indeed, the effect of 'too much sitting' (i.e., contractile inactivity) is not the same as too little exercise (i.e., defined type of physical activity as recommended by current guidelines)[388]. **Figure 7.2** adapted from Craft et al[388] illustrates that even individuals who exercised ~ 300 min/week (2X the recommended levels), their sitting time and total daily non-exercise activity were not different from those who did not exercise. Studies based on accelerometer measured sedentary time also highlighted the inverse relationship between breaks in sedentary time with waist circumference[389] and BMI[390]. Findings, although plausible, also support the possibility that frequent muscular contraction throughout the day may be necessary to short-circuit molecular signalling that could lead to metabolic disease[322]. Overall, however, existing evidence supporting total sedentary time rather than the pattern of how sedentary time is accumulated in relation to metabolic risk is stronger[323].

Interestingly, the association of low physical activity levels with increased central adiposity (waist circumference and anthropometry-derived VAT) was the strongest in Arab-Emiratis, compared to individuals of European, South Asian, or African descent. These findings reiterate the contribution of central adiposity towards the increased burden of T2D and insulin resistance in the Arab-Emirati population. Therefore, my findings do not refute the important association of physical inactivity with T2D risk in the Arab-Emirati population; there are several reasons and limitations that must be acknowledged when deriving conclusions from my observations, particularly concerning the small sample size and thus limited power to detect an effect size lower than 1.4.

Figure 7.2 Time spent sitting, standing, incidental stepping when not exercising (non-exercise stepping), and exercising as defined by the federal physical activity guidelines. Adapted from Craft L (2012) [388].



There is an increased exposure to physical inactivity across among Arab-Emiratis, and that was highlighted in this study. Yet, the impact of inactivity on the development of T2D might be related to adiposity (particularly central adiposity), other unmeasured adiposity, dietary factors, or patterns of physical activity throughout longer periods of time. The members of the Diabetes Prevention Program Research Group for example, showed that lifestyle intervention was particularly effective with consistent physical activity, with one case of diabetes prevented per seven persons treated for three years[88]. On the contrary, in this study, only current physical activity levels were estimated, with no knowledge of how active/inactive individuals had been in the last few years or through lifetime. In addition, the risk reduction associated with increased physical activity appears to be greatest in those at increased baseline risk, such as individuals with obesity, positive family history, and with impaired glucose metabolism[391]. Whereas individuals at low baseline risk of T2D, remain at low risk whether they are active or not[391]. Thus, an individual's baseline risk must be considered to quantify the amount of physical activity required to confer low risk of T2D. Stratified analyses on the protective role of physical activity on T2D risk among Arab-Emirati individuals with increased baseline risk would thus yield more helpful insight.

Errors with estimation of physical activity using self-reported questionnaires in epidemiological studies are well-documented[339, 340]. The use of improved technology such as accelerometers, to directly measure total energy expenditure, can refine our understanding of the protective mechanisms of physical activity towards T2D risk. This can be achieved through the availability of large amounts of detailed information, including body motion, posture changes, sleep patterns, heart rate and movement that accelerometers can measure/store[392]. The use of Actigraph monitors in a relatively small sample in this study showed no evidence to discredit my self-reported questionnaire derived observations. But

their application to large Arab-Emirati epidemiological studies will better capture the association of physical activity to risk of T2D in this population.

7.2.1 Future work and implications for public health

Given the unique cultural and climatic barriers for physical activity that Arab-Emiratis face on a daily basis[333], adhering to physical activity recommendations can be of great challenge. One of the most high-impact questions that is yet to be resolved is what behaviours will be most effective and practical to help increase physical activity levels in this population. Taken together, my results highlighted the potential positive impact of reduced sedentary time on T2D risk in the Arab-Emirati population. This raises important implications about why sedentary behaviour is implicated in disease, and which type of interventions should be evaluated in providing the optimal solution for this particular population, especially for people who cannot (e.g., elderly) or are unwilling to exercise (e.g., overweight or obese individuals) at much higher intensities.

Future public health approaches that emphasise on non-fatiguing and very safe types of physical activity in the lower end of the physical activity continuum, will not only replace dozens of hours of sedentary time each week, but will also be more achievable in a larger proportion of people, and thereby provide high impact at the population level. For example, workplace behavioural expectations can be targeted through: (1) introducing height adjustable workstations, (2) drink fountains and bathrooms at a further distance from workplace stations, (3) encouraging the use of stairs instead of elevators, and (4) organising walking groups at lunch time.

7.3 The environment and T2D risk

I have assessed the contribution of adiposity measures and physical inactivity to the increased risk of T2D amongst Arab-Emiratis compared to Europeans. However, to fully evaluate the role of the environment to T2D risk, other important factors need to be considered. Environmental influences differ between populations and within the same population, in terms of which and how much of the environmental risk factors they are exposed to, as well as how strong the effects are on risk of T2D[393]. An individual is exposed to and interacts with multiple layers of the environment: the physical environment (e.g., chemical exposures, light and dark cycles), the "built" environment (the part of the physical environment that is constructed or modified by human activity), the food environment, the psychosocial as well as the socio-economic environment. The emphasis on the built environment for example is necessary in public health, as encouraging people to be more physically active in an environment that is not supportive is less likely to be effective[394]. This is particularly true where the climate includes extremes of temperature. Thus, barriers to physical activity in the UAE must be considered[333]. In addition, the physical environment can include exposures to air and water pollution, which have been shown to interfere in the functioning of the endocrine system (endocrine disrupting chemicals)[395]. For example, oestrogen receptor agonists such as bisphenol A, which is derived from polycarbonate plastics for coatings and food containers, has been implicated in the development of obesity and insulin resistance[396]. Due to the exploitation of the UAE's natural resources, its rapid population growth and high energy demand, the country has relatively high levels of air pollution[397], but the contribution of this towards the increased burden of T2D is not yet investigated.

Environmental drivers of excessive eating also need to be considered. These include the presentation of large portions of food, the extensive food availability, social acceptance of frequent eating, and susceptibility to food marketing tools[393]. Although measured estimates of these factors have not been

documented for the UAE, with background knowledge of the people's cultural practice and beliefs, we can presume that such factors are relatively prevalent in this population. In addition, previous studies have highlighted the causal role of sleep disturbances in energy imbalance and insulin dysfunction. For example, obstructive sleep apnoea (OSA) alters glucose metabolism, promotes insulin resistance, and is associated with increased risk of T2D[398]. The estimated prevalence of OSA in the UAE is 23% in males and 19.5% in females, and that increases to 70% among individuals with obesity[399]. The contribution of OSA and other sleep-related factors toward obesity, eating behaviour and T2D amongst Arab-Emiratis merit further evaluation in the setting of large epidemiological study.

7.4 Genetic risk factors and T2D

In my genome-wide association analyses, I investigated the contribution of known and novel genetic variants towards the risk of T2D amongst Arab-Emiratis. My GWAS identified 48 loci to be significantly associated with T2D in Arab-Emiratis, 40 of which were novel observations. I replicated variants within four previously identified loci, including *TCF7L2*, *HNF1B*, *CDKAL1*, *PPARG*, and *DGKB*, and found that they have similar allele frequencies with overall lower effect sizes compared to Europeans or South Asians. These findings do not support a central role for previously-reported T2D genetic susceptibility factors underlying the increased risk of T2D observed in Arab-Emiratis compared to Europeans. Possible reasons for the failure to replicate many of the associations among Arab-Emiratis may relate to limitations of sample size and power, allelic heterogeneity and differences in patterns of linkage disequilibrium between known T2D variants and the actual causal variants at the identified loci in different ethnic populations. Alternatively, differences in exposure to certain environmental risk factors in different populations, which has been highlighted in this study too, may alter the impact of susceptibility loci on risk of T2D through gene-environmental interactions[393].

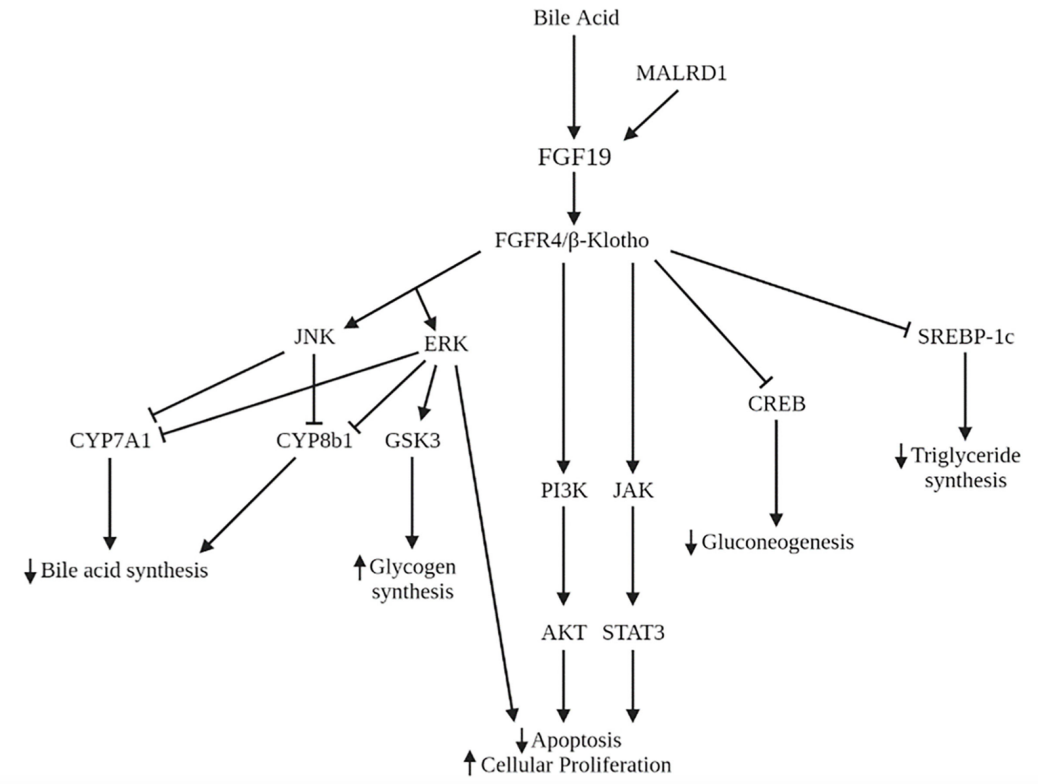
Evidence from a range of studies suggests that the development of obesity and T2D may result from a synergistic interaction between genes and the environment. For example, Pima Indians living in Arizona (USA), exposed to the Western living environment, were reported to have striking high rates of T2D in the early 2000s, with more than 50% of adults being affected in this population[400]. Segregation analyses demonstrated the strong familial aggregation of T2D in Pima Indian families, suggesting a genetic basis for the disease[401]. On the other hand, Pima Indians living in Northern Mexico, who are genetically similar to those living in Arizona, but rather retained a healthier and more traditional lifestyle, had much lower rates of obesity and T2D[402]. These observations imply a strong genetic predisposition to obesity and T2D making them exceptionally susceptible to westernised environments. However, clear-cut evidence supporting such hypotheses is lacking, largely due to lack of power, and sensitivity of gene-environmental interaction effects to confounding of correlated environmental exposures[393]. It is also possible that variants that have an impact on T2D risk through their interaction with behavioural factors may differ from variants known to predispose individuals to development of disease, irrespective of interaction with behavioural factors. More recent studies support this view. Analysis of ~340,000 individuals identified five variant-dietary trait pairs associated with HbA1c, with four of them not showing a genetic main effect and which, therefore, could not have been identified in a main effect GWAS for HbA1c[403].

In addition to the shared variants across ethnic groups, my study highlighted the association of known T2D or glycaemic traits loci with T2D in Arab-Emiratis, however signals within these loci were highlighted by different lead SNPs from those previously identified in Europeans. For example, variants within the *IGF2R* locus (e.g. rs622217) were previously identified to be associated with T2D in Europeans[148, 404], but this signal was highlighted by a different SNP in Arab-Emiratis (rs635551). *IGF2R* is a gene encoding insulin-like growth factor type 2 (IGF2) receptor. The IGF system and its binding proteins overall play

important roles in growth, development, and energy homeostasis. *IGF2R* has been implicated in diverse biological mechanisms, acting as an *IGF2* scavenger[405], including glucose homeostasis[406]. In addition, circulating soluble IGF2R was previously reported to be associated with degree of insulin resistance, as assessed by the HOMA model, independently of age and BMI[405].

Among the 40 novel loci identified in my GWAS, the strongest association was highlighted by rs10827154 in *MALRD1* (stronger than *TCF7L2*), a newly identified intestinal protein that modulates levels of fibroblast growth factor 19 (FGF19)[360]. Numerous studies have demonstrated the important role of FGF19 in maintaining the tight regulation of cyclical process of bile acid synthesis in the liver, transport, and reabsorption from the ileum[407]. Alterations in this homeostasis affects hepatic metabolic processes, resulting in inflammation and development of diseases that include cholestatic liver diseases, dyslipidaemia, and diabetes[360]. **Figure 7.3** demonstrates the signalling pathways of FGF19 and MALRD1. FGF19 acts as a post-prandial hormone that enhances glycogen synthesis and inhibits gluconeogenesis independent of insulin levels[360]. In addition, FGF19 activates the transcription of genes, which are necessary for fatty acid synthesis and inhibition of lipogenic enzyme synthesis[360].

Figure 7.3 Signalling pathways of FGF19 and MALRD1. Adapted from Wang L (2022)[360].



Based on the T2D lead SNPs I identified in this study, I generated an Arab-Emirati-specific polygenic risk score (PRS). Testing and comparing the efficacy of this Arab-Emirati-specific PRS with European- and South Asian-specific PRSs, in identifying individuals at risk of developing T2D in a separate Arab-Emirati sample would be essential. This will further highlight the importance of extending genome-wide association studies to populations in the Middle East and the Gulf region.

My trans-ethnic meta-analyses of Arab-Emirati and South Asian data reinforced the importance of extending the investigation of susceptibility to T2D to non-European populations. Some of the novel T2D risk loci, originally identified here in the Arab-Emirati GWAS, were characterised by overall stronger association with T2D in the trans-ethnic meta-analysis, compared to their South Asian-specific results. These findings suggest that some of these variants could be at a lower allele frequency in South Asian or

European populations, and there is potential for their replication/discovery through trans-ethnic meta-analysis in larger samples. Even though the Arab-Emirati sample in this study was relatively small (n = 991), I was able to identify additional SNPs in or near *UBE2E2* (Ubiquitin Conjugating Enzyme E2 E2), *IGF2BP2* (Insulin Like Growth Factor 2 MRNA Binding Protein 2), and *PPARG* (Peroxisome Proliferator Activated Receptor Gamma). These loci were previously reported to be associated with T2D risk, but the additional identified variants from my trans-ethnic meta-analyses were independent of these known lead SNPs (i.e., not in linkage disequilibrium).

To date, more than 400 genetic loci have been identified for association with risk of T2D, and in line with the findings in this study, the majority exhibit similar effect sizes and frequencies across ancestries. Altogether these variants explain approximately 20% of T2D heritability. This 'missing heritability' is thought to be primarily due to unidentified loci with additive genetic effects (i.e., the infinitesimal model stating that the heritability of disease underlies a large number of small-effect common variants across the entire allele frequency spectrum)[408]. Alternatively, it might be explained by a large number of large-effect rare variants (the rare allele model), or a combination of genotypic, environmental and epigenetic interactions (the broad sense heritability model)[409]. The identification of further common genetic variants in this study lends support to the view that much of the overall variance in T2D susceptibility can be attributed to the impact of many common causal variants, most of modest effects.

The rare allele model argues that population processes operative in the human lineage would be more likely to favour the existence of multiple rare variations contributing to disease. For example, common variations are likely to be older than rarer variants, but given that they reached appreciable frequencies in the populations, they are therefore not as likely to have been subjected to negative selection[410]. Nevertheless, evaluating the contribution of rare variants to complex diseases like T2D remains a major

challenge relating to the identification of enough copies of the rare alleles to permit informative statistical analysis[411].

The high consanguinity rates in populations like the UAE has likely enriched rare functional variants including complex disease risk alleles, which could be rare and undetectable in other populations, but are amplified by cultural factors among Arab-Emiratis to discoverable levels. This can provide novel insights into disease mechanisms that may be applicable to other populations, particularly in the region. For the Arab-Emirati population more specifically, this can have important implications for future efforts to understand genetic risk and facilitate the development of population-specific therapeutic interventions[247], including metabolic disorders like T2D. Whole genome sequencing endeavours, such as the Arab-Emirati genome program which catalogues genetic variation amongst Arab-Emiratis[412] will permit accurate imputation of both shared and Arab-Emirati specific variants, which will inform design of a genotyping microarray specific for this population. This will enhance the search for both rare and common genetic variants underlying susceptibility to T2D and other diseases in the UAE population.

In addition to genetic variant-explained T2D susceptibility, the role of epigenetic modifications on risk of T2D must be acknowledged, particularly when evaluating the 'missing heritability'. Epigenetic regulation refers to heritable changes in gene expression and phenotype that are not determined by changes in DNA sequence[190]. Previous birth cohort studies showed that maternal undernutrition, low birth-weight, and rapid postnatal child growth increase off-spring risk of developing T2D[413]. This risk is thought to be mediated through epigenetic regulations (i.e., may be trans-generational), and may contribute to age-dependent changes in key metabolic genes, and increased susceptibility to T2D. In South Asian populations for example, risk of T2D was found to be highest amongst individuals who had shorter mothers, parents with lower BMI, low birth weight, and thinness in infancy, independent of their adult

BMI[413]. In addition, in one of the very few prospective studies, Chambers et al.[414], conducted a trans-ethnic epigenome-wide association analysis, and identified five differentially methylated loci in Europeans and South Asians. Interestingly, the methylation levels at these loci differed between the two populations, suggesting that measurement of DNA methylation might help explain the increased risk of T2D among South Asians[414]. Future epigenome-wide association analyses among Arab-Emiratis are needed to investigate the contribution of epigenetic factors to the increased risk of T2D in this population.

7.4.1 Future work and implications for public health

Although the vast majority of people with T2D are of non-European ethnicity, our current knowledge of T2D genetic architecture remains Eurocentric. The expansion of GWASs to other ethnic groups have been very successful in providing insights into disease mechanisms and drug targets, although not effective in identifying mechanisms that make non-Europeans more susceptible to T2D. Large-scale GWASs and whole genome/exome sequencing in non-Europeans is expected to refine our current knowledge about population-specific disease mechanisms.

Given that T2D is a disease with a heterogeneous spectrum of metabolic disturbance, any ethnicity-specific contribution of genetic background to individual susceptibility, needs to be considered together with information on internal (e.g., microbiome), as well as the external environment (e.g., lifestyle). One way to achieve that as suggested by Yaghootkar et al.[100] is through designing a multi-omic approach that includes measuring (1) genetic variants and which pathways they are involved in (e.g., subcutaneous adipose tissue storage, β-cell function, insulin resistance), and (2) transcriptome-, proteome-, or metabolome-measured intermediate phenotypes. With this approach, any variation in the degree of

pathogenesis between ethnic groups will be highlighted by enrichment of different 'base colours', that may symbolise a combination of specific pathways, genetic variants, and/or metabolic biomarkers[100].

7.5 Strengths and limitations

To my knowledge, this is the first trans-ethnic study conducted to establish the contribution of environmental and genetic risk factors towards the increased risk of T2D in the Arab-Emirati population. This study adds to efforts aiming to reduce the potential for bias in the representation of the Arab-Emirati and other understudied populations worldwide, and eventually help attenuate healthcare disparities[415]. Although Arab-Emirati samples (EM1 and EM2) were recruited from a diabetes centre (limitations discussed below), EM1 was still a good representative sample of the UAE's age structure, as well as age-specific T2D and obesity rates. Therefore, controls in the EM2 sample population, who were recruited from the same source population as EM1 (patients attending ICLDC for general health screening), can also be considered a good representative of the general UAE population. Given the challenges of recruiting people free from diabetes from a diabetes centre like ICLDC, I was still capable of achieving a case: control ratio of 1:1.3, with similar male: female ratio (0.9:1), and minimal differences in age. Physical activity questionnaire data completeness was high, with over 98.9% of all individuals completing all 16 questions. In addition, UKBB data access has provided me with a huge opportunity to investigate the primary aim of this study, with big sample sizes for Europeans, as well as extending my analyses to other ethnic groups (South Asians and Afro-Caribbeans). Lastly, despite the lack of power in my GWAS study, the replication of several known T2D or glycaemic traits loci in this study give us some confidence on the novel observed associations. The Emirati Genome Project will provide a huge opportunity to validate the novel loci identified in the primary Arab-Emirati GWAS.

Findings that I present in this thesis, like any observational study, suffer from two major limitations. Firstly, both the Arab-Emirati sample populations and the UKBB cohort might be subject to 'healthy volunteer' selection bias; and secondly, neither of the cohorts are fully representative of their original populations. With that taken into account, prevalence rates of T2D, obesity, and physical inactivity observed here were closely comparable to those estimated in more representable samples[205, 284 262]. This has been discussed in the respective results chapters.

T2D cases, particularly those attending a diabetes centre like ICLDC on a regular basis, may have taken measures to reduce weight, modify dietary patterns or increase their physical activity levels following diagnosis, as recommended by their physicians or dieticians. Moreover, some anti-diabetic medications are associated with weight gain, others are weight-neutral, and some can help with weight loss[416]. For example, sulfonylurea use has been linked to significant weight gain, and conversely, metformin has been shown to reduce weight as compared to sulfonylureas and to also mitigate the weight gain seen with sulfonylureas and repaglinide[416]. In addition to that, all analyses were limited to specific adiposity measures and physical activity, with no data on some other primary lifestyle (e.g., dietary intake, sleep patterns), socioeconomic (e.g., education, employment, social support), or other cultural (e.g., hospitality, activity barriers), and climatic factors. We should also consider the fact that individuals from the UKBB are all UK residents, and thus, results may vary if South Asian and Afro-Caribbean individuals were residing in their home countries. Altogether, these factors may have confounded associations of BMI, waist circumference, and WHR with T2D observed in this study.

Arab-Emiratis living across the seven Emirates of the UAE share the same religion, language and cultural history. Nevertheless, the high rates of endogamy and consanguinity within the UAE population is expected to produce genetic homogeneity and partitioning of gene pools[248, 249]. Unlike many other

countries, where multi-ethnic immigration has increased mixing of disparate gene pools, the Arab-Emirati population observes tribal endogamy and marriage with non-Emiratis is usually discouraged[248, 249]. Recent work by Eliott et al. found that the tribal culture in the UAE has created a genetic structure in the population[244]. They found that the inter-Emirati genetic differences to be broadly reflective of the different geographical locations of the individual Arab-Emirati populations[244]. In addition to that, the genetic architecture of the Arab-Emirati population captures the occurrence of admixture events[239], which may have occurred thousands of years ago, possibly related to movement of people in the Middle East after major cultural transitions. This must be considered when interpreting the results I present here. Both EM1 and EM2 include individuals predominantly from Abu Dhabi or Al Ain (> 70%). Therefore, my findings are best applicable in the context of Arab-Emirati men and women living in Abu Dhabi or Al Ain, and relevance to particular Arab-Emirati tribes or those from other Emirates is unclear. Larger samples that are inclusive of individuals from all seven Emirates will be required to produce more representative results of the UAE population.

7.6 Conclusions

I have investigated the contribution of obesity, physical inactivity and genetic risk variants to the increased risk of T2D among Arab-Emiratis compared to Europeans. I demonstrated an around 13-fold higher risk of T2D in the Arab-Emirati population compared to Europeans, unexplained by differences in BMI. I highlighted the stronger association of central adiposity towards T2D risk among Arab-Emiratis, compared to Europeans, South Asians, and Afro-Caribbeans, that may explain a larger proportion of the differences in risk between populations, as compared to BMI. My study failed to capture a direct association of physical inactivity with T2D risk in Arab-Emiratis but shed some light on its impact on central adiposity and insulin resistance. Lastly, my genome-wide association analyses identified 40

potentially novel loci in Arab-Emiratis and suggested that known genetic variants do not account for the increased T2D risk observed in the UAE. The findings in this study urgently call for an international programme that addresses the factors underlying the increased T2D prevalence among Arab-Emiratis and other high-risk populations in the region.

Bibliography

1. Lin X, Xu Y, Pan X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Sci Rep*. Sep 8 2020;10(1):14790. doi:10.1038/s41598-020-71908-9

2. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*. Dec 14 2006;444(7121):840-6. doi:10.1038/nature05482

3. Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract*. Jan 2022;183:109119. doi:10.1016/j.diabres.2021.109119

4. Tuomi T, Groop LC, Zimmet PZ, Rowley MJ, Knowles W, Mackay IR. Antibodies to glutamic acid decarboxylase reveal latent autoimmune diabetes mellitus in adults with a non-insulin-dependent onset of disease. *Diabetes*. Feb 1993;42(2):359-62. doi:10.2337/diab.42.2.359

5. Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The Lancet Diabetes & Endocrinology*. 2018/05/01/ 2018;6(5):361-369. doi:https://doi.org/10.1016/S2213-8587(18)30051-2

6. Pigeyre M, Hess S, Gomez MF, et al. Validation of the classification for type 2 diabetes into five subgroups: a report from the ORIGIN trial. *Diabetologia*. Jan 2022;65(1):206-215. doi:10.1007/s00125-021-05567-4

7. Zou X, Zhou X, Zhu Z, Ji L. Novel subgroups of patients with adult-onset diabetes in Chinese and US populations. *The Lancet Diabetes & Endocrinology*. 2019/01/01/ 2019;7(1):9-11. doi:https://doi.org/10.1016/S2213-8587(18)30316-4

8. Anjana RM, Baskar V, Nair ATN, et al. Novel subgroups of type 2 diabetes and their association with microvascular outcomes in an Asian Indian population: a data-driven cluster analysis: the INSPIRED study. *BMJ Open Diabetes Research & Care*. 2020;8(1):e001506. doi:10.1136/bmjdrc-2020-001506

9. Tanabe H, Saito H, Kudo A, et al. Factors Associated with Risk of Diabetic Complications in Novel Cluster-Based Diabetes Subgroups: A Japanese Retrospective Cohort Study. *J Clin Med*. Jul 2 2020;9(7)doi:10.3390/jcm9072083

10. Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: a high-risk state for diabetes development. *Lancet*. Jun 16 2012;379(9833):2279-90. doi:10.1016/s0140-6736(12)60283-9

11. Danaei G, Finucane MM, Lu Y, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. *Lancet*. Jul 2 2011;378(9785):31-40. doi:10.1016/s0140-6736(11)60679-x

12. Bansal N. Prediabetes diagnosis and treatment: A review. *World J Diabetes*. Mar 15 2015;6(2):296-303. doi:10.4239/wjd.v6.i2.296

13. Dabelea D, Mayer-Davis EJ, Saydah S, et al. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *Jama*. May 7 2014;311(17):1778-86. doi:10.1001/jama.2014.3201

14. Gakidou E, Afshin A, Abajobir AA, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2017/09/16/ 2017;390(10100):1345-1422. doi:https://doi.org/10.1016/S0140-6736(17)32366-8

15. Yang JJ, Yu D, Wen W, et al. Association of Diabetes With All-Cause and Cause-Specific Mortality in Asia: A Pooled Analysis of More Than 1 Million Participants. *JAMA Network Open*. 2019;2(4):e192696-e192696. doi:10.1001/jamanetworkopen.2019.2696

16. Rao Kondapally Seshasai S, Kaptoge S, Thompson A, et al. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med*. Mar 3 2011;364(9):829-841. doi:10.1056/NEJMoa1008862

17. Yang JK, Feng Y, Yuan MY, et al. Plasma glucose levels and diabetes are independent predictors for mortality and morbidity in patients with SARS. *Diabet Med*. Jun 2006;23(6):623-8. doi:10.1111/j.1464-5491.2006.01861.x

18. Guan W-j, Ni Z-y, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*. 2020;382(18):1708-1720. doi:10.1056/NEJMoa2002032

19. Chen H, Chen G, Zheng X, Guo Y. Contribution of specific diseases and injuries to changes in health adjusted life expectancy in 187 countries from 1990 to 2013: retrospective observational study. *Bmj*. Mar 27 2019;364:l969. doi:10.1136/bmj.l969

20. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol*. Feb 2018;14(2):88-98. doi:10.1038/nrendo.2017.151

21. Association AD. Standards of medical care in diabetes--2013. *Diabetes Care*. Jan 2013;36 Suppl 1(Suppl 1):S11-66. doi:10.2337/dc13-S011

22. Kotwas A, Karakiewicz B, Zabielska P, Wieder-Huszla S, Jurczak A. Epidemiological factors for type 2 diabetes mellitus: evidence from the Global Burden of Disease. *Arch Public Health*. Jun 22 2021;79(1):110. doi:10.1186/s13690-021-00632-1

23. Galicia-Garcia U, Benito-Vicente A, Jebari S, et al. Pathophysiology of Type 2 Diabetes Mellitus. *Int J Mol Sci*. Aug 30 2020;21(17)doi:10.3390/ijms21176275

24. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*. Apr 9-15 2005;365(9467):1333-46. doi:10.1016/s0140-6736(05)61032-x

25. Czech MP. Insulin action and resistance in obesity and type 2 diabetes. *Nat Med*. Jul 11 2017;23(7):804-814. doi:10.1038/nm.4350

26. Pearson T, Wattis JA, King JR, MacDonald IA, Mazzatti DJ. The Effects of Insulin Resistance on Individual Tissues: An Application of a Mathematical Model of Metabolism in Humans. *Bull Math Biol*. Jun 2016;78(6):1189-217. doi:10.1007/s11538-016-0181-1

27. Mu W, Cheng XF, Liu Y, et al. Potential Nexus of Non-alcoholic Fatty Liver Disease and Type 2 Diabetes Mellitus: Insulin Resistance Between Hepatic and Peripheral Tissues. *Front Pharmacol*. 2018;9:1566. doi:10.3389/fphar.2018.01566

28. Satoh T. Molecular mechanisms for the regulation of insulin-stimulated glucose uptake by small guanosine triphosphatases in skeletal muscle and adipocytes. *Int J Mol Sci*. Oct 16 2014;15(10):18677-92. doi:10.3390/ijms151018677

29. DeFronzo RA. Lilly lecture 1987. The triumvirate: beta-cell, muscle, liver. A collusion responsible for NIDDM. *Diabetes*. Jun 1988;37(6):667-87. doi:10.2337/diab.37.6.667

30. Venkatasamy VV, Pericherla S, Manthuruthil S, Mishra S, Hanno R. Effect of Physical activity on Insulin Resistance, Inflammation and Oxidative Stress in Diabetes Mellitus. *J Clin Diagn Res*. Aug 2013;7(8):1764-6. doi:10.7860/jcdr/2013/6518.3306

31. Kershaw EE, Flier JS. Adipose tissue as an endocrine organ. *J Clin Endocrinol Metab*. Jun 2004;89(6):2548-56. doi:10.1210/jc.2004-0395

32. Scherer PE. Adipose tissue: from lipid storage compartment to endocrine organ. *Diabetes*. Jun 2006;55(6):1537-45. doi:10.2337/db06-0263

33. Brehm A, Krssak M, Schmid AI, Nowotny P, Waldhäusl W, Roden M. Increased lipid availability impairs insulin-stimulated ATP synthesis in human skeletal muscle. *Diabetes*. Jan 2006;55(1):136-40.

34. Scherer PE. The Multifaceted Roles of Adipose Tissue-Therapeutic Targets for Diabetes and Beyond: The 2015 Banting Lecture. *Diabetes*. Jun 2016;65(6):1452-61. doi:10.2337/db16-0339

35. Ravussin Y, Leibel RL, Ferrante AW, Jr. A missing link in body weight homeostasis: the catabolic signal of the overfed state. *Cell Metab*. Oct 7 2014;20(4):565-72. doi:10.1016/j.cmet.2014.09.002

36. Saponaro C, Gaggini M, Carli F, Gastaldelli A. The Subtle Balance between Lipolysis and Lipogenesis: A Critical Point in Metabolic Homeostasis. *Nutrients*. Nov 13 2015;7(11):9453-74. doi:10.3390/nu7115475

37. Boden G. Obesity and free fatty acids. *Endocrinol Metab Clin North Am*. Sep 2008;37(3):635-46, viii-ix. doi:10.1016/j.ecl.2008.06.007

38. Roden M, Price TB, Perseghin G, et al. Mechanism of free fatty acid-induced insulin resistance in humans. *J Clin Invest*. Jun 15 1996;97(12):2859-65. doi:10.1172/jci118742

39. Belfort R, Mandarino L, Kashyap S, et al. Dose-response effect of elevated plasma free fatty acid on insulin signaling. *Diabetes*. Jun 2005;54(6):1640-8. doi:10.2337/diabetes.54.6.1640

40. Cherrington AD, Moore MC, Sindelar DK, Edgerton DS. Insulin action on the liver in vivo. *Biochem Soc Trans*. Nov 2007;35(Pt 5):1171-4. doi:10.1042/bst0351171

41. Santoleri D, Titchenell PM. Resolving the Paradox of Hepatic Insulin Resistance. *Cell Mol Gastroenterol Hepatol*. 2019;7(2):447-456. doi:10.1016/j.jcmgh.2018.10.016

42. Prentki M, Nolan CJ. Islet beta cell failure in type 2 diabetes. *J Clin Invest*. Jul 2006;116(7):1802-12. doi:10.1172/jci29103

43. Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC. Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. *Diabetes*. Jan 2003;52(1):102-10. doi:10.2337/diabetes.52.1.102

44. Christensen AA, Gannon M. The Beta Cell in Type 2 Diabetes. *Curr Diab Rep*. Aug 9 2019;19(9):81. doi:10.1007/s11892-019-1196-4

45. Halban PA, Polonsky KS, Bowden DW, et al. β-cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *Diabetes Care*. Jun 2014;37(6):1751-8. doi:10.2337/dc14-0396

46. Cerf ME. Beta cell dysfunction and insulin resistance. *Front Endocrinol (Lausanne)*. 2013;4:37. doi:10.3389/fendo.2013.00037

47. 16. UKpds. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. Prospective Diabetes Study Group. *Diabetes*. Nov 1995;44(11):1249-58.

48. Ladwa M, Hakim O, Amiel SA, Goff LM. A Systematic Review of Beta Cell Function in Adults of Black African Ethnicity. *J Diabetes Res*. 2019;2019:7891359. doi:10.1155/2019/7891359

49. Ladwa M, Bello O, Hakim O, et al. Ethnic differences in beta cell function occur independently of insulin sensitivity and pancreatic fat in black and white men. *BMJ Open Diabetes Res Care*. Mar 2021;9(1)doi:10.1136/bmjdrc-2020-002034

50. McCarthy MI. Painting a new picture of personalised medicine for diabetes. *Diabetologia*. May 2017;60(5):793-799. doi:10.1007/s00125-017-4210-x

51. Beulens JWJ, Pinho MGM, Abreu TC, et al. Environmental risk factors of type 2 diabetes-an exposome approach. *Diabetologia*. Feb 2022;65(2):263-274. doi:10.1007/s00125-021-05618-w

52. Saavedra JM. Obesity--a risk factor or a disease: What can exercise do for obese children? *Indian J Med Res*. May 2014;139(5):661-2.

53. Organisation WH. Obesity and overweight. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

54. Ofei F. Obesity - a preventable disease. *Ghana Med J*. Sep 2005;39(3):98-101.

55. Jehan S, Zizi F, Pandi-Perumal SR, et al. Obstructive Sleep Apnea and Obesity: Implications for Public Health. *Sleep Med Disord*. 2017;1(4)

56. Emerenziani S, Rescio MP, Guarino MP, Cicala M. Gastro-esophageal reflux disease and obesity, where is the link? *World J Gastroenterol*. Oct 21 2013;19(39):6536-9. doi:10.3748/wjg.v19.i39.6536

57. Fuselier A, Hanberry J, Margaret Lovin J, Gomelsky A. Obesity and Stress Urinary Incontinence: Impact on Pathophysiology and Treatment. *Curr Urol Rep*. Feb 22 2018;19(1):10. doi:10.1007/s11934-018-0762-7

58. Wells JC. Sexual dimorphism of body composition. *Best Pract Res Clin Endocrinol Metab*. Sep 2007;21(3):415-30. doi:10.1016/j.beem.2007.04.007

59. Horber FF, Zürcher RM, Herren H, Crivelli MA, Robotti G, Frey FJ. Altered body fat distribution in patients with glucocorticoid treatment and in patients on long-term dialysis. *Am J Clin Nutr*. May 1986;43(5):758-69. doi:10.1093/ajcn/43.5.758

60. Malis C, Rasmussen EL, Poulsen P, et al. Total and regional fat distribution is strongly influenced by genetic factors in young and elderly twins. *Obes Res*. Dec 2005;13(12):2139-45. doi:10.1038/oby.2005.265

61. Hilton C, Karpe F, Pinnick KE. Role of developmental transcription factors in white, brown and beige adipose tissues. *Biochim Biophys Acta*. May 2015;1851(5):686-96. doi:10.1016/j.bbalip.2015.02.003

62. Deurenberg P, Deurenberg Yap M, Wang J, Lin FP, Schmidt G. The impact of body build on the relationship between body mass index and percent body fat. *Int J Obes Relat Metab Disord*. May 1999;23(5):537-42. doi:10.1038/sj.ijo.0800868

63. Gurunathan U, Myles PS. Limitations of body mass index as an obesity measure of perioperative risk. *Br J Anaesth*. Mar 2016;116(3):319-21. doi:10.1093/bja/aev541

64. Chan JM, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Obesity, Fat Distribution, and Weight Gain as Risk Factors for Clinical Diabetes in Men. *Diabetes Care*. 1994;17(9):961-969. doi:10.2337/diacare.17.9.961

65. Stefan N, Häring HU, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol*. Oct 2013;1(2):152-62. doi:10.1016/s2213-8587(13)70062-7

66. Chen DL, Liess C, Poljak A, et al. Phenotypic Characterization of Insulin-Resistant and Insulin-Sensitive Obesity. *J Clin Endocrinol Metab*. Nov 2015;100(11):4082-91. doi:10.1210/jc.2015-2712

67. Kartheuser AH, Leonard DF, Penninckx F, et al. Waist circumference and waist/hip ratio are better predictive risk factors for mortality and morbidity after colorectal surgery than body mass index and body surface area. *Ann Surg*. Nov 2013;258(5):722-30. doi:10.1097/SLA.0b013e3182a6605a

68. Bays H. Central obesity as a clinical marker of adiposopathy; increased visceral adiposity as a surrogate marker for global fat dysfunction. *Curr Opin Endocrinol Diabetes Obes*. Oct 2014;21(5):345-51. doi:10.1097/med.0000000000000093

69. Pouliot MC, Després JP, Lemieux S, et al. Waist circumference and abdominal sagittal diameter: best simple anthropometric indexes of abdominal visceral adipose tissue accumulation and related cardiovascular risk in men and women. *Am J Cardiol*. Mar 1 1994;73(7):460-8. doi:10.1016/0002-9149(94)90676-9

70. van der Kooy K, Seidell JC. Techniques for the measurement of visceral fat: a practical guide. *Int J Obes Relat Metab Disord*. Apr 1993;17(4):187-96.

71. Shuster A, Patlas M, Pinthus JH, Mourtzakis M. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. *Br J Radiol*. Jan 2012;85(1009):1-10. doi:10.1259/bjr/38447238

72. Onat A, Avci GS, Barlan MM, Uyarel H, Uzunlar B, Sansoy V. Measures of abdominal obesity assessed for visceral adiposity and relation to coronary risk. *Int J Obes Relat Metab Disord*. Aug 2004;28(8):1018-25. doi:10.1038/sj.ijo.0802695

73. Ozhan H, Alemdar R, Caglar O, et al. Performance of bioelectrical impedance analysis in the diagnosis of metabolic syndrome. *J Investig Med*. Mar 2012;60(3):587-91. doi:10.2310/JIM.0b013e318244e2d9

74. Wang H, Chen YE, Eitzman DT. Imaging body fat: techniques and cardiometabolic implications. *Arterioscler Thromb Vasc Biol*. Oct 2014;34(10):2217-23. doi:10.1161/atvbaha.114.303036

75. Steinberger J, Daniels SR. Obesity, insulin resistance, diabetes, and cardiovascular risk in children: an American Heart Association scientific statement from the Atherosclerosis, Hypertension, and Obesity in the Young Committee (Council on Cardiovascular Disease in the Young) and the Diabetes Committee (Council on Nutrition, Physical Activity, and Metabolism). *Circulation*. Mar 18 2003;107(10):1448-53. doi:10.1161/01.cir.0000060923.07573.f2

76. Halpern A, Mancini MC. Diabesity: are weight loss medications effective? *Treat Endocrinol*. 2005;4(2):65-74. doi:10.2165/00024677-200504020-00001

77. Sattar N, Wannamethee SG, Forouhi NG. Novel biochemical risk factors for type 2 diabetes: pathogenic insights or prediction possibilities? *Diabetologia*. Jun 2008;51(6):926-40. doi:10.1007/s00125-008-0954-7

78. Sattar N, Gill JM. Type 2 diabetes as a disease of ectopic fat? *BMC Med*. Aug 26 2014;12:123. doi:10.1186/s12916-014-0123-4

79. Shaten BJ, Smith GD, Kuller LH, Neaton JD. Risk factors for the development of type II diabetes among men enrolled in the usual care group of the Multiple Risk Factor Intervention Trial. *Diabetes Care*. Oct 1993;16(10):1331-9. doi:10.2337/diacare.16.10.1331

80. Cassano PA, Rosner B, Vokonas PS, Weiss ST. Obesity and body fat distribution in relation to the incidence of non-insulin-dependent diabetes mellitus. A prospective cohort study of men in the normative aging study. *Am J Epidemiol*. Dec 15 1992;136(12):1474-86. doi:10.1093/oxfordjournals.aje.a116468

81. Colditz GA, Willett WC, Stampfer MJ, et al. Weight as a risk factor for clinical diabetes in women. *Am J Epidemiol*. Sep 1990;132(3):501-13. doi:10.1093/oxfordjournals.aje.a115686

82. Hu FB, Manson JE, Stampfer MJ, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *N Engl J Med*. Sep 13 2001;345(11):790-7. doi:10.1056/NEJMoa010492

83. Carey VJ, Walters EE, Colditz GA, et al. Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women. The Nurses' Health Study. *Am J Epidemiol*. Apr 1 1997;145(7):614-9. doi:10.1093/oxfordjournals.aje.a009158

84. Eriksson KF, Lindgärde F. Prevention of type 2 (non-insulin-dependent) diabetes mellitus by diet and physical exercise. The 6-year Malmö feasibility study. *Diabetologia*. Dec 1991;34(12):891-8. doi:10.1007/bf00400196

85. Moore LL, Visioni AJ, Wilson PW, D'Agostino RB, Finkle WD, Ellison RC. Can sustained weight loss in overweight individuals reduce the risk of diabetes mellitus? *Epidemiology*. May 2000;11(3):269-73. doi:10.1097/00001648-200005000-00007

86. Resnick HE, Valsania P, Halter JB, Lin X. Relation of weight gain and weight loss on subsequent diabetes risk in overweight adults. *J Epidemiol Community Health*. Aug 2000;54(8):596-602. doi:10.1136/jech.54.8.596

87. Tuomilehto J, Lindström J, Eriksson JG, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med*. May 3 2001;344(18):1343-50. doi:10.1056/nejm200105033441801

88. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. Feb 7 2002;346(6):393-403. doi:10.1056/NEJMoa012512

89. Anderson JW, Konz EC, Frederich RC, Wood CL. Long-term weight-loss maintenance: a meta-analysis of US studies. *Am J Clin Nutr*. Nov 2001;74(5):579-84. doi:10.1093/ajcn/74.5.579

90. Jung UJ, Choi MS. Obesity and its metabolic complications: the role of adipokines and the relationship between obesity, inflammation, insulin resistance, dyslipidemia and nonalcoholic fatty liver disease. *Int J Mol Sci*. Apr 11 2014;15(4):6184-223. doi:10.3390/ijms15046184

91. López-Domènech S, Abad-Jiménez Z, Iannantuoni F, et al. Moderate weight loss attenuates chronic endoplasmic reticulum stress and mitochondrial dysfunction in human obesity. *Mol Metab*. Jan 2019;19:24-33. doi:10.1016/j.molmet.2018.10.005

92. Thomas EL, Fitzpatrick JA, Malik SJ, Taylor-Robinson SD, Bell JD. Whole body fat: content and distribution. *Prog Nucl Magn Reson Spectrosc*. Aug 2013;73:56-80. doi:10.1016/j.pnmrs.2013.04.001

93. Marcadenti A, de Abreu-Silva EO. Different adipose tissue depots: Metabolic implications and effects of surgical removal. *Endocrinol Nutr*. Nov 2015;62(9):458-64. doi:10.1016/j.endonu.2015.05.010

94. Fox CS, Massaro JM, Hoffmann U, et al. Abdominal visceral and subcutaneous adipose tissue compartments: association with metabolic risk factors in the Framingham Heart Study. *Circulation*. Jul 3 2007;116(1):39-48. doi:10.1161/circulationaha.106.675355

95. Nicklas BJ, Penninx BW, Cesari M, et al. Association of visceral adipose tissue with incident myocardial infarction in older men and women: the Health, Aging and Body Composition Study. *Am J Epidemiol*. Oct 15 2004;160(8):741-9. doi:10.1093/aje/kwh281

96. Cnop M, Landchild MJ, Vidal J, et al. The concurrent accumulation of intra-abdominal and subcutaneous fat explains the association between insulin resistance and plasma leptin concentrations : distinct metabolic effects of two fat compartments. *Diabetes*. Apr 2002;51(4):1005-15. doi:10.2337/diabetes.51.4.1005

97. Pouliot MC, Després JP, Nadeau A, et al. Visceral obesity in men. Associations with glucose tolerance, plasma insulin, and lipoprotein levels. *Diabetes*. Jul 1992;41(7):826-34. doi:10.2337/diab.41.7.826

98. Cusi K. The role of adipose tissue and lipotoxicity in the pathogenesis of type 2 diabetes. *Curr Diab Rep*. Aug 2010;10(4):306-15. doi:10.1007/s11892-010-0122-6

99. Huang-Doran I, Sleigh A, Rochford JJ, O'Rahilly S, Savage DB. Lipodystrophy: metabolic insights from a rare disorder. *J Endocrinol*. Dec 2010;207(3):245-55. doi:10.1677/joe-10-0272

100. Stefan N, Kantartzis K, Machann J, et al. Identification and characterization of metabolically benign obesity in humans. *Arch Intern Med*. Aug 11 2008;168(15):1609-16. doi:10.1001/archinte.168.15.1609

101. Yaghootkar H, Whitcher B, Bell JD, Thomas EL. Ethnic differences in adiposity and diabetes risk - insights from genetic studies. *J Intern Med*. Sep 2020;288(3):271-283. doi:10.1111/joim.13082

102. Arner E, Westermark PO, Spalding KL, et al. Adipocyte turnover: relevance to human adipose tissue morphology. *Diabetes*. Jan 2010;59(1):105-9. doi:10.2337/db09-0942

103. Lundgren M, Svensson M, Lindmark S, Renström F, Ruge T, Eriksson JW. Fat cell enlargement is an independent marker of insulin resistance and 'hyperleptinaemia'. *Diabetologia*. Mar 2007;50(3):625-33. doi:10.1007/s00125-006-0572-1

104. Weyer C, Foley JE, Bogardus C, Tataranni PA, Pratley RE. Enlarged subcutaneous abdominal adipocyte size, but not obesity itself, predicts type II diabetes independent of insulin resistance. *Diabetologia*. Dec 2000;43(12):1498-506. doi:10.1007/s001250051560

105. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *Lancet Glob Health*. Oct 2018;6(10):e1077-e1086. doi:10.1016/s2214-109x(18)30357-7

106. Physical Activity Guidelines Advisory Committee report, 2008. To the Secretary of Health and Human Services. Part A: executive summary. *Nutr Rev*. Feb 2009;67(2):114-20. doi:10.1111/j.1753-4887.2008.00136.x

107. WHO. Global action plan for the prevention and control of noncommunicable diseases 2013–2020. *World Health Organization, Geneva2013*. 2013;

108. Westerterp KR. Physical activity and physical activity induced energy expenditure in humans: measurement, determinants, and effects. *Front Physiol*. 2013;4:90. doi:10.3389/fphys.2013.00090

109. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep*. Mar-Apr 1985;100(2):126-31.

110. WHO. Global recommendations on physical activity for health.
World Health Organization, Geneva2010. 2010;

111. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. Practical guide to measuring physical activity. *J Acad Nutr Diet*. Feb 2014;114(2):199-208. doi:10.1016/j.jand.2013.09.018

112. Melanson EL, Jr., Freedson PS. Physical activity assessment: a review of methods. *Crit Rev Food Sci Nutr*. May 1996;36(5):385-96. doi:10.1080/10408399609527732

113. Castillo-Retamal M, Hinckson EA. Measuring physical activity and sedentary behaviour at work: a review. *Work*. 2011;40(4):345-57. doi:10.3233/wor-2011-1246

114. Westerterp KR. Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol*. Apr 2009;105(6):823-8. doi:10.1007/s00421-009-1000-2

115. Jacobs DR, Jr., Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc*. Jan 1993;25(1):81-91. doi:10.1249/00005768-199301000-00012

116. Klesges RC, Eck LH, Mellon MW, Fulliton W, Somes GW, Hanson CL. The accuracy of self-reports of physical activity. *Med Sci Sports Exerc*. Oct 1990;22(5):690-7. doi:10.1249/00005768-199010000-00022

117. Uitenbroek DG. Seasonal variation in leisure time physical activity. *Med Sci Sports Exerc*. Jun 1993;25(6):755-60.

118. Chen KY, Bassett DR, Jr. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*. Nov 2005;37(11 Suppl):S490-500. doi:10.1249/01.mss.0000185571.49104.82

119. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc*. Nov 2005;37(11 Suppl):S523-30. doi:10.1249/01.mss.0000185658.28284.ba

120. CDC. General Physical Activities Defined by Level of Intensity.

121. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clin Cardiol*. Aug 1990;13(8):555-65. doi:10.1002/clc.4960130809

122. Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc*. Aug 2011;43(8):1575-81. doi:10.1249/MSS.0b013e31821ece12

123. Haskell WL, Lee IM, Pate RR, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc*. Aug 2007;39(8):1423-34. doi:10.1249/mss.0b013e3180616b27

124. Helmrich SP, Ragland DR, Leung RW, Paffenbarger RS, Jr. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *N Engl J Med*. Jul 18 1991;325(3):147-52. doi:10.1056/nejm199107183250302

125. Laaksonen DE, Lindström J, Lakka TA, et al. Physical activity in the prevention of type 2 diabetes: the Finnish diabetes prevention study. *Diabetes*. Jan 2005;54(1):158-65. doi:10.2337/diabetes.54.1.158

126. Hu FB, Sigal RJ, Rich-Edwards JW, et al. Walking compared with vigorous physical activity and risk of type 2 diabetes in women: a prospective study. *Jama*. Oct 20 1999;282(15):1433-9. doi:10.1001/jama.282.15.1433

127. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med*. Jun 25 2001;161(12):1542-8. doi:10.1001/archinte.161.12.1542

128. Colberg SR, Sigal RJ, Yardley JE, et al. Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association. *Diabetes Care*. Nov 2016;39(11):2065-2079. doi:10.2337/dc16-1728

129. Owen N, Sugiyama T, Eakin EE, Gardiner PA, Tremblay MS, Sallis JF. Adults' sedentary behavior determinants and interventions. *Am J Prev Med*. Aug 2011;41(2):189-96. doi:10.1016/j.amepre.2011.05.013

130. Fritschi C, Park H, Richardson A, et al. Association Between Daily Time Spent in Sedentary Behavior and Duration of Hyperglycemia in Type 2 Diabetes. *Biol Res Nurs*. Mar 2016;18(2):160-6. doi:10.1177/1099800415600065

131. Dunstan DW, Salmon J, Healy GN, et al. Association of television viewing with fasting and 2-h postchallenge plasma glucose levels in adults without diagnosed diabetes. *Diabetes Care*. Mar 2007;30(3):516-22. doi:10.2337/dc06-1996

132. Gaesser GA, Angadi SS, Sawyer BJ. Exercise and diet, independent of weight loss, improve cardiometabolic risk profile in overweight and obese individuals. *Phys Sportsmed*. May 2011;39(2):87-97. doi:10.3810/psm.2011.05.1898

133. Roberts CK, Hevener AL, Barnard RJ. Metabolic syndrome and insulin resistance: underlying causes and modification by exercise training. *Compr Physiol*. Jan 2013;3(1):1-58. doi:10.1002/cphy.c110062

134. Magkos F, Tsekouras Y, Kavouras SA, Mittendorfer B, Sidossis LS. Improved insulin sensitivity after a single bout of exercise is curvilinearly related to exercise energy expenditure. *Clin Sci (Lond)*. Jan 2008;114(1):59-64. doi:10.1042/cs20070134

135. Wang X, Patterson BW, Smith GI, et al. A ~60-min brisk walk increases insulin-stimulated glucose disposal but has no effect on hepatic and adipose tissue insulin sensitivity in older women. *J Appl Physiol (1985)*. Jun 2013;114(11):1563-8. doi:10.1152/japplphysiol.01364.2012

136. Snowling NJ, Hopkins WG. Effects of different modes of exercise training on glucose control and risk factors for complications in type 2 diabetic patients: a meta-analysis. *Diabetes Care*. Nov 2006;29(11):2518-27. doi:10.2337/dc06-1317

137. Jelleyman C, Yates T, O'Donovan G, et al. The effects of high-intensity interval training on glucose regulation and insulin resistance: a meta-analysis. *Obes Rev*. Nov 2015;16(11):942-61. doi:10.1111/obr.12317

138. Hudish LI, Reusch JE, Sussel L. β Cell dysfunction during progression of metabolic syndrome to type 2 diabetes. *J Clin Invest*. Oct 1 2019;129(10):4001-4008. doi:10.1172/jci129188

139. Venables MC, Jeukendrup AE. Physical inactivity and obesity: links with insulin resistance and type 2 diabetes mellitus. *Diabetes Metab Res Rev*. Sep 2009;25 Suppl 1:S18-23. doi:10.1002/dmrr.983

140. Eckardt K, Taube A, Eckel J. Obesity-associated insulin resistance in skeletal muscle: role of lipid accumulation and physical inactivity. *Rev Endocr Metab Disord*. Sep 2011;12(3):163-72. doi:10.1007/s11154-011-9168-2

141. Ali O. Genetics of type 2 diabetes. *World J Diabetes*. Aug 15 2013;4(4):114-23. doi:10.4239/wjd.v4.i4.114

142. Permutt MA, Wasson J, Cox N. Genetic epidemiology of diabetes. *J Clin Invest*. Jun 2005;115(6):1431-9. doi:10.1172/jci24758

143. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet*. Jan 18 2012;13(2):135-45. doi:10.1038/nrg3118

144. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature*. Aug 4 2016;536(7614):41-47. doi:10.1038/nature18642

145. Vaxillaire M, Bonnefond A, Froguel P. The lessons of early-onset monogenic diabetes for the understanding of diabetes pathogenesis. *Best Pract Res Clin Endocrinol Metab*. Apr 2012;26(2):171-87. doi:10.1016/j.beem.2011.12.001

146. Scott RA, Scott LJ, Mägi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes*. Nov 2017;66(11):2888-2902. doi:10.2337/db16-1253

147. Mahajan A, Wessel J, Willems SM, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet*. Apr 2018;50(4):559-571. doi:10.1038/s41588-018-0084-1

148. Xue A, Wu Y, Zhu Z, et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat Commun*. Jul 27 2018;9(1):2941. doi:10.1038/s41467-018-04951-w

149. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet*. Nov 2018;50(11):1505-1513. doi:10.1038/s41588-018-0241-6

150. Spracklen CN, Horikoshi M, Kim YJ, et al. Identification of type 2 diabetes loci in 433,540 East Asian individuals. *Nature*. Jun 2020;582(7811):240-245. doi:10.1038/s41586-020-2263-3

151. Goodarzi MO, Rotter JI. Genetics Insights in the Relationship Between Type 2 Diabetes and Coronary Heart Disease. *Circ Res*. May 22 2020;126(11):1526-1548. doi:10.1161/circresaha.119.316065

152. Hanis CL, Boerwinkle E, Chakraborty R, et al. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet*. Jun 1996;13(2):161-6. doi:10.1038/ng0696-161

153. Duggirala R, Blangero J, Almasy L, et al. Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. *Am J Hum Genet*. Apr 1999;64(4):1127-40. doi:10.1086/302316

154. Grant SF, Thorleifsson G, Reynisdottir I, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet*. Mar 2006;38(3):320-3. doi:10.1038/ng1732

155. Tong Y, Lin Y, Zhang Y, et al. Association between TCF7L2 gene polymorphisms and susceptibility to type 2 diabetes mellitus: a large Human Genome Epidemiology (HuGE) review and meta-analysis. *BMC Med Genet*. Feb 19 2009;10:15. doi:10.1186/1471-2350-10-15

156. Ruchat SM, Weisnagel SJ, Vohl MC, Rankinen T, Bouchard C, Pérusse L. Evidence for interaction between PPARG Pro12Ala and PPARGC1A Gly482Ser polymorphisms in determining type 2 diabetes

intermediate phenotypes in overweight subjects. *Exp Clin Endocrinol Diabetes*. Oct 2009;117(9):455-9. doi:10.1055/s-0029-1216352

157. Le Fur S, Le Stunff C, Bougnères P. Increased insulin resistance in obese children who have both 972 IRS-1 and 1057 IRS-2 polymorphisms. *Diabetes*. Dec 2002;51 Suppl 3:S304-7. doi:10.2337/diabetes.51.2007.s304

158. Hani EH, Boutin P, Durand E, et al. Missense mutations in the pancreatic islet beta cell inwardly rectifying K+ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of Type II diabetes mellitus in Caucasians. *Diabetologia*. Dec 1998;41(12):1511-5. doi:10.1007/s001250051098

159. Consortium IH. A haplotype map of the human genome. *Nature*. Oct 27 2005;437(7063):1299-320. doi:10.1038/nature04226

160. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. Jun 1 2007;316(5829):1336-41. doi:10.1126/science.1142364

161. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science*. Jun 1 2007;316(5829):1341-5. doi:10.1126/science.1142382

162. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet*. Jun 2007;39(6):770-5. doi:10.1038/ng2043

163. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. May 2008;40(5):638-45. doi:10.1038/ng.120

164. Simonis-Bik AM, Nijpels G, van Haeften TW, et al. Gene variants in the novel type 2 diabetes loci CDC123/CAMK1D, THADA, ADAMTS9, BCL11A, and MTNR1B affect different aspects of pancreatic beta-cell function. *Diabetes*. Jan 2010;59(1):293-301. doi:10.2337/db09-1048

165. Grarup N, Andersen G, Krarup NT, et al. Association testing of novel type 2 diabetes risk alleles in the JAZF1, CDC123/CAMK1D, TSPAN8, THADA, ADAMTS9, and NOTCH2 loci with insulin release, insulin sensitivity, and obesity in a population-based sample of 4,516 glucose-tolerant middle-aged Danes. *Diabetes*. Sep 2008;57(9):2534-40. doi:10.2337/db08-0436

166. Langenberg C, Lotta LA. Genomic insights into the causes of type 2 diabetes. *Lancet*. Jun 16 2018;391(10138):2463-2474. doi:10.1016/s0140-6736(18)31132-2

167. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. Oct 2016;48(10):1279-83. doi:10.1038/ng.3643

168. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. Feb 2010;42(2):105-16. doi:10.1038/ng.520

169. Prokopenko I, Langenberg C, Florez JC, et al. Variants in MTNR1B influence fasting glucose levels. *Nat Genet*. Jan 2009;41(1):77-81. doi:10.1038/ng.290

170. Saxena R, Hivert MF, Langenberg C, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nat Genet*. Feb 2010;42(2):142-8. doi:10.1038/ng.521

171. Prokopenko I, Poon W, Mägi R, et al. A central role for GRB10 in regulation of islet function in man. *PLoS Genet*. Apr 2014;10(4):e1004235. doi:10.1371/journal.pgen.1004235

172. Scott RA, Fall T, Pasko D, et al. Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. *Diabetes*. Dec 2014;63(12):4378-4387. doi:10.2337/db14-0319

173. Morris AP. Progress in defining the genetic contribution to type 2 diabetes susceptibility. *Curr Opin Genet Dev*. Jun 2018;50:41-51. doi:10.1016/j.gde.2018.02.003

174. Walford GA, Gustafsson S, Rybin D, et al. Genome-Wide Association Study of the Modified Stumvoll Insulin Sensitivity Index Identifies BCL2 and FAM19A2 as Novel Insulin Sensitivity Loci. *Diabetes*. Oct 2016;65(10):3200-11. doi:10.2337/db16-0199

175. Knowles JW, Xie W, Zhang Z, et al. Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. *J Clin Invest*. Apr 2015;125(4):1739-51. doi:10.1172/jci74692

176. Dimas AS, Lagou V, Barker A, et al. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes*. Jun 2014;63(6):2158-71. doi:10.2337/db13-0949

177. Lyssenko V, Lupi R, Marchetti P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *J Clin Invest*. Aug 2007;117(8):2155-63. doi:10.1172/jci30706

178. Villareal DT, Robertson H, Bell GI, et al. TCF7L2 variant rs7903146 affects the risk of type 2 diabetes by modulating incretin action. *Diabetes*. Feb 2010;59(2):479-85. doi:10.2337/db09-1169

179. Le Bacquer O, Kerr-Conte J, Gargani S, et al. TCF7L2 rs7903146 impairs islet function and morphology in non-diabetic individuals. *Diabetologia*. Oct 2012;55(10):2677-2681. doi:10.1007/s00125-012-2660-8

180. Florez JC, Hirschhorn J, Altshuler D. The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet*. 2003;4:257-91. doi:10.1146/annurev.genom.4.070802.110436

181. Udler MS, McCarthy MI, Florez JC, Mahajan A. Genetic Risk Scores for Diabetes Diagnosis and Precision Medicine. *Endocr Rev*. Dec 1 2019;40(6):1500-1520. doi:10.1210/er.2019-00088

182. Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock SJ, Park JH. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. *Nat Genet*. Apr 2013;45(4):400-5, 405e1-3. doi:10.1038/ng.2579

183. Hyttinen V, Kaprio J, Kinnunen L, Koskenvuo M, Tuomilehto J. Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. *Diabetes*. Apr 2003;52(4):1052-5. doi:10.2337/diabetes.52.4.1052

184. Meigs JB, Cupples LA, Wilson PW. Parental transmission of type 2 diabetes: the Framingham Offspring Study. *Diabetes*. Dec 2000;49(12):2201-7. doi:10.2337/diabetes.49.12.2201

185. Poulsen P, Kyvik KO, Vaag A, Beck-Nielsen H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia*. Feb 1999;42(2):139-45. doi:10.1007/s001250051131

186. Tillil H, Köbberling J. Age-corrected empirical genetic risk estimates for first-degree relatives of IDDM patients. *Diabetes*. Jan 1987;36(1):93-9. doi:10.2337/diab.36.1.93

187. Kaprio J, Tuomilehto J, Koskenvuo M, et al. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia*. Nov 1992;35(11):1060-7. doi:10.1007/bf02221682

188. Almgren P, Lehtovirta M, Isomaa B, et al. Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia*. Nov 2011;54(11):2811-9. doi:10.1007/s00125-011-2267-5

189. Ling C, Bacos K, Rönn T. Epigenetics of type 2 diabetes mellitus and weight change — a tool for precision medicine? *Nature Reviews Endocrinology*. 2022/07/01 2022;18(7):433-448. doi:10.1038/s41574-022-00671-w

190. Gluckman PD, Hanson MA, Buklijas T, Low FM, Beedle AS. Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. *Nat Rev Endocrinol*. Jul 2009;5(7):401-8. doi:10.1038/nrendo.2009.102

191. Ling C, Rönn T. Epigenetics in Human Obesity and Type 2 Diabetes. *Cell Metab*. May 7 2019;29(5):1028-1044. doi:10.1016/j.cmet.2019.03.009

192. Hall E, Volkov P, Dayeh T, et al. Sex differences in the genome-wide DNA methylation pattern and impact on gene expression, microRNA levels and insulin secretion in human pancreatic islets. *Genome Biology*. 2014/12/03 2014;15(12):522. doi:10.1186/s13059-014-0522-z

193. Barrès R, Osler ME, Yan J, et al. Non-CpG methylation of the PGC-1alpha promoter through DNMT3B controls mitochondrial density. *Cell Metab*. Sep 2009;10(3):189-98. doi:10.1016/j.cmet.2009.07.011

194. Ling C, Del Guerra S, Lupi R, et al. Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia*. Apr 2008;51(4):615-22. doi:10.1007/s00125-007-0916-5

195. Volkov P, Bacos K, Ofori JK, et al. Whole-Genome Bisulfite Sequencing of Human Pancreatic Islets Reveals Novel Differentially Methylated Regions in Type 2 Diabetes Pathogenesis. *Diabetes*. Apr 2017;66(4):1074-1085. doi:10.2337/db16-0996

196. Brøns C, Jacobsen S, Nilsson E, et al. Deoxyribonucleic acid methylation and gene expression of PPARGC1A in human muscle is influenced by high-fat overfeeding in a birth-weight-dependent manner. *J Clin Endocrinol Metab*. Jun 2010;95(6):3048-56. doi:10.1210/jc.2009-2413

197. Bysani M, Perfilyev A, de Mello VD, et al. Epigenetic alterations in blood mirror age-associated DNA methylation and gene expression changes in human liver. *Epigenomics*. Feb 2017;9(2):105-122. doi:10.2217/epi-2016-0087

198. Yang BT, Dayeh TA, Volkov PA, et al. Increased DNA methylation and decreased expression of PDX-1 in pancreatic islets from patients with type 2 diabetes. *Mol Endocrinol*. Jul 2012;26(7):1203-12. doi:10.1210/me.2012-1004

199. Nitert MD, Dayeh T, Volkov P, et al. Impact of an exercise intervention on DNA methylation in skeletal muscle from first-degree relatives of patients with type 2 diabetes. *Diabetes*. Dec 2012;61(12):3322-32. doi:10.2337/db11-1653

200. Lindholm ME, Marabita F, Gomez-Cabrero D, et al. An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training. *Epigenetics*. Dec 2014;9(12):1557-69. doi:10.4161/15592294.2014.982445

201. Sun D, Heianza Y, Li X, et al. Genetic, epigenetic and transcriptional variations at NFATC2IP locus with weight loss in response to diet interventions: The POUNDS Lost Trial. *Diabetes Obes Metab*. Sep 2018;20(9):2298-2303. doi:10.1111/dom.13333

202. Pan H, Lin X, Wu Y, et al. HIF3A association with adiposity: the story begins before birth. *Epigenomics*. 2015;7(6):937-50. doi:10.2217/epi.15.45

203. Wang S, Song J, Yang Y, Zhang Y, Wang H, Ma J. HIF3A DNA Methylation Is Associated with Childhood Obesity and ALT. *PLoS One*. 2015;10(12):e0145944. doi:10.1371/journal.pone.0145944

204. Richmond RC, Sharp GC, Ward ME, et al. DNA Methylation and BMI: Investigating Identified Methylation Sites at HIF3A in a Causal Framework. *Diabetes*. May 2016;65(5):1231-44. doi:10.2337/db15-0996

205. (NCD-RisC) NRFC. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. *Lancet*. Apr 9 2016;387(10027):1513-1530. doi:10.1016/s0140-6736(16)00618-8

206. Sekikawa A, Eguchi H, Tominaga M, et al. Prevalence of type 2 diabetes mellitus and impaired glucose tolerance in a rural area of Japan. The Funagata diabetes study. *J Diabetes Complications*. Mar-Apr 2000;14(2):78-83. doi:10.1016/s1056-8727(00)00074-x

207. Abate N, Chandalia M. The impact of ethnicity on type 2 diabetes. *J Diabetes Complications*. Jan-Feb 2003;17(1):39-58. doi:10.1016/s1056-8727(02)00190-3

208. Golden SH, Yajnik C, Phatak S, Hanson RL, Knowler WC. Racial/ethnic differences in the burden of type 2 diabetes over the life course: a focus on the USA and India. *Diabetologia*. Oct 2019;62(10):1751-1760. doi:10.1007/s00125-019-4968-0

209. Cowie CC, Casagrande SS, Geiss LS. Prevalence and Incidence of Type 2 Diabetes and Prediabetes. In: Cowie CC, Casagrande SS, Menke A, et al, eds. *Diabetes in America*. National Institute of Diabetes and Digestive and Kidney Diseases (US); 2018.

210. Anand SS, Yusuf S, Vuksan V, et al. Differences in risk factors, atherosclerosis, and cardiovascular disease between ethnic groups in Canada: the Study of Health Assessment and Risk in Ethnic groups (SHARE). *Lancet*. Jul 22 2000;356(9226):279-84. doi:10.1016/s0140-6736(00)02502-2

211. Wulan SN, Westerterp KR, Plasqui G. Ethnic differences in body composition and the associated metabolic profile: a comparative study between Asians and Caucasians. *Maturitas*. Apr 2010;65(4):315-9. doi:10.1016/j.maturitas.2009.12.012

212. Anand SS, Tarnopolsky MA, Rashid S, et al. Adipocyte hypertrophy, fatty liver and metabolic risk factors in South Asians: the Molecular Study of Health and Risk in Ethnic Groups (mol-SHARE). *PLoS One*. 2011;6(7):e22112. doi:10.1371/journal.pone.0022112

213. Sniderman AD, Bhopal R, Prabhakaran D, Sarrafzadegan N, Tchernof A. Why might South Asians be so susceptible to central obesity and its atherogenic consequences? The adipose tissue overflow hypothesis. *Int J Epidemiol*. Feb 2007;36(1):220-5. doi:10.1093/ije/dyl245

214. Morimoto Y, Maskarinec G, Conroy SM, Lim U, Shepherd J, Novotny R. Asian ethnicity is associated with a higher trunk/peripheral fat ratio in women and adolescent girls. *J Epidemiol*. 2012;22(2):130-5. doi:10.2188/jea.je20110100

215. He Q, Horlick M, Thornton J, et al. Sex and race differences in fat distribution among Asian, African-American, and Caucasian prepubertal children. *J Clin Endocrinol Metab*. May 2002;87(5):2164-70. doi:10.1210/jcem.87.5.8452

216. Goff LM, Ladwa M, Hakim O, Bello O. Ethnic distinctions in the pathophysiology of type 2 diabetes: a focus on black African-Caribbean populations. *Proc Nutr Soc*. May 2020;79(2):184-193. doi:10.1017/s0029665119001034

217. Sumner AE, Micklesfield LK, Ricks M, et al. Waist circumference, BMI, and visceral adipose tissue in white women and women of African descent. *Obesity (Silver Spring)*. Mar 2011;19(3):671-4. doi:10.1038/oby.2010.201

218. Ji Y, Yiorkas AM, Frau F, et al. Genome-Wide and Abdominal MRI Data Provide Evidence That a Genetically Determined Favorable Adiposity Phenotype Is Characterized by Lower Ectopic Liver Fat and Lower Risk of Type 2 Diabetes, Heart Disease, and Hypertension. *Diabetes*. Jan 2019;68(1):207-219. doi:10.2337/db18-0708

219. Yaghootkar H, Lotta LA, Tyrrell J, et al. Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease. *Diabetes*. Aug 2016;65(8):2448-60. doi:10.2337/db15-1671

220. Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat Genet*. Jan 2017;49(1):17-26. doi:10.1038/ng.3714

221. Haffner SM, D'Agostino R, Saad MF, et al. Increased insulin resistance and insulin secretion in nondiabetic African-Americans and Hispanics compared with non-Hispanic whites. The Insulin Resistance Atherosclerosis Study. *Diabetes*. Jun 1996;45(6):742-8. doi:10.2337/diab.45.6.742

222. Kodama K, Tojjar D, Yamada S, Toda K, Patel CJ, Butte AJ. Ethnic differences in the relationship between insulin sensitivity and insulin response: a systematic review and meta-analysis. *Diabetes Care*. Jun 2013;36(6):1789-96. doi:10.2337/dc12-1235

223. Gujral UP, Mohan V, Pradeepa R, et al. Ethnic Variations in Diabetes and Prediabetes Prevalence and the roles of Insulin Resistance and β-cell Function: The CARRS and NHANES Studies. *J Clin Transl Endocrinol*. Jun 1 2016;4:19-27. doi:10.1016/j.jcte.2016.02.004

224. Tan VM, Lee YS, Venkataraman K, et al. Ethnic differences in insulin sensitivity and beta-cell function among Asian men. *Nutr Diabetes*. Jul 20 2015;5(7):e173. doi:10.1038/nutd.2015.24

225. Cho YS, Chen CH, Hu C, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat Genet*. Dec 11 2011;44(1):67-72. doi:10.1038/ng.1019

226. Imamura M, Takahashi A, Yamauchi T, et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nat Commun*. Jan 28 2016;7:10531. doi:10.1038/ncomms10531

227. Kooner JS, Saleheen D, Sim X, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat Genet*. Aug 28 2011;43(10):984-9. doi:10.1038/ng.921

228. Parra EJ, Below JE, Krithika S, et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. *Diabetologia*. Aug 2011;54(8):2038-46. doi:10.1007/s00125-011-2172-y

229. Ng MC, Shriner D, Chen BH, et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet*. Aug 2014;10(8):e1004517. doi:10.1371/journal.pgen.1004517

230. Waters KM, Stram DO, Hassanein MT, et al. Consistent association of type 2 diabetes risk variants found in europeans in diverse racial and ethnic groups. *PLoS Genet*. Aug 26 2010;6(8)doi:10.1371/journal.pgen.1001078

231. Mahajan A, Go MJ, Zhang W, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet*. Mar 2014;46(3):234-44. doi:10.1038/ng.2897

232. Estrada K, Aukrust I, Bjørkhaug L, et al. Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *Jama*. Jun 11 2014;311(22):2305-14. doi:10.1001/jama.2014.6511

233. Mercader JM, Liao RG, Bell AD, et al. A Loss-of-Function Splice Acceptor Variant in IGF2 Is Protective for Type 2 Diabetes. *Diabetes*. Nov 2017;66(11):2903-2914. doi:10.2337/db17-0187

234. Williams AL, Jacobs SB, Moreno-Macías H, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature*. Feb 6 2014;506(7486):97-101. doi:10.1038/nature12828

235. Andersen MK, Pedersen CE, Moltke I, Hansen T, Albrechtsen A, Grarup N. Genetics of Type 2 Diabetes: the Power of Isolated Populations. *Curr Diab Rep*. Jul 2016;16(7):65. doi:10.1007/s11892-016-0757-z

236. Masel J. Genetic drift. *Curr Biol*. Oct 25 2011;21(20):R837-8. doi:10.1016/j.cub.2011.08.007

237. Moltke I, Grarup N, Jørgensen ME, et al. A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature*. Aug 14 2014;512(7513):190-3. doi:10.1038/nature13425

238. Grarup N, Moltke I, Andersen MK, et al. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nat Genet*. Feb 2018;50(2):172-174. doi:10.1038/s41588-017-0022-7

239. Tay GK, Henschel A, Daw Elbait G, Al Safar HS. Genetic Diversity and Low Stratification of the Population of the United Arab Emirates. *Front Genet*. 2020;11:608. doi:10.3389/fgene.2020.00608

240. Hunter-Zinck H, Musharoff S, Salit J, et al. Population genetic structure of the people of Qatar. *Am J Hum Genet*. Jul 9 2010;87(1):17-25. doi:10.1016/j.ajhg.2010.05.018

241. Mohammad T, Xue Y, Evison M, Tyler-Smith C. Genetic structure of nomadic Bedouin from Kuwait. *Heredity (Edinb)*. Nov 2009;103(5):425-33. doi:10.1038/hdy.2009.72

242. Alsmadi O, John SE, Thareja G, et al. Genome at juncture of early human migration: a systematic analysis of two whole genomes and thirteen exomes from Kuwaiti population subgroup of inferred Saudi Arabian tribe ancestry. *PLoS One*. 2014;9(6):e99069. doi:10.1371/journal.pone.0099069

243. John SE, Thareja G, Hebbar P, Behbehani K, Thanaraj TA, Alsmadi O. Kuwaiti population subgroup of nomadic Bedouin ancestry-Whole genome sequence and analysis. *Genom Data*. Mar 2015;3:116-27. doi:10.1016/j.gdata.2014.11.016

244. Elliott KS, Haber M, Daggag H, et al. Fine-Scale Genetic Structure in the United Arab Emirates Reflects Endogamous and Consanguineous Culture, Population History, and Geography. *Mol Biol Evol*. Mar 2 2022;39(3)doi:10.1093/molbev/msac039

245. Weiss H, Courty MA, Wetterstrom W, et al. The genesis and collapse of third millennium north mesopotamian civilization. *Science*. Aug 20 1993;261(5124):995-1004. doi:10.1126/science.261.5124.995

246. Almarri MA, Haber M, Lootah RA, et al. The genomic history of the Middle East. *Cell*. Sep 2 2021;184(18):4612-4625.e14. doi:10.1016/j.cell.2021.07.013

247. Zayed H. The Arab genome: Health and wealth. *Gene*. Nov 5 2016;592(2):239-43. doi:10.1016/j.gene.2016.07.007

248. Tadmouri GO, Nair P, Obeid T, Al Ali MT, Al Khaja N, Hamamy HA. Consanguinity and reproductive health among Arabs. *Reprod Health*. Oct 8 2009;6:17. doi:10.1186/1742-4755-6-17

249. Hamamy H. Consanguineous marriages : Preconception consultation in primary health care settings. *J Community Genet*. Jul 2012;3(3):185-92. doi:10.1007/s12687-011-0072-y

250. al-Mahroos F, McKeigue PM. High prevalence of diabetes in Bahrainis. Associations with ethnicity and raised plasma cholesterol. *Diabetes Care*. Jun 1998;21(6):936-42. doi:10.2337/diacare.21.6.936

251. Abdella N, Khogali M, al-Ali S, Gumaa K, Bajaj J. Known type 2 diabetes mellitus among the Kuwaiti population. A prevalence study. *Acta Diabetol*. Jul 1996;33(2):145-9. doi:10.1007/bf00569425

252. Al-Lawati JA, Al Riyami AM, Mohammed AJ, Jousilahti P. Increasing prevalence of diabetes mellitus in Oman. *Diabet Med*. Nov 2002;19(11):954-7. doi:10.1046/j.1464-5491.2002.00818.x

253. Al Riyami A, Elaty MA, Morsi M, Al Kharusi H, Al Shukaily W, Jaju S. Oman world health survey: part 1 - methodology, sociodemographic profile and epidemiology of non-communicable diseases in oman. *Oman Med J*. Sep 2012;27(5):425-43.

254. Bener A, Zirie M, Janahi IM, Al-Hamaq AO, Musallam M, Wareham NJ. Prevalence of diagnosed and undiagnosed diabetes mellitus and its risk factors in a population-based study of Qatar. *Diabetes Res Clin Pract*. Apr 2009;84(1):99-106. doi:10.1016/j.diabres.2009.02.003

255. El-Hazmi MA, Warsy AS, Al-Swailem AR, Al-Swailem AM, Sulaimani R, Al-Meshari AA. Diabetes mellitus and impaired glucose tolerance in Saudi Arabia. *Ann Saudi Med*. Jul 1996;16(4):381-5. doi:10.5144/0256-4947.1996.381

256. Al-Nuaim AR. Prevalence of glucose intolerance in urban and rural communities in Saudi Arabia. *Diabet Med*. Jul 1997;14(7):595-602. doi:10.1002/(sici)1096-9136(199707)14:7<595::Aid-dia377>3.0.Co;2-c

257. Saadi H, Carruthers SG, Nagelkerke N, et al. Prevalence of diabetes mellitus and its complications in a population-based sample in Al Ain, United Arab Emirates. *Diabetes Res Clin Pract*. Dec 2007;78(3):369-77. doi:10.1016/j.diabres.2007.04.008

258. Alhyas L, McKay A, Majeed A. 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*. 2012;7(8):e40948. doi:10.1371/journal.pone.0040948

259. Khalil AB, Beshyah SA, Abdella N, et al. Diabesity in the Arabian Gulf: Challenges and Opportunities. *Oman Med J*. Jul 2018;33(4):273-282. doi:10.5001/omj.2018.53

260. (NCD-RisC) NRFC. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *Lancet*. Dec 16 2017;390(10113):2627-2642. doi:10.1016/s0140-6736(17)32129-3

261. Al-Lawati JA, Tuomilehto J. Diabetes risk score in Oman: a tool to identify prevalent type 2 diabetes among Arabs of the Middle East. *Diabetes Res Clin Pract*. Sep 2007;77(3):438-44. doi:10.1016/j.diabres.2007.01.013

262. Mabry RM, Reeves MM, Eakin EG, Owen N. Evidence of physical activity participation among men and women in the countries of the Gulf cooperation council: a review. *Obes Rev*. Jun 2010;11(6):457-64. doi:10.1111/j.1467-789X.2009.00655.x

263. Mabry RM, Winkler EA, Reeves MM, Eakin EG, Owen N. Associations of physical activity and sitting time with the metabolic syndrome among Omani adults. *Obesity (Silver Spring)*. Nov 2012;20(11):2290-5. doi:10.1038/oby.2012.26

264. Midhet FM, Al-Mohaimeed AA, Sharaf FK. Lifestyle related risk factors of type 2 diabetes mellitus in Saudi Arabia. *Saudi Med J*. Jul 2010;31(7):768-74.

265. Al-Daghri NM, Alkharfy KM, Alokail MS, et al. Assessing the contribution of 38 genetic loci to the risk of type 2 diabetes in the Saudi Arabian Population. *Clin Endocrinol (Oxf)*. Apr 2014;80(4):532-7. doi:10.1111/cen.12187

266. Alsmadi O, Al-Rubeaan K, Wakil SM, et al. Genetic study of Saudi diabetes (GSSD): significant association of the KCNJ11 E23K polymorphism with type 2 diabetes. *Diabetes Metab Res Rev*. Feb 2008;24(2):137-40. doi:10.1002/dmrr.777

267. Al-Sinani S, Woodhouse N, Al-Mamari A, et al. Association of gene variants with susceptibility to type 2 diabetes among Omanis. *World J Diabetes*. Mar 15 2015;6(2):358-66. doi:10.4239/wjd.v6.i2.358

268. Al Safar HS, Cordell HJ, Jafer O, et al. A genome-wide search for type 2 diabetes susceptibility genes in an extended Arab family. *Ann Hum Genet*. Nov 2013;77(6):488-503. doi:10.1111/ahg.12036

269. Osman W, Hassoun A, Jelinek HF, et al. Genetics of type 2 diabetes and coronary artery disease and their associations with twelve cardiometabolic traits in the United Arab Emirates population. *Gene*. Aug 5 2020;750:144722. doi:10.1016/j.gene.2020.144722

270. Zhang X, Geiss LS, Cheng YJ, Beckles GL, Gregg EW, Kahn HS. The missed patient with diabetes: how access to health care affects the detection of diabetes. *Diabetes Care*. Sep 2008;31(9):1748-53. doi:10.2337/dc08-0572

271. Gakidou E, Mallinger L, Abbott-Klafter J, et al. Management of diabetes and associated cardiovascular risk factors in seven countries: a comparison of data from national health examination surveys. *Bull World Health Organ*. Mar 1 2011;89(3):172-83. doi:10.2471/blt.10.080820

272. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Jama*. Feb 1 1995;273(5):402-7. doi:10.1001/jama.273.5.402

273. (NCD-RisC) NRFC. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. *Lancet*. Apr 2 2016;387(10026):1377-1396. doi:10.1016/s0140-6736(16)30054-x

274. (NCD-RisC) NRFC. Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. *Lancet*. Jan 7 2017;389(10064):37-55. doi:10.1016/s0140-6736(16)31919-5

275. Ahmad O, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. 2001;

276. Collins R. What makes UK Biobank special? *Lancet*. Mar 31 2012;379(9822):1173-4. doi:10.1016/s0140-6736(12)60404-8

277. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. Mar 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779

278. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *Am J Epidemiol*. Nov 1 2017;186(9):1026-1034. doi:10.1093/aje/kwx246

279. Anderson JJ, Welsh P, Ho FK, et al. Ethnic differences in prevalence of actionable HbA1c levels in UK Biobank: implications for screening. *BMJ Open Diabetes Res Care*. Aug 2021;9(1)doi:10.1136/bmjdrc-2021-002176

280. Lin L, Dekkers IA, Tao Q, Lamb HJ. Novel artificial neural network and linear regression based equation for estimating visceral adipose tissue volume. *Clin Nutr*. Oct 2020;39(10):3182-3188. doi:10.1016/j.clnu.2020.02.013

281. IPAQ. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) - Short and Long Forms. 2005;

282. IPAQ. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) - Short Form. 2004;

283. Sehmi J. *The Determinants of the Excess Risk of Type-2 Diabetes Amongst Indian Asians Compared to Europeans*. Imperial College London; 2013.

284. Hajat C, Harrison O, Al Siksek Z. Weqaya: a population-wide cardiovascular screening program in Abu Dhabi, United Arab Emirates. *Am J Public Health*. May 2012;102(5):909-14. doi:10.2105/ajph.2011.300290

285. Keating XD, Zhou K, Liu X, et al. Reliability and Concurrent Validity of Global Physical Activity Questionnaire (GPAQ): A Systematic Review. *Int J Environ Res Public Health*. Oct 26 2019;16(21)doi:10.3390/ijerph16214128

286. Ács P, Betlehem J, Oláh A, et al. Cross-cultural adaptation and validation of the Global Physical Activity Questionnaire among healthy Hungarian adults. *BMC Public Health*. Aug 17 2020;20(Suppl 1):1056. doi:10.1186/s12889-020-08477-z

287. GPAQ. Global Physical Activity Questionnaire (GPAQ) Analysis Guide.

288. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc*. May 1998;30(5):777-81. doi:10.1097/00005768-199805000-00021

289. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*. Jul 1985;28(7):412-9. doi:10.1007/bf00280883

290. Macherey-Nagel. User Manual. Genomic DNA from blood.

291. Rayner NW, Robertson, N., Mahajan, A., McCarthy, M. I. A Suite of Programs For Pre- and Post-imputation Data Checking. University of Oxford.

292. Alkaf B, Blakemore AI, Järvelin MR, Lessan N. Secondary analyses of global datasets: do obesity and physical activity explain variation in diabetes risk across populations? *Int J Obes (Lond)*. May 2021;45(5):944-956. doi:10.1038/s41366-021-00764-y

293. Nakagami T, Qiao Q, Carstensen B, et al. Age, body mass index and Type 2 diabetes-associations modified by ethnicity. *Diabetologia*. Aug 2003;46(8):1063-70. doi:10.1007/s00125-003-1158-9

294. Lovejoy JC, de la Bretonne JA, Klemperer M, Tulley R. Abdominal fat distribution and metabolic risk factors: effects of race. *Metabolism*. Sep 1996;45(9):1119-24. doi:10.1016/s0026-0495(96)90011-6

295. Ujcic-Voortman JK, Schram MT, Jacobs-van der Bruggen MA, Verhoeff AP, Baan CA. Diabetes prevalence and risk factors among ethnic minorities. *Eur J Public Health*. Oct 2009;19(5):511-5. doi:10.1093/eurpub/ckp096

296. Admiraal WM, van Valkengoed IG, JS LdM, Stronks K, Hoekstra JB, Holleman F. The association of physical inactivity with Type 2 diabetes among different ethnic groups. *Diabet Med*. Jun 2011;28(6):668-72. doi:10.1111/j.1464-5491.2011.03248.x

297. Norgan NG. Population differences in body composition in relation to the body mass index. *Eur J Clin Nutr*. Nov 1994;48 Suppl 3:S10-25; discussion S26-7.

298. (NICE) NIfHaCE. BMI: preventing ill health and premature death in black, Asian and other minority ethnic groups.

299. Banerji MA, Faridi N, Atluri R, Chaiken RL, Lebovitz HE. Body composition, visceral fat, leptin, and insulin resistance in Asian Indian men. *J Clin Endocrinol Metab*. Jan 1999;84(1):137-44. doi:10.1210/jcem.84.1.5371

300. McKeigue PM, Pierpoint T, Ferrie JE, Marmot MG. Relationship of glucose intolerance and hyperinsulinaemia to body fat pattern in south Asians and Europeans. *Diabetologia*. Aug 1992;35(8):785-91. doi:10.1007/bf00429101

301. Karter AJ, Mayer-Davis EJ, Selby JV, et al. Insulin sensitivity and abdominal obesity in African-American, Hispanic, and non-Hispanic white men and women. The Insulin Resistance and Atherosclerosis Study. *Diabetes*. Nov 1996;45(11):1547-55. doi:10.2337/diab.45.11.1547

302. Kautzky-Willer A, Harreiter J, Pacini G. Sex and Gender Differences in Risk, Pathophysiology and Complications of Type 2 Diabetes Mellitus. *Endocr Rev*. Jun 2016;37(3):278-316. doi:10.1210/er.2015-1137

303. Tuomilehto J. The emerging global epidemic of type 1 diabetes. *Curr Diab Rep*. Dec 2013;13(6):795-804. doi:10.1007/s11892-013-0433-5

304. Chambers JC, Eda S, Bassett P, et al. C-reactive protein, insulin resistance, central obesity, and coronary heart disease risk in Indian Asians from the United Kingdom compared with European whites. *Circulation*. Jul 10 2001;104(2):145-50. doi:10.1161/01.cir.104.2.145

305. Patel A, MacMahon S, Chalmers J, et al. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med*. Jun 12 2008;358(24):2560-72. doi:10.1056/NEJMoa0802987

306. King H, Rewers M. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. WHO Ad Hoc Diabetes Reporting Group. *Diabetes Care*. Jan 1993;16(1):157-77. doi:10.2337/diacare.16.1.157

307. Cumming G. Inference by eye: reading the overlap of independent confidence intervals. *Stat Med*. Jan 30 2009;28(2):205-20. doi:10.1002/sim.3471

308. Siegel AF, Wagner, M. F. *Practical Business Statistics (Eighth Edition)*. 2022.

309. Greendale GA, Sternfeld B, Huang M, et al. Changes in body composition and weight during the menopause transition. *JCI Insight*. Mar 7 2019;4(5)doi:10.1172/jci.insight.124865

310. González-Muniesa P, Mártinez-González MA, Hu FB, et al. Obesity. *Nat Rev Dis Primers*. Jun 15 2017;3:17034. doi:10.1038/nrdp.2017.34

311. Tchernof A, Després JP. Pathophysiology of human visceral obesity: an update. *Physiol Rev*. Jan 2013;93(1):359-404. doi:10.1152/physrev.00033.2011

312. Tomiyama AJ, Hunger JM, Nguyen-Cuu J, Wells C. Misclassification of cardiometabolic health when using body mass index categories in NHANES 2005-2012. *Int J Obes (Lond)*. May 2016;40(5):883-6. doi:10.1038/ijo.2016.17

313. Kantartzis K, Machann J, Schick F, Fritsche A, Häring HU, Stefan N. The impact of liver fat vs visceral fat in determining categories of prediabetes. *Diabetologia*. May 2010;53(5):882-9. doi:10.1007/s00125-010-1663-6

314. Heni M, Machann J, Staiger H, et al. Pancreatic fat is negatively associated with insulin secretion in individuals with impaired fasting glucose and/or impaired glucose tolerance: a nuclear magnetic resonance study. *Diabetes Metab Res Rev*. Mar 2010;26(3):200-5. doi:10.1002/dmrr.1073

315. Alenaini W, Parkinson JRC, McCarthy JP, et al. Ethnic Differences in Body Fat Deposition and Liver Fat Content in Two UK-Based Cohorts. *Obesity (Silver Spring)*. Nov 2020;28(11):2142-2152. doi:10.1002/oby.22948

316. Linnemann AK, Baan M, Davis DB. Pancreatic β-cell proliferation in obesity. *Adv Nutr*. May 2014;5(3):278-88. doi:10.3945/an.113.005488

317. van Vliet S, Koh HE, Patterson BW, et al. Obesity Is Associated With Increased Basal and Postprandial β-Cell Insulin Secretion Even in the Absence of Insulin Resistance. *Diabetes*. Oct 2020;69(10):2112-2119. doi:10.2337/db20-0377

318. Jeon CY, Lokken RP, Hu FB, van Dam RM. Physical activity of moderate intensity and risk of type 2 diabetes: a systematic review. *Diabetes Care*. Mar 2007;30(3):744-52. doi:10.2337/dc06-1842

319. Ramachandran A, Snehalatha C, Mary S, Mukesh B, Bhaskar AD, Vijay V. The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). *Diabetologia*. Feb 2006;49(2):289-97. doi:10.1007/s00125-005-0097-z

320. Colberg SR, Sigal RJ, Fernhall B, et al. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement. *Diabetes Care*. Dec 2010;33(12):e147-67. doi:10.2337/dc10-9990

321. Hamilton MT, Hamilton DG, Zderic TW. Exercise physiology versus inactivity physiology: an essential concept for understanding lipoprotein lipase regulation. *Exerc Sport Sci Rev*. Oct 2004;32(4):161-6. doi:10.1097/00003677-200410000-00007

322. Hamilton MT, Hamilton DG, Zderic TW. Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. *Diabetes*. Nov 2007;56(11):2655-67. doi:10.2337/db07-0882

323. Hamilton MT, Hamilton DG, Zderic TW. Sedentary behavior as a mediator of type 2 diabetes. *Med Sport Sci*. 2014;60:11-26. doi:10.1159/000357332

324. Irwin ML, Mayer-Davis EJ, Addy CL, et al. Moderate-intensity physical activity and fasting insulin levels in women: the Cross-Cultural Activity Participation Study. *Diabetes Care*. Apr 2000;23(4):449-54. doi:10.2337/diacare.23.4.449

325. Monda KL, Ballantyne CM, North KE. Longitudinal impact of physical activity on lipid profiles in middle-aged adults: the Atherosclerosis Risk in Communities Study. *J Lipid Res*. Aug 2009;50(8):1685-91. doi:10.1194/jlr.P900029-JLR200

326. Ekelund U, Franks PW, Sharp S, Brage S, Wareham NJ. Increase in physical activity energy expenditure is associated with reduced metabolic risk independent of change in fatness and fitness. *Diabetes Care*. Aug 2007;30(8):2101-6. doi:10.2337/dc07-0719

327. Kriska AM, Pereira MA, Hanson RL, et al. Association of physical activity and serum insulin concentrations in two populations at high risk for type 2 diabetes but differing by BMI. *Diabetes Care*. Jul 2001;24(7):1175-80. doi:10.2337/diacare.24.7.1175

328. Bradbury KE, Guo W, Cairns BJ, Armstrong ME, Key TJ. Association between physical activity and body fat percentage, with adjustment for BMI: a large cross-sectional analysis of UK Biobank. *BMJ Open*. Mar 24 2017;7(3):e011843. doi:10.1136/bmjopen-2016-011843

329. Doyle C, Khan A, Burton N. Reliability and validity of a self-administered Arabic version of the Global Physical Activity Questionnaire (GPAQ-A). *J Sports Med Phys Fitness*. Jul 2019;59(7):1221-1228. doi:10.23736/s0022-4707.18.09186-7

330. Guthold R. Report from the WHO Physical Activity Global Surveillance Meeting. *World Health Organization, Geneva*. 2009;

331. Samara A, Andersen PT, Aro AR. Health Promotion and Obesity in the Arab Gulf States: Challenges and Good Practices. *J Obes*. 2019;2019:4756260. doi:10.1155/2019/4756260

332. Samara A, Nistrup A, Al-Rammah TY, Aro AR. Lack of facilities rather than sociocultural factors as the primary barrier to physical activity among female Saudi university students. *Int J Womens Health*. 2015;7:279-86. doi:10.2147/ijwh.S80680

333. Al-Kaabi J, Al-Maskari F, Saadi H, Afandi B, Parkar H, Nagelkerke N. Physical activity and reported barriers to activity among type 2 diabetic patients in the United arab emirates. *Rev Diabet Stud*. Winter 2009;6(4):271-8. doi:10.1900/rds.2009.6.271

334. Goodyear LJ, Kahn BB. Exercise, glucose transport, and insulin sensitivity. *Annu Rev Med*. 1998;49:235-61. doi:10.1146/annurev.med.49.1.235

335. Yaribeygi H, Maleki M, Sathyapalan T, Jamialahmadi T, Sahebkar A. Pathophysiology of Physical Inactivity-Dependent Insulin Resistance: A Theoretical Mechanistic Review Emphasizing Clinical Evidence. *J Diabetes Res*. 2021;2021:7796727. doi:10.1155/2021/7796727

336. Crichton GE, Alkerwi A. Physical activity, sedentary behavior time and lipid levels in the Observation of Cardiovascular Risk Factors in Luxembourg study. *Lipids Health Dis*. Aug 11 2015;14:87. doi:10.1186/s12944-015-0085-3

337. Joseph JJ, Echouffo-Tcheugui JB, Golden SH, et al. Physical activity, sedentary behaviors and the incidence of type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA). *BMJ Open Diabetes Res Care*. 2016;4(1):e000185. doi:10.1136/bmjdrc-2015-000185

338. Fernandes RA, Zanesco A. Early sport practice is related to lower prevalence of cardiovascular and metabolic outcomes in adults independently of overweight and current physical activity. *Medicina (Kaunas)*. 2015;51(6):336-42. doi:10.1016/j.medici.2015.10.003

339. Celis-Morales CA, Perez-Bravo F, Ibañez L, Salas C, Bailey ME, Gill JM. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. *PLoS One*. 2012;7(5):e36345. doi:10.1371/journal.pone.0036345

340. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med*. Jun 2003;37(3):197-206; discussion 206. doi:10.1136/bjsm.37.3.197

341. Digital N. Statistics on Obesity, Physical Activity and Diet, England, 2019. 2019;

342. Yasuda K, Miyake K, Horikawa Y, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet*. Sep 2008;40(9):1092-7. doi:10.1038/ng.207

343. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. Oct 13 2016;538(7624):161-164. doi:10.1038/538161a

344. Mahajan A, Spracklen CN, Zhang W, et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. *Nat Genet*. May 2022;54(5):560-572. doi:10.1038/s41588-022-01058-3

345. Loh M, Zhang W, Ng HK, et al. Identification of genetic effects underlying type 2 diabetes in South Asian and European populations. *Commun Biol*. Apr 7 2022;5(1):329. doi:10.1038/s42003-022-03248-5

346. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. *Cell*. May 2 2019;177(4):1080. doi:10.1016/j.cell.2019.04.032

347. Pasaniuc B, Zaitlen N, Shi H, et al. Fast and accurate imputation of summary statistics enhances evidence of functional enrichment. *Bioinformatics*. Oct 15 2014;30(20):2906-14. doi:10.1093/bioinformatics/btu416

348. Banda Y, Kvale MN, Hoffmann TJ, et al. Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics*. Aug 2015;200(4):1285-95. doi:10.1534/genetics.115.178616

349. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. Oct 2018;562(7726):203-209. doi:10.1038/s41586-018-0579-z

350. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. Oct 2016;48(10):1284-1287. doi:10.1038/ng.3656

351. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*. Mar 18 2012;44(4):369-75, s1-3. doi:10.1038/ng.2213

352. Geeketics. LocusZooms. Github. https://github.com/Geeketics/LocusZooms/commit/d34b65be030a1c74dac83582bb31330dd5e69500#diff-dbfb0088cfa05321739ebe775cd752d8b1c271bf393720c1653390cab4adf02c

353. Patel CJ, Chen R, Kodama K, Ioannidis JP, Butte AJ. Systematic identification of interaction effects between genome- and environment-wide associations in type 2 diabetes mellitus. *Hum Genet*. May 2013;132(5):495-508. doi:10.1007/s00439-012-1258-z

354. Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. Feb 22 2007;445(7130):881-5. doi:10.1038/nature05616

355. Del Bosque-Plata L, Martínez-Martínez E, Espinoza-Camacho M, Gragnoli C. The Role of TCF7L2 in Type 2 Diabetes. *Diabetes*. Jun 2021;70(6):1220-1228. doi:10.2337/db20-0573

356. Fujita Y, Tanaka D, Tatsuoka H, et al. A novel splice-site mutation of the HNF1B gene in a family with maturity onset diabetes of the young type 5 (MODY5). *Endocrinol Diabetes Metab Case Rep*. Sep 23 2020;2020doi:10.1530/edm-20-0092

357. Kaneko YK, Ishikawa T. Diacylglycerol Signaling Pathway in Pancreatic β-Cells: An Essential Role of Diacylglycerol Kinase in the Regulation of Insulin Secretion. *Biol Pharm Bull*. 2015;38(5):669-73. doi:10.1248/bpb.b15-00060

358. Florez JC, Jablonski KA, Sun MW, et al. Effects of the type 2 diabetes-associated PPARG P12A polymorphism on progression to diabetes and response to troglitazone. *J Clin Endocrinol Metab*. Apr 2007;92(4):1502-9. doi:10.1210/jc.2006-2275

359. Palmer CJ, Bruckner RJ, Paulo JA, et al. Cdkal1, a type 2 diabetes susceptibility gene, regulates mitochondrial function in adipose tissue. *Mol Metab*. Oct 2017;6(10):1212-1225. doi:10.1016/j.molmet.2017.07.013

360. Wang LX, Frey MR, Kohli R. The Role of FGF19 and MALRD1 in Enterohepatic Bile Acid Signaling. *Front Endocrinol (Lausanne)*. 2021;12:799648. doi:10.3389/fendo.2021.799648

361. Anderson D, Cordell HJ, Fakiola M, et al. First genome-wide association study in an Australian aboriginal population provides insights into genetic risk factors for body mass index and type 2 diabetes. *PLoS One*. 2015;10(3):e0119333. doi:10.1371/journal.pone.0119333

362. Kathiresan S, Manning AK, Demissie S, et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med Genet*. Sep 19 2007;8 Suppl 1(Suppl 1):S17. doi:10.1186/1471-2350-8-s1-s17

363. Fox CS, Liu Y, White CC, et al. Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet*. 2012;8(5):e1002695. doi:10.1371/journal.pgen.1002695

364. Wan JY, Goodman DL, Willems EL, et al. Genome-wide association analysis of metabolic syndrome quantitative traits in the GENNID multiethnic family study. *Diabetol Metab Syndr*. Jun 1 2021;13(1):59. doi:10.1186/s13098-021-00670-3

365. Hoffmann TJ, Ehret GB, Nandakumar P, et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat Genet*. Jan 2017;49(1):54-64. doi:10.1038/ng.3715

366. Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science*. Dec 1 2006;314(5804):1461-3. doi:10.1126/science.1135245

367. Chen J, Spracklen CN, Marenne G, et al. The trans-ancestral genomic architecture of glycemic traits. *Nat Genet*. Jun 2021;53(6):840-860. doi:10.1038/s41588-021-00852-9

368. Kong AP, Luk AO, Chan JC. Detecting people at high risk of type 2 diabetes- How do we find them and who should be treated? *Best Pract Res Clin Endocrinol Metab*. Jun 2016;30(3):345-55. doi:10.1016/j.beem.2016.06.003

369. Wildman RP, Muntner P, Reynolds K, et al. The obese without cardiometabolic risk factor clustering and the normal weight with cardiometabolic risk factor clustering: prevalence and correlates of 2 phenotypes among the US population (NHANES 1999-2004). *Arch Intern Med*. Aug 11 2008;168(15):1617-24. doi:10.1001/archinte.168.15.1617

370. Kanter R, Caballero B. Global gender disparities in obesity: a review. *Adv Nutr*. Jul 1 2012;3(4):491-8. doi:10.3945/an.112.002063

371. Demerath EW, Sun SS, Rogers N, et al. Anatomical patterning of visceral adipose tissue: race, sex, and age variation. *Obesity (Silver Spring)*. Dec 2007;15(12):2984-93. doi:10.1038/oby.2007.356

372. Goodman-Gruen D, Barrett-Connor E. Sex differences in measures of body fat and body fat distribution in the elderly. *Am J Epidemiol*. May 1 1996;143(9):898-906. doi:10.1093/oxfordjournals.aje.a008833

373. Nookaew I, Svensson PA, Jacobson P, et al. Adipose tissue resting energy expenditure and expression of genes involved in mitochondrial function are higher in women than in men. *J Clin Endocrinol Metab*. Feb 2013;98(2):E370-8. doi:10.1210/jc.2012-2764

374. Brussaard HE, Gevers Leuven JA, Frölich M, Kluft C, Krans HM. Short-term oestrogen replacement therapy improves insulin resistance, lipids and fibrinolysis in postmenopausal women with NIDDM. *Diabetologia*. Jul 1997;40(7):843-9. doi:10.1007/s001250050758

375. Grove KL, Fried SK, Greenberg AS, Xiao XQ, Clegg DJ. A microarray analysis of sexual dimorphism of adipose tissues in high-fat-diet-induced obese mice. *Int J Obes (Lond)*. Jun 2010;34(6):989-1000. doi:10.1038/ijo.2010.12

376. Danforth E, Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nat Genet*. Sep 2000;26(1):13. doi:10.1038/79111

377. Gregor MF, Hotamisligil GS. Thematic review series: Adipocyte Biology. Adipocyte stress: the endoplasmic reticulum and metabolic disease. *J Lipid Res*. Sep 2007;48(9):1905-14. doi:10.1194/jlr.R700007-JLR200

378. Müller MJ, Langemann D, Gehrke I, et al. Effect of constitution on mass of individual organs and their association with metabolic rate in humans--a detailed view on allometric scaling. *PLoS One*. 2011;6(7):e22732. doi:10.1371/journal.pone.0022732

379. Korenblat KM, Fabbrini E, Mohammed BS, Klein S. Liver, muscle, and adipose tissue insulin action is directly related to intrahepatic triglyceride content in obese subjects. *Gastroenterology*. May 2008;134(5):1369-75. doi:10.1053/j.gastro.2008.01.075

380. Alswat K, Aljumah AA, Sanai FM, et al. Nonalcoholic fatty liver disease burden - Saudi Arabia and United Arab Emirates, 2017-2030. *Saudi J Gastroenterol*. Jul-Aug 2018;24(4):211-219. doi:10.4103/sjg.SJG_122_18

381. Cholongitas E, Pavlopoulou I, Papatheodoridi M, et al. Epidemiology of nonalcoholic fatty liver disease in Europe: a systematic review and meta-analysis. *Ann Gastroenterol*. 2021;34(3):404-414. doi:10.20524/aog.2021.0604

382. Kral JG, Schaffner F, Pierson RN, Jr., Wang J. Body fat topography as an independent predictor of fatty liver. *Metabolism*. May 1993;42(5):548-51. doi:10.1016/0026-0495(93)90210-f

383. Lê KA, Ventura EE, Fisher JQ, et al. Ethnic differences in pancreatic fat accumulation and its relationship with other fat depots and inflammatory markers. *Diabetes Care*. Feb 2011;34(2):485-90. doi:10.2337/dc10-0760

384. Mongraw-Chaffin M, Foster MC, Anderson CAM, et al. Metabolically Healthy Obesity, Transition to Metabolic Syndrome, and Cardiovascular Risk. *J Am Coll Cardiol*. May 1 2018;71(17):1857-1865. doi:10.1016/j.jacc.2018.02.055

385. Garvey WT, Mechanick JI, Brett EM, et al. AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS AND AMERICAN COLLEGE OF ENDOCRINOLOGY COMPREHENSIVE CLINICAL PRACTICE GUIDELINES FOR MEDICAL CARE OF PATIENTS WITH OBESITY. *Endocr Pract*. Jul 2016;22 Suppl 3:1-203. doi:10.4158/ep161365.Gl

386. Jensen MD, Ryan DH, Apovian CM, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *J Am Coll Cardiol*. Jul 1 2014;63(25 Pt B):2985-3023. doi:10.1016/j.jacc.2013.11.004

387. Abusnana S, Fargaly M, Alfardan SH, et al. Clinical Practice Recommendations for the Management of Obesity in the United Arab Emirates. *Obes Facts*. 2018;11(5):413-428. doi:10.1159/000491796

388. Craft LL, Zderic TW, Gapstur SM, et al. Evidence that women meeting physical activity guidelines do not sit less: an observational inclinometry study. *Int J Behav Nutr Phys Act*. Oct 4 2012;9:122. doi:10.1186/1479-5868-9-122

389. Henson J, Yates T, Biddle SJ, et al. Associations of objectively measured sedentary behaviour and physical activity with markers of cardiometabolic health. *Diabetologia*. May 2013;56(5):1012-20. doi:10.1007/s00125-013-2845-9

390. Healy GN, Dunstan DW, Salmon J, et al. Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care*. Apr 2008;31(4):661-6. doi:10.2337/dc07-2046

391. Gill JM, Cooper AR. Physical activity and prevention of type 2 diabetes mellitus. *Sports Med*. 2008;38(10):807-24. doi:10.2165/00007256-200838100-00002

392. Zhang K, Pi-Sunyer FX, Boozer CN. Improving energy expenditure estimation for physical activity. *Med Sci Sports Exerc*. May 2004;36(5):883-9. doi:10.1249/01.mss.0000126585.40962.22

393. Sørensen TIA, Metz S, Kilpeläinen TO. Do gene-environment interactions have implications for the precision prevention of type 2 diabetes? *Diabetologia*. Jan 7 2022;doi:10.1007/s00125-021-05639-5

394. Renalds A, Smith TH, Hale PJ. A systematic review of built environment and health. *Fam Community Health*. Jan-Mar 2010;33(1):68-78. doi:10.1097/FCH.0b013e3181c4e2e5

395. Darbre PD. Overview of air pollution and endocrine disorders. *Int J Gen Med*. 2018;11:191-207. doi:10.2147/ijgm.S102230

396. vom Saal FS, Myers JP. Bisphenol A and risk of metabolic disorders. *Jama*. Sep 17 2008;300(11):1353-5. doi:10.1001/jama.300.11.1353

397. MacDonald Gibson J, Thomsen J, Launay F, Harder E, DeFelice N. Deaths and medical visits attributable to environmental pollution in the United Arab Emirates. *PLoS One*. 2013;8(3):e57536. doi:10.1371/journal.pone.0057536

398. Muraki I, Wada H, Tanigawa T. Sleep apnea and type 2 diabetes. *J Diabetes Investig*. Sep 2018;9(5):991-997. doi:10.1111/jdi.12823

399. London KsCH. Sleep Apnea. https://kingscollegehospitaldubai.com/service/pulmonologist-dubai/sleep-medicine/sleep-apnea-lab/#:~:text=The%20estimated%20prevalence%20in%20UAE,high%20risk%20of%20Sleep%20Apnea.

400. Knowler WC, Pettitt DJ, Saad MF, Bennett PH. Diabetes mellitus in the Pima Indians: incidence, risk factors and pathogenesis. *Diabetes Metab Rev*. Feb 1990;6(1):1-27. doi:10.1002/dmr.5610060101

401. Hanson RL, Elston RC, Pettitt DJ, Bennett PH, Knowler WC. Segregation analysis of non-insulin-dependent diabetes mellitus in Pima Indians: evidence for a major-gene effect. *Am J Hum Genet*. Jul 1995;57(1):160-70.

402. Schulz LO, Bennett PH, Ravussin E, et al. Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S. *Diabetes Care*. Aug 2006;29(8):1866-71. doi:10.2337/dc06-0138

403. Westerman KE, Miao J, Chasman DI, et al. Genome-wide gene-diet interaction analysis in the UK Biobank identifies novel effects on hemoglobin A1c. *Hum Mol Genet*. Aug 28 2021;30(18):1773-1783. doi:10.1093/hmg/ddab109

404. Villuendas G, Botella-Carretero JI, López-Bermejo A, et al. The ACAA-insertion/deletion polymorphism at the 3' UTR of the IGF-II receptor gene is associated with type 2 diabetes and surrogate markers of insulin resistance. *Eur J Endocrinol*. Aug 2006;155(2):331-6. doi:10.1530/eje.1.02217

405. Chanprasertyothin S, Jongjaroenprasert W, Ongphiphadhanakul B. The association of soluble IGF2R and IGF2R gene polymorphism with type 2 diabetes. *J Diabetes Res*. 2015;2015:216383. doi:10.1155/2015/216383

406. Ong KK, Dunger DB. Developmental aspects in the pathogenesis of type 2 diabetes. *Mol Cell Endocrinol*. Dec 20 2001;185(1-2):145-9. doi:10.1016/s0303-7207(01)00625-6

407. Reue K, Lee JM, Vergnes L. Diet1 is a regulator of fibroblast growth factor 15/19-dependent bile acid synthesis. *Dig Dis*. 2015;33(3):307-13. doi:10.1159/000371649

408. Yang J, Bakshi A, Zhu Z, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet*. Oct 2015;47(10):1114-20. doi:10.1038/ng.3390

409. Zaitlen N, Kraft P. Heritability in the genome-wide association era. *Hum Genet*. Oct 2012;131(10):1655-64. doi:10.1007/s00439-012-1199-6

410. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev*. Jun 2009;19(3):212-9. doi:10.1016/j.gde.2009.04.010

411. Jun G, Manning A, Almeida M, et al. Evaluating the contribution of rare variants to type 2 diabetes and related traits using pedigrees. *Proc Natl Acad Sci U S A*. Jan 9 2018;115(2):379-384. doi:10.1073/pnas.1705859115

412. Portal TUAEG. The Emirati Genome Programme. https://u.ae/en/information-and-services/health-and-fitness/research-in-the-field-of-health

413. Yajnik CS, Fall CH, Coyaji KJ, et al. Neonatal anthropometry: the thin-fat Indian baby. The Pune Maternal Nutrition Study. *Int J Obes Relat Metab Disord*. Feb 2003;27(2):173-80. doi:10.1038/sj.ijo.802219

414. Chambers JC, Loh M, Lehne B, et al. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol*. Jul 2015;3(7):526-534. doi:10.1016/s2213-8587(15)00127-8

415. Wojcik GL, Graff M, Nishimura KK, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature*. Jun 2019;570(7762):514-518. doi:10.1038/s41586-019-1310-4

416. Kalra S, Kalra B, Unnikrishnan A, Agrawal N, Kumar S. Optimizing weight control in diabetes: antidiabetic drug selection. *Diabetes Metab Syndr Obes*. Aug 24 2010;3:297-9. doi:10.2147/dmsott.S11941

Appendix 1

مـركـز إمـبـريـال كـولـيـدج لـنـدن للسـكـري
IMPERIAL COLLEGE LONDON DIABETES CENTRE

A Mubadala Company

PARTICIPANT INFORMATION SHEET AND CONSENT FORM

Title: Genetic and epigenetic contribution to increased risk of T2DM in the UAE.

Introduction

You are invited to take part in a research study held by Imperial College London Diabetes Centre (ICLDC) in Abu Dhabi, UAE. Please read the following information carefully; it is important that you understand why the research is being done and what it is you are being asked to do. Please discuss it with your doctor, family or friends if you wish to; you do not need to decide whether to take part immediately.

Why is the study being done?

Type 2 Diabetes Mellitus (T2DM) is a very common disease in the Middle East, with the UAE being one of the top 20 countries with the highest rates of disease worldwide. This increased spread could be due to lifestyle changes of the Emirati population, such as reduced physical activity, and increased intake of total calories in terms of saturated fat (especially in fast food) and sugar-sweetened beverages. On the other hand, there could be genes that are unique to the Emirati population, which make them at a higher risk of developing T2DM compared to other populations.

In this study, we would like to investigate the reasons underlying increased risk of T2DM amongst the Emiratis in comparison to other populations, such as Europeans. This study will help us identify environmental and genetic factors that make the Emiratis at a higher risk of developing disease. Firstly, we will investigate how obesity, physical activity, and dietary intake can increase risk of T2DM in the UAE. Secondly, we will carry out genetic testing of Emiratis whole genomes to identify genetic variations that are found in subjects with T2DM but not in subjects free from diabetes. Additionally, in the genomes of subjects with T2DM, chemical modifications that can switch on or off genes critical in disease genesis will be identified. Results of this study will help in explaining the high risk of T2DM amongst the Emiratis, and improving disease diagnosis, especially in the UAE.

Why have I been chosen?

We are inviting Emirati patients who attend ICLDC to take part in this study, whether they have T2DM or are free from diabetes.

Do I have to take part?

It is your choice. If you do not take part, your care will not be affected in any way. If you do enter the study, you can withdraw at any time; you do not have to give any reason and there will be no penalty or disadvantage for you doing so, and your data and samples will be destroyed immediately.

What will happen if I take part?

You will be asked to sign a consent form, so that we have a written record of your agreement to participate. As part of your normal medical consultation, you are going to give the hospital some blood samples. We would like to take three more blood samples, which are about the size of your finger, at the same time. This means that no extra needle pricks will be required and you will only need to have blood taken all at once. Your blood samples will be used for genetic testing.

Additionally, you will be asked some questions about your physical activity and dietary intake. You will also be provided with activity monitors that will measure your activity for one week.

What will happen if incidental findings are discovered during the study?

Many of the discoveries of whole genome methods identify variations in genes that are responsible for very small increases in disease risk. Even for valid relationships between specific genes and disease, the significance for individuals can be very hard to interpret, especially for complex disorders, like T2DM, where very many genes, and environmental factors all have a part in the origin of disease. Thus, while this study can identify novel genes that can contribute to increased risk of T2DM amongst Emiratis, results are unlikely to have a direct impact on patients' current clinical management.

Would results of this study possibly identify subjects at risk of developing disease other than T2DM?

This study focuses on testing the relationship between variations in gene sequences with T2DM amongst the Emiratis. This will be based on identifying specific variations that are found in subjects with T2DM, but not in healthy subjects. Relationships of genes with other disease will not be tested.

How will my participation be kept confidential?

Your medical information will be protected by the confidential information policies of ICLDC. No information that could identify you will be released. We remove your name, date of birth and other information, and a code will be used on your sample for external testing to prevent whoever sees your sample from identifying you.

Who should I contact for future questions about the study?

Should you have any further questions please feel free to ask one of our staff members or you can call on (02) 4040800, extension 971, during our working hours.

Certificate of Consent

Genetic and epigenetic contribution to increased risk of T2DM in the UAE.

Patient Full Name		
Patient ICLDC Number		
Patient Contact Number		
Patient Date of Birth		
Patient Gender	☐ ☐**Male**	☐ ☐**Female**

If you would like to participate, please respond to each of the following questions and sign at the bottom of the page;

1. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.	☐**I agree**
2. I understand that I may be re-contacted by ICLDC (e.g. to answer some more questions and/or attend another assessment visit), but this is optional.	☐**I agree**
3. I give permission for access to my medical and other health-related records, and for long-term storage and use of this and other information about me, for health-related research purposes (even after my incapacity or death).	☐**I agree**
4. I give permission for long-term storage and use of any of my samples for health-related research purposes (even after my incapacity or death), and relinquish all rights to these samples, which I am donating for this study.	☐**I agree**
5. I understand that none of my results will be given to me (except for some measurements during this visit) and that I will not benefit financially from taking part (e.g. if research leads to commercial development of a new treatment).	☐**I agree**
6. I agree to take part in this study.	☐**I agree**

I have read the information, or it has been read to me. I have had the opportunity to ask questions about it and my questions have been answered to my satisfaction. I consent voluntarily to have my samples stored in the manner and for the purpose indicated above.

Patient/ Parent/Guardian Name		**Signature**	**Date**

I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understands the information.

I confirm that the participant was given an opportunity to ask questions and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

A copy of this consent and/or information sheet may be provided to the participant (If requested).

ICLDC Staff Name		**Signature**	**Date**

Appendix 2: Research Proposal (as presented at Early Stage Review)

1.1 Research problem and hypothesis

There is an increasing concern about the rising tide of T2D and its related comorbidities in the Gulf region in particular. Despite the significant association of established lifestyle risk factors with diabetes in the Gulf as presented in this review, populations in the region seem to have an increased susceptibility to developing diabetes. Indeed, various other regions in the world with similar environmental exposures, have much lower rates of diabetes. Whether a genetic component underlies this increased susceptibility is likely, but there is little data in the literature that supports this view.

In my research, I aim to investigate the contribution of environmental, genetic, and epigenetic factors to the increased risk of T2D in the Arab-Emirati population compared to Europeans. Results of this study will help us understand whether environmental exposures, including obesity and lack of physical activity, may explain the increased risk of T2D in the UAE, or whether Arab-Emiratis and Europeans respond to environmental exposures differently based on differences in the frequencies of genetic and/or epigenetic risk variants.

9.1 Aims and Objectives

1. Collect a sample of Arab-Emirati T2D cases and controls to enable investigation of genetic and environmental factors for T2D in the UAE.

Study design: case-control study

Subjects: 600 Arab-Emirati adults with T2D, and 600 age-sex matched controls free from T2D, recruited from Imperial College London Diabetes Centre (ICLDC) based in Abu Dhabi, UAE. Participants include men and women aged ≥18 years. They are provided with a complete information letter about the study and are asked to sign a written consent to be included in the study.

Methods: participants undergo health examination conducted by nurses and are interviewed by myself. Examinations include a brief medical history, and anthropometric measures. Blood samples are collected after an 8-hour fast for biochemical analysis, including measuring glucose, HbA1c, insulin, and lipid profile.

Diabetes diagnosis is defined using 1999 WHO criteria. Obesity is defined by BMI, with subjects classified as obese if they have a BMI ≥ 30 kg/m^2 and overweight if they have a 25 kg/ m^2 ≤ BMI < 30 kg/ m^2. Central obesity is defined by WHR > 0.9 for males and > 0.85 for females. Participants also fill the Global Physical Activity Questionnaire (GPAQ), which assesses their activity duration and intensity at work, transport, and recreational times during a typical week. A subset of 200 participants, 100 free from T2D and 100 with T2D, are asked to wear Actigraph activity monitors that measure their physical activity levels for 7 days.

Analysis: data will be examined for completeness, consistency, and response rates and for evidence of responder bias. Data collected using GPAQ will be examined following GPAQ guidelines [189]. Activity levels and durations will be categorized into; Low Physical Activity (LPA); Moderate Physical Activity (MPA); and High Physical Activity (HPA). Data collected using Actigraph GT9X is validated in ActiLife software. Participants were included in the analysis if they wore the monitors for 7 days including 2 weekend days, for at least 10 hours a day. Minutes per day of moderate to vigorous physical activity (MVPA) based on accelerometry will also be calculated in ActiLife software. These measures will be used to assess the validity of data collected using GPAQ in this sample population.

2. Examine the contribution of environmental risk factors to the increased risk of T2D in the UAE

Study design: case-control study

Subjects: the sample of Arab-Emirati T2D cases and controls collected (specific aim 1), as well as ~5,000 Europeans (including 3000 free from diabetes and 2000 with T2D) from London Life Sciences Prospective Population (LOLIPOP) Study.

Analysis: assess differences in BMI, waist and hip circumference, and physical activity levels among Arab-Emiratis and Europeans who are free from diabetes (or have diabetes). This allows us to investigate whether differences in the prevalence rates of T2D in Arab-Emirati and European subjects reported in the literature are paralleled with different rates of environmental risk factors. Furthermore, using logistic regression analysis, we will assess differences in the strength of association of overall obesity, central obesity, and physical inactivity with risk of T2D in Arab-Emiratis and Europeans, and assess differences between the samples. Results will allow us to infer whether environmental risk factors can explain at least part of the difference in T2D magnitude of risk between the Arab-Emiratis and Europeans.

3. Carry out genome-wide association analysis to identify novel and known genetic loci influencing, and contributing to increased risk of T2D in the UAE.

Study design: case-control study

Subjects: the cohort of Arab-Emirati T2D cases and controls collected (specific aim 1), as well as ~5,000 Europeans (Including 3000 free from diabetes and 2000 with T2D) from LOLIPOP Study. We are also going to explore the opportunity and feasibility of analysing 5300 from Northern Finland Birth Cohort (NFBC), which includes 800 with T2D. In addition, we are looking to include samples from ongoing cohort studies in the GULF region. These GWAS analyses are also conducted for training purposes. These results can be further meta-analysed with the existing data from publicly available databases, including the UK Biobank.

Analysis: allele frequencies will be compared between Arab-Emirati subjects with T2D and healthy controls to identify known and novel genetic loci that are associated with T2D in the UAE. The strength of association of genetic variants with T2D will be analysed with logistic regression under an additive genetic model to estimate effect sizes (ORs, CI 95%). The heterogeneity in allelic effects of lead SNPs identified across Arab-Emiratis and Europeans. Effect sizes will also be compared across the 2 populations to assess whether identified genetic variants can explain part of the increased risk of T2D among Arab-Emiratis. In addition, I will use linear regression to identify SNPs associated with continuous variables of fasting glucose levels, fasting insulin levels, triglycerides levels, insulin sensitivity or β-cell function, under an additive genetic model, and effect sizes will be compared to data from Europeans. Furthermore, the combined effect of genetic loci influencing similar T2D- related phenotypic traits will be tested across the two populations. This will be through evaluating the effect of carrying increasing number of risk alleles on a T2D intermediate phenotype in Arab-Emiratis compared to Europeans, and adjusting for case-control status. In addition, based on existing genetic biobanks, we will create a polygenic risk score that could be illustrative of the UAE population to assess the differences in risk of T2D between Arab-Emiratis and Europeans (UK – LOLIPOP and Finnish – NFBC). As suggested above then further meta-analyse with estimates from existing public databases.

International Journal of Obesity
https://doi.org/10.1038/s41366-021-00764-y

REVIEW ARTICLE

Epidemiology and population health

Secondary analyses of global datasets: do obesity and physical activity explain variation in diabetes risk across populations?

Budour Alkaf[1,2] · Alexandra I. Blakemore[3,4] · Marjo-Riitta Järvelin[2,3,5,6] · Nader Lessan[1]

Received: 21 May 2020 / Revised: 9 December 2020 / Accepted: 20 January 2021
© The Author(s) 2021. This article is published with open access

Abstract

Type 2 diabetes rates vary significantly across geographic regions. These differences are sometimes assumed to be entirely driven by differential distribution of environmental triggers, including obesity and insufficient physical activity (IPA). In this review, we discuss data which conflicts with this supposition. We carried out a secondary analysis of publicly available data to unravel the relative contribution of obesity and IPA towards diabetes risk across different populations. We used sex-specific, age-standardized estimates from Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) on diabetes (1980–2014) and obesity (1975–2016) rates, in 200 countries, and from WHO on IPA rates in 168 countries in the year 2016. NCD-RisC and WHO organized countries into nine super-regions. All analyses were region- and sex-specific. Although obesity has been increasing since 1975 in every part of the world, this was not reflected in a proportional increase in diabetes rates in several regions, including Central and Eastern Europe, and High-income western countries region. Similarly, the association of physical inactivity with diabetes is not homogeneous across regions. Countries from different regions across the world could have very similar rates of diabetes, despite falling on opposite ends of IPA rate spectrum. The combined effect of obesity and IPA on diabetes risk was analyzed at the worldwide and country level. The overall findings highlighted the larger impact of obesity on disease risk; low IPA rates do not seem to be protective of diabetes, when obesity rates are high. Despite that, some countries deviate from this overall observation. Sex differences were observed across all our analyses. Overall, data presented in this review indicate that different populations, while experiencing similar environmental shifts, are apparently differentially subject to diabetes risk. Sex-related differences observed suggest that males and females are either subject to different risk factor exposures or have different responses to them.

Introduction

Type 2 diabetes (T2D) is a complex metabolic disorder with worryingly increasing rates in several regions in the world. The development of T2D reflects a complex interplay of adverse lifestyle exposures with potentially increased genetic predisposition. The rapid worldwide increase in T2D rates is primarily a consequence of ageing in most populations, and the increase of obesity levels and lack of physical activity [1].

The link between obesity and T2D has been recognized as early as the 1920's [2]. Various measures of being overweight have repeatedly been shown to strongly associate with increased risk of T2D [3–6]. Similarly, the impact of physical activity on preventing and managing T2D has been established through cross-sectional and prospective studies. Overall, these studies have demonstrated the [1] independent association of physical inactivity with T2D risk, [2] association of increased sedentary behaviors

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41366-021-00764-y.

✉ Budour Alkaf
ba2013@ic.ac.uk

1 Imperial College London Diabetes Centre, Abu Dhabi, United Arab Emirates

2 Department of Epidemiology and Biostatistics, Imperial College London, London, UK

3 College of Health, Medicine, and Life Sciences, Brunel University London, Uxbridge, UK

4 Department of Medicine, Imperial College London, London, UK

5 Centre for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland

6 Unit of Primary Health Care and Medical Research Centre, Oulu University Hospital, Oulu, Finland

Published online: 11 February 2021

with elevated T2D risk, and [3] independent-beneficial effect of moderate and high occupational physical activity in reducing T2D risk [7–15]. These associations were largely based on individuals of White ethnicity.

T2D rates vary significantly between countries, geographic regions, and ethnic groups [16]. An ethnic group is a group of people who share cultural ideas and beliefs that have been a part of their community for generations, including language, religion, history, types of food, beliefs, and celebrations. People from the same ethnic group do not always share the same religion, and on the other hand, people from different ethnic groups may share the same religion, despite being from different cultures. Southwest Asia, for example, is home to many different ethnic groups who share similar religions.

Understanding difference in T2D risk across ethnic groups has been very challenging and not straightforward. Although the differential distribution of established behavioral and environmental exposures across geographic regions and ethnic populations can account for part of this variation, there is still a big part that remains unexplained. Part of the ethnic disparity in T2D risk is thought to be due to differences in glycemic control mechanisms between ethnicities [17, 18]. For example, Indian Asians are more insulin resistant than most other ethnic groups, while T2D pathogenesis among African Americans appears to be largely driven by β-cell dysfunction, rather than by insulin resistance [19, 20].

Preferential distribution of fat storage was also shown to contribute to difference in risk of developing T2D, both within populations of the same ethnicity, and across different ethnic groups [21–23]. For a given amount of body fat, individuals with excessive intra-abdominal or visceral adipose tissue have a substantially greater risk of being insulin resistant and developing T2D [24, 25]. Whereas, higher relative distribution of subcutaneous adipose tissue plays a protective role on T2D risk [26]. Fat storage distribution differs across ethnicities. Two populations may have identical BMI distributions, but could still have large differences in intra-abdominal fat accumulation [21–23]. The Asian Indian population, for example, seems to preferentially accumulate truncal fat, which has been proposed to explain the excessive insulin resistance and high T2D prevalence in this ethnic group, despite the absence of significant obesity as assessed by BMI [19, 20].

Differences in T2D risk across populations could also be explained by differential genetic susceptibility. Although a whole population may be subject to a changed environment, genetic factors influence how each individual within a population responds to that exposure: some people carry genetic variants that increase susceptibility to obesity and/or T2D. Such genetic susceptibility factors may be more common in some ethnic groups than others, and this may contribute to the variation in T2D risk observed between different populations. However, the prevalence of genetic risk factors will not have changed during the relatively short time period under consideration, so recent differences observed in populations of the same or similar ethnicity are likely to reflect environmental exposures. In particular, there are very different rates of obesity and physical activity between different populations and these may contribute to the observed differences in T2D risk.

Several trans-ethnic studies were directed to elucidate the impact of ethnicity on the association of age, BMI, and physical inactivity with T2D. Overall, their findings confirmed large ethnic, regional- and sex-dependent differences in age-adjusted prevalence of T2D, but also showed that ethnic-specific variations were not fully explained by differences in BMI between ethnic groups [21]. In addition, they reported ethnic-specific variation in T2D risk conferred by BMI and weight gain, with the biggest detrimental effects among Asians in comparison to other ethnic groups studied [27, 28]. Cross-sectional studies also demonstrated that differences in T2D rates between ethnic groups were not explained by different levels of physical activity [29–31]. Interestingly, reduced physical activity levels were not associated with T2D risk in several ethnic groups studied. A more recent prospective study supported these findings [32].

In spite of these efforts, there is still an underrepresentation of various ethnic populations in existing studies and a lack of studies that cover a wider range of ethnic populations. Networks and organizations like the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC) and World Health Organization (WHO) provide an opportunity to fill this research gap. Together, they provide worldwide rigorous and timely data on risk factors for noncommunicable diseases, including diabetes, obesity, and physical inactivity. In 2016 and 2017, NCD-RisC [33, 34] provided the lengthiest and most complete estimates of worldwide adult diabetes rates, and a complete picture of trends in mean BMI and prevalence rates of obesity for all countries in the world with the longest observation period. NCD-RisC diabetes estimates presented in this review are based on all diabetes cases worldwide, including people with type 1 diabetes (T1D) however, T2D accounts for roughly 90% of these cases. In addition to NCD-RisC data, in 2018, a group from WHO [35] provided the most complete description of global-, regional-, and country levels of insufficient physical activity, and presented regional and global trends over time.

In our secondary-analysis review, we aimed to compile publicly available estimates by NCD-RisC and WHO on diabetes, obesity, and insufficient physical activity rates and trends in adults, to unravel the relative contribution of obesity and physical inactivity on diabetes risk across different populations. We present and discuss region- and

sex-specific [1] changes in obesity and diabetes trends across 35–40 years, and [2] the association of obesity and insufficient physical activity with diabetes. We also analyzed country- and sex-specific joint effects of obesity and insufficient physical activity on diabetes risk.

We have conducted a systematic review as part of this review paper. Details are presented in supplementary information (Supplementary Table 1–3).

Materials and methods

Data retrieval: publicly available databases for estimates of diabetes, obesity, and insufficient physical activity rates

Region-, country-, and sex-specific estimates of [1] diabetes prevalence rates from 1980 to 2014, and [2] obesity prevalence rates from 1975 to 2016, were downloaded from www.ncdrisc.org. All estimates downloaded and used in this study covered males and females aged ≥18 years. Region-, country-, and sex-specific estimates of insufficient physical activity prevalence rates in 2016 were retrieved from a recent publication by a group from WHO [35].

Estimates of diabetes, obesity, and insufficient physical activity rates were available for 200 (diabetes and obesity) and 168 (insufficient physical activity) countries, respectively. Estimates were available as age-standardized prevalence rates with upper and lower 95% uncertainty intervals, and arranged by sex, country, region, and year. NCD-RisC and WHO organized countries into 21 regions and nine super-regions (Tables 1, 2), mostly on the basis of geography and national income. The exception was a region consisting of High-income English-speaking countries: cardiometabolic risk factors, including BMI, have very similar trends in these countries, which may be distinct from other countries within their geographic region [33–35]. Region definitions by NCD-RisC and WHO were equivalent. Region-specific analyses here were based on the nine super-regions, namely [1] Central Asia, Middle East and North Africa, [2] Central and Eastern Europe, [3] East and South East Asia, [4] High-income Asia Pacific, [5] High-income Western countries, [6] Latin America and Caribbean, [7] Oceania, [8] South Asia, and [9] Sub-Saharan Africa.

Description of Non-communicable Disease Risk Factor Collaboration (NCD-RisC) data sources

NCD-RisC is a worldwide network of health researchers and practitioners who aim to document systematically worldwide trends and variations in noncommunicable diseases (NCDs). Here we used yearly data for diabetes and obesity prevalence rates.

Diabetes was defined as fasting plasma glucose of 7.0 mmol/L or higher, history of diagnosis with diabetes, or use of insulin or oral hypoglycemic drugs [33]. To estimate diabetes prevalence rates, NCD-RisC identified, assessed, and reanalyzed population-based health-examination surveys that had measured at least one diabetes biomarker, and then converted diabetes prevalence in sources that had defined diabetes through 2hOGTT or HbA1c or used cutoff other than 7.0 mmol/L for fasting plasma glucose, to a corresponding prevalence based on the definition above [33]. Data sources included were representative of a national, subnational, or community population. Data sources that relied entirely on self-reported history of diagnosis were excluded. More detailed methods for identifying and accessing data sources were described in here [33].

Obesity was defined as BMI ≥ 30 kg/m^2. To estimate obesity prevalence rates in adults, NCD-RisC pooled and analyzed population-based studies that measured height and weight in adults aged 20 years and older [34]. They included data collected on samples of a national, subnational, or community populations, which had measured weight and height. Data sources that were based on self-reported weight and height were excluded, due to being subject to biases that vary by geography, time, age, sex, and socioeconomic characteristics [34].

The statistical models used by NCD-RisC to estimate prevalence by country, year, sex, and age are described in detail in their statistical paper and related substantive papers [33, 34, 36–38].

Description of World Health Organization (WHO) data sources

WHO estimated prevalence rates of insufficient physical activity in adults aged 18 years and older in 168 countries for three World Bank income groups, nine regions (Table 2), and globally from 2001 to 2016. Data were published in Lancet Global Health journal in 2018 [35]. Insufficient physical activity was defined as adults not meeting the WHO recommendations on physical activity for health—i.e., at least 150 min of moderate-intensity, or 75 min of vigorous-intensity physical activity per week, or any equivalent combination of the two.

Data sources included fulfilled the following criteria: [1] the survey questionnaire explicitly included physical activity across four key domains—i.e., work in the household (paid or unpaid), transport to get to and from places (i.e., walking and cycling), leisure time activities (i.e., sports and active recreation); [2] data were collected through random sampling with a sample size of at least 200, and were representative of a national or defined subnational

Table 1 List of countries included in NCD-RisC analyses to estimate diabetes and obesity prevalence rates, by super-region and region.

Super-region	Region	Number of countries	Countries
Sub-Saharan Africa	Central Africa	6	Angola, Central African Republic, Congo, DR Congo, Equatorial Guinea, Gabon
	East Africa	17	Burundi, Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Rwanda, Seychelles, Somalia, Sudan (former), Tanzania, Uganda, Zambia
	Southern Africa	6	Botswana, Lesotho, Namibia, South Africa, Swaziland, Zimbabwe
	West Africa	19	Benin, Burkina Faso, Cabo Verde, Cameroon, Chad, Cote d'Ivoire, Gambia, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Sao Tome and Principe, Senegal, Sierra Leone, Togo
Central Asia, Middle East and North Africa	Central Asia	9	Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Mongolia, Tajikistan, Turkmenistan, Uzbekistan
	Middle East and North Africa	19	Algeria, Bahrain, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Occupied Palestinian Territory, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, Tunisia, Turkey, United Arab Emirates, Yemen
South Asia	South Asia	6	Afghanistan, Bangladesh, Bhutan, India, Nepal, Pakistan
East and South East Asia	East Asia	4	China, China (Hong Kong SAR), North Korea, Taiwan
	South East Asia	12	Brunei Darussalam, Cambodia, Indonesia, Lao PDR, Malaysia, Maldives, Myanmar, Philippines, Sri Lanka, Thailand, Timor-Leste, Viet Nam
Oceania	Polynesia and Micronesia	13	American Samoa, Cook Islands, French Polynesia, Kiribati, Marshall Islands, Micronesia (Federated States of), Nauru, Niue, Palau, Samoa, Tokelau, Tonga, Tuvalu
	Melanesia	4	Fiji, Papua New Guinea, Solomon Islands, Vanuatu
High-income Asia Pacific	High-income Asia Pacific	3	Japan, Singapore, South Korea
Latin America and Caribbean	Andean Latin America	3	Bolivia, Ecuador, Peru
	Caribbean	18	Antigua and Barbuda, Bahamas, Barbados, Belize, Bermuda, Cuba, Dominica, Dominican Republic, Grenada, Guyana, Haiti, Jamaica, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago
	Central Latin America	9	Colombia, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Venezuela
	Southern Latin America	5	Argentina, Brazil, Chile, Paraguay, Uruguay
High-income Western countries	High-income English-speaking countries	6	Australia, Canada, Ireland, New Zealand, United Kingdom, United States of America
	North Western Europe	12	Austria, Belgium, Denmark, Finland, Germany, Greenland, Iceland, Luxembourg, Netherlands, Norway, Sweden, Switzerland
	South Western Europe	9	Andorra, Cyprus, France, Greece, Israel, Italy, Malta, Portugal, Spain
Central and Eastern Europe	Central Europe	13	Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Hungary, Macedonia (TFYR), Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia
	Eastern Europe	7	Belarus, Estonia, Latvia, Lithuania, Moldova, Russian Federation, Ukraine

Table 2 List of countries included in the analysis performed in Guthold R paper, by super-region.

Super-region	Number of countries	Countries
Central Asia, Middle East and North Africa	23	Algeria, Armenia, Egypt, Georgia, Iran (Islamic Republic of), Iraq, Jordan, Kazakhstan, Kuwait, Kyrgyzstan, Lebanon, Libya, Mongolia, Morocco, Oman, Qatar, Saudi Arabia, State of Palestine, Tajikistan, Tunisia, Turkey, United Arab Emirates, Uzbekistan
Central and Eastern Europe	17	Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Republic of Moldova, Romania, Russian Federation, Serbia, Slovakia, Slovenia, Ukraine
East and South East Asia	13	Brunei Darussalam, Cambodia, China, Indonesia, Lao People's Democratic Republic, Malaysia, Maldives, Myanmar, Philippines, Sri Lanka, Thailand, Timor-Leste, Viet Nam
High-income Asia Pacific	3	Japan, Republic of Korea, Singapore
High-income Western countries	24	Andorra, Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States of America
Latin America and Caribbean	25	Argentina, Bahamas, Barbados, Bermuda, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Grenada, Guatemala, Jamaica, Mexico, Paraguay, Saint Kitts and Nevis, Saint Lucia, Suriname, Trinidad and Tobago, Uruguay, Venezuela (Bolivarian Republic of)
Oceania	17	American Samoa, Cook Islands, Fiji, French Polynesia, Kiribati, Marshall Islands, Micronesia (Federated States of), Nauru, Niue, Palau, Papua New Guinea, Samoa, Solomon Islands, Tokelau, Tonga, Tuvalu, Vanuatu
South Asia	5	Bangladesh, Bhutan, India, Nepal, Pakistan
Sub-Saharan Africa	41	Benin, Botswana, Burkina Faso, Cabo Verde, Cameroon, Central African Republic, Chad, Comoros, Congo, Côte d'Ivoire, Democratic Republic of the Congo, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, São Tomé and Principe, Senegal, Seychelles, Sierra Leone, South Africa, Swaziland, Togo, Uganda, United Republic of Tanzania, Zambia, Zimbabwe

population[3]; prevalence of insufficient physical activity was reported by age and sex, according to current WHO [39], or former physical activity recommendations [35, 40]. Physical activity data collected using wearable devices were not included due to limited comparability with self-reported data.

The statistical models used by WHO to derive global and regional estimates were described in detail in their paper [35].

Current data analysis approaches

The following describes the current data analysis approaches used in this secondary-analysis review. Region- and sex-specific differences in diabetes prevalence rates and trends were assessed using scatter plots to compare the distribution of diabetes prevalence rates in males against that in females across 200 countries. This was done separately for the year 1980 and 2014, and for the change in diabetes prevalence throughout the 34 years (change in diabetes prevalence throughout the 34 years (1980–2014) studied = diabetes prevalence in 2014 − diabetes prevalence in 1980). These analyses were repeated to assess region- and sex-specific differences in obesity prevalence rates in 1975 and 2016, and the change in obesity prevalence throughout the 41 years (1975–2016) studied (change in obesity prevalence throughout the 41 years = obesity prevalence in 2016 − obesity prevalence in 1975). Region- and sex-specific correlation analyses of obesity with diabetes were based on obesity and diabetes prevalence rates in 200 countries in years 2016 and 2014, respectively. Region- and sex-specific correlation analyses of insufficient physical activity with diabetes were based on insufficient physical activity and diabetes prevalence rates in 163 countries in the year 2016 and 2014, respectively. Shapiro-Wild tests were performed to test for the normality of diabetes, obesity, and insufficient physical activity prevalence rate distribution. Correlation was tested using Pearson Correlation test. Significance was set at $p < 0.05$. Country- and sex-specific joint effects of obesity and insufficient physical activity prevalence rates with diabetes were based on estimates in 163 countries in years 2014 (diabetes) and 2016 (obesity and insufficient physical activity). Estimate of obesity, insufficient physical activity, and diabetes rates were each split into three tertiles. For each of obesity and insufficient physical activity rates, countries that fell in the first, second, and third tertiles were coded as low, moderate, and high, respectively. Diabetes risk was defined on the basis of the tertile split, i.e., countries that fell in the first tertile (with the lowest rates of diabetes) were coded as low risk, and those that fell in the second and third tertiles, were coded as moderate and high risk, respectively. All analyses were performed in Prism (Version 8).

Results

Region- and sex-specific trends in diabetes rates from 1980 to 2014

Based on our analysis of NCD-RisC data, in 2014, age-standardized diabetes prevalence was high in Oceania among both males (15.8%) and females (14.4%) (Fig. 1A, B). During the 34 years studied (1980–2014), Oceania and Central Asia, Middle East and North Africa regions were estimated to have the highest rates of diabetes across all years, particularly after 1985 (Fig. 1A, B). Sex-specific differences in diabetes prevalence increased with the years; in 1980, overall diabetes rates were almost identical across sexes (Supplementary Figure 1A). By 2014, diabetes rates were higher around the world, rising more rapidly in some countries/regions than others. Overall, increases in diabetes prevalence were higher in males compared to females (Supplementary Figure 1C). In both males and females, during the 34 years (1980–2014), the largest increase in diabetes prevalence was in Oceania and Central Asia, Middle East and North Africa, whereas the lowest increase in diabetes prevalence was in the High-Income Western countries region. There were sex differences between populations too: rates were higher in males, in High-income Western countries, and higher in females among countries in Latin America and Caribbean (Supplementary Figure 1B).

Region- and sex-specific trends in obesity rates from 1975 to 2016

Throughout the 41 years studied (1975–2016), obesity rates in males were high in High-income Western countries, and lowest in South Asia, particularly post-2000 (Fig. 1C). In females, obesity rates were high in Central and Eastern Europe in the first 15 years, but this increasing trend slowed after 1990 (Fig. 1D). Between 1990 and 2016, obesity rates in females were high in Central Asia Middle East and North Africa region, and low in High-income Asia Pacific region, particularly in the last 5 years (2010–2016) (Fig. 1D). Overall, females had higher rates of obesity compared to males across all regions both in 1975 and 2016 (Supplementary Figure 2A, B). In both males and females, during the 41 years (1975–2016), the largest increase in obesity prevalence was in Oceania, whereas the lowest increase in obesity prevalence was in South Asia and High-income Asia Pacific regions (Supplementary Figure 2C).

Contribution of obesity towards explaining differences in diabetes risk across populations

In order to assess region- and sex-specific correlation of obesity with diabetes, we used NCD-RisC estimates of

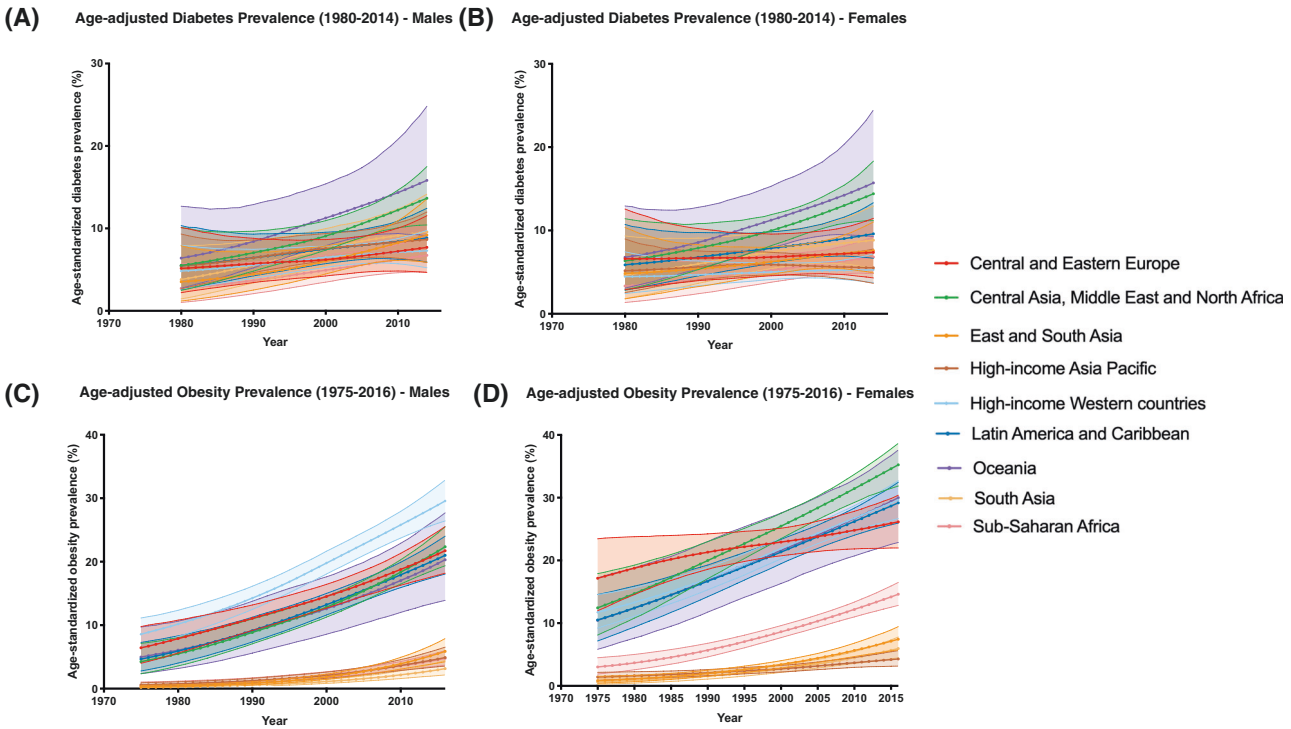


Fig. 1 Region and sex-specifc trends of diabetes and obesity in the last 35-40 years. A Age-adjusted diabetes prevalence (1980–2014) in males, **B** Age-adjusted diabetes prevalence (1980–2014) in females, **C** Age-adjusted obesity prevalence (1975–2016) in males, and **D** Age-adjusted obesity prevalence (1975–2016) in females. Age-standardized diabetes and obesity prevalence rates as estimated by Non-Communicable Disease risk Collaboration (NCD-RisC) group, for each year from 1980 to 2014, and 1975 to 2016, respectively. NCD-RisC group reported diabetes and obesity prevalence estimates in 200 countries, which were categorized into nine super-regions, including Central and Eastern Europe, Central Asia Middle East and North Africa, East and South East Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa.

diabetes prevalence rates in 2014 and obesity prevalence rates in 2016 for 200 countries in nine regions. In males, obesity rates were significantly correlated with diabetes rates in only four regions, with the strongest correlation in Oceania, followed by Central Asia Middle East and North Africa, East and South East Asia, and Sub-Saharan Africa (Table 3). Whereas in females, a significant correlation was observed across all regions, except for Central and Eastern Europe, and High-income Asia Pacific. As in males, the strongest correlation was observed in Oceania and Central Asia Middle East and North Africa regions (Table 3).

In line with these observations, a majority of countries with obesity prevalence rates lower than the worldwide mean (10.8% in males, 14.9% in females), also had lower diabetes prevalence rates compared to the worldwide mean (9% in males, 7.9% in females), and vice versa. This was observed in males and females (Fig. 2). Despite this well-established association, several countries, particularly in Latin America and Caribbean and Central and Eastern Europe regions yield surprising data: they have some of the highest rates of obesity, but still have some of the lowest rates of diabetes. When comparing diabetes and obesity rate trends between 1980 and 2016 (Supplementary Figure 3), the increasing obesity trends in some regions, like Central Asia, Middle East and North Africa reflect the increasing diabetes trend in the region. In other regions, however, including Central Eastern Europe and High-income Western countries regions, diabetes rates have been consistent during that time period, despite the increasing trends of obesity in these regions.

Contribution of insufficient physical activity towards explaining differences in diabetes risk across populations

To assess region- and sex-specific correlation of insufficient physical activity with diabetes, we extracted estimates from year 2016 of insufficient physical activity rates for 163 countries in nine regions, as reported by Guthold et al. [35]. Diabetes prevalence rates were based on NCD-RisC estimates in 2014. Prevalence rates of insufficient physical activity were higher in females compared to males across all regions (Supplementary Figure 4). In males, insufficient physical activity rates were significantly correlated with diabetes rates in only three regions, with the strongest correlation in Central Asia Middle East and North Africa,

Table 3 Pearson correlation test for region-specific correlation of obesity and insufficient physical activity (IPA) with diabetes.

Males	Obesity		IPA	
Region	R-squared	*P* value[†]	R-squared	*P* value[†]
Central and Eastern Europe	0.019	0.567	0.007	0.765
Central Asia Middle East and North Africa	0.747	<0.0001	0.486	0.0002
East and South East Asia	0.491	0.0025	0.281	0.062
High-income Asia Pacific	0.098	0.797	Too few pairs (only two countries)	
High-income Western Countries	0.134	0.06	0.243	0.014
Latin America and Caribbean	0.101	0.064	0.077	0.2
Oceania	0.827	<0.0001	0.134	0.015
South Asia	0.504	0.114	0.042	0.74
Sub-Saharan Africa	0.36	<0.0001	0.137	0.019
Females				
Region	R-squared	*P* value[†]	R-squared	*P* value[†]
Central and Eastern Europe	0.045	0.372	0.671	0.0001
Central Asia Middle East and North Africa	0.669	<0.0001	0.396	0.0013
East and South East Asia	0.619	0.0003	0.105	0.281
High-income Asia Pacific	0.575	0.452	Too few pairs (only two countries)	
High-income Western Countries	0.593	<0.0001	0.348	0.0024
Latin America and Caribbean	0.347	0.0002	0.003	0.816
Oceania	0.7	<0.0001	0.114	0.184
South Asia	0.737	0.029	0	0.96
Sub-Saharan Africa	0.55	<0.0001	0.106	0.041

IPA Insufficient physical activity.

[†]*P* value not corrected for multiple testing.

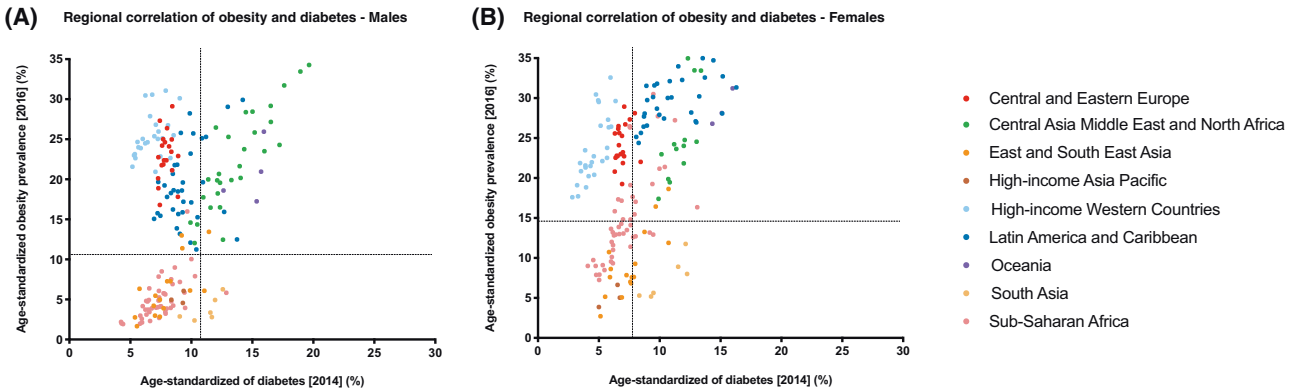


Fig. 2 Region- and sex-specific correlation of obesity with diabetes. **A** Males and **B** Females. The data presented was based on Non-Communicable Disease risk Collaboration (NCD-RisC) group estimates of diabetes and obesity, in 2014 and 2016, respectively. Scattergraphs A and B represent the age-standardized prevalence rates of diabetes in 2014 (*x*-axis) against obesity in 2016 (*y*-axis), across 200 countries, in males and females, respectively. For each graph, the vertical dotted line represents the worldwide age-standardized prevalence rate of diabetes in 2014, and the horizontal line represents the worldwide age-standardized prevalence rate of obesity in 2016. The 200 countries were categorized into nine super-regions and color-coded in the figure, which include Central and Eastern Europe, Central Asia Middle East and North Africa, East and South East Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa.

followed by High-income Western countries and Sub-Saharan Africa (Table 3). Whereas in females, the correlation of insufficient physical activity rates with diabetes rates was only significant in four regions, with the strongest correlation in Central and Eastern Europe, followed by Central Asia Middle East and North Africa, High-income Western countries, and Sub-Saharan Africa (Table 3).

Figure 3 demonstrates the correlation of insufficient physical activity with diabetes across all nine regions in males (Fig. 3A) and females (Fig. 3B). A majority of countries with insufficient physical activity rates lower than the worldwide mean (23.4% in males, 31.7% in females) also had lower diabetes rates compared to the worldwide mean (9% in males, 7.9% in females), and vice versa. This was observed in males and females (Fig. 3). Despite that, several countries, particularly in High-income Western countries and Central and Eastern Europe regions, have some of the highest rates of insufficient physical activity, but still have some of the lowest rates of diabetes. Other countries, in Oceania in particular, have some of the lowest rates of insufficient physical activity despite having some of the highest rates of diabetes in the world.

Joint effects of obesity and physical inactivity levels with diabetes

In order to assess country- and sex-specific joint effects of obesity and physical inactivity on diabetes, we analyzed the distribution of obesity and insufficient physical activity rates in 163 countries in relation to diabetes risk (Fig. 4). Overall, males and females across all countries have varying risk of diabetes (low, moderate, or high), irrespective of differences in their insufficient physical activity levels. Also, countries with obesity rates above 30% have high diabetes risk, irrespective of their insufficient physical activity rates. However, there is no such clear cutoff for insufficient physical activity (Fig. 4).

In spite of this overall picture, there are some countries that are outliers of this observation, and there seems to be a difference between sexes. A majority of males with high obesity rates have a high risk of diabetes; however, there are males in various countries with high diabetes risk despite having low or moderate rates of obesity. That is not the case in females, where the majority of countries with high obesity rates also have high risk of diabetes (Fig. 4). In addition, males but not females, in some countries of South Asia, including Bhutan, Bangladesh, and Nepal, have high diabetes risk despite having low rates of both obesity and insufficient physical activity (Table 4). In contrast to this scenario, males and females in some countries, particularly in High-income Western region, have a relatively low risk of diabetes despite having high rates of both obesity and insufficient physical activity (Table 4).

Discussion

The rapid increase of diabetes across the world, particularly in Westernized populations, has been associated with environmental changes related to obesity. Excessive caloric intake and reduced physical activity are the main features of a high-risk environment for diabetes development, and the relationship of these environmental stressors with disease risk was sometimes considered to be straightforward. Here, we present and discuss data which conflicts with this supposition. Although obesity

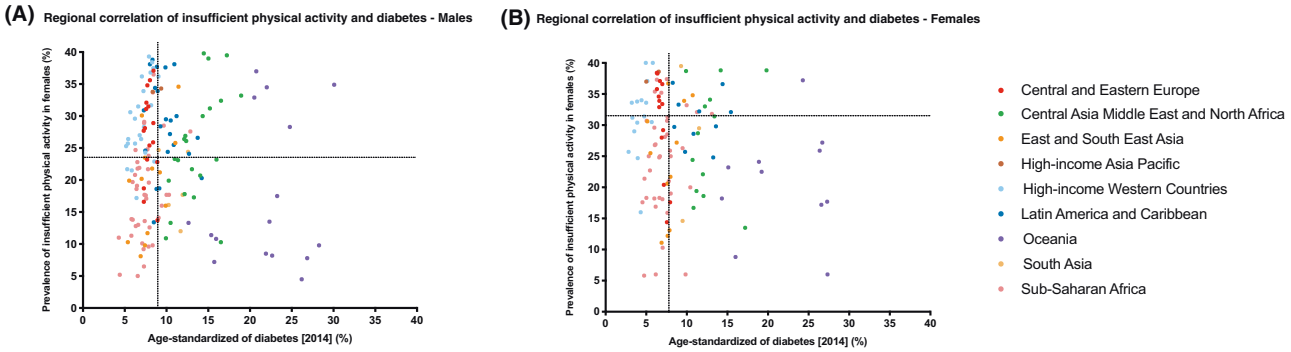


Fig. 3 Region- and sex-specific correlation of insufficient physical activity with diabetes. A Males and **B** Females. The data presented was based on diabetes prevalence rates estimates by Non-Communicable Disease risk Collaboration (NCD-RisC) group in 2014, and insufficient physical activity prevalence rates estimates by World Health Organization (WHO) in 2016. Scattergraphs A and B represent the age-standardized prevalence rates of diabetes in 2014 (*x*-axis) against insufficient physical activity 2016 (*y*-axis), across 163 countries, in males and females, respectively. For each graph, the vertical dotted line represents the worldwide age-standardized prevalence rate of diabetes in 2014, and the horizontal line represents the worldwide age-standardized prevalence rate of insufficient physical activity in 2016. The 163 countries were categorized into nine super-regions and color-coded in the figure, which include Central and Eastern Europe, Central Asia Middle East and North Africa, East and South East Asia, High-income Asia Pacific, High-income Western countries, Latin America and Caribbean, Oceania, South Asia, and Sub-Saharan Africa.

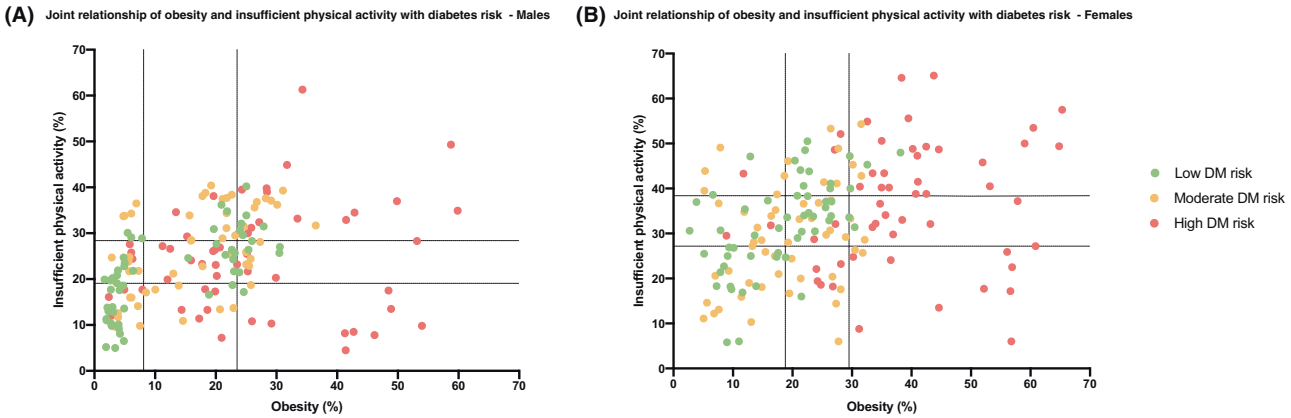


Fig. 4 Joint effects of obesity and insufficient physical activity levels with diabetes risk. A Males and **B** Females. The diabetes and obesity rates were estimated by Non-Communicable Disease Risk Collaboration (NCD-RisC) group in 2014 and 2016, respectively. Insufficient physical activity rates were estimated by World Health Organization (WHO) in 2016. Scattergraphs represent distribution of age-standardized prevalence rates of obesity (*x*-axis) against insufficient physical activity rates in 163 countries. Estimates of obesity, insufficient physical actiivty, and diabetes were grouped into tertiles. For diabetes rates, countries that fell in the first tertile (with the lowest rates of diabetes) were coded as low risk (green), and those that fell in the second and third tertiles, were coded as moderate (yellow) and high (red) risk, respectively. Dotted vertical and horizontal lines represent the first and second tertiles of obesity and insufficient physical activity rates, respectively.

has been increasing since 1980 in every part of the world, this was not reflected in a proportional increase in diabetes rates in several regions, including Europe and High-income Western Countries. These data suggest that certain populations, such as those in the Middle East, are exposed to similar environmental stressors to some Westernized populations, yet have much higher rates of diabetes. Our findings, however, must be interpreted with caution. Obesity rates reported here are based on the universal BMI cut-off point of ≥30 kg/m^2, and that could have resulted in underestimating obesity levels in some regions, and unreliable comparisons with other regions.

BMI is a simple and noninvasive method for assessing obesity and excessive fat stores within a single population, but less reliable when comparing adiposity measures between populations [41, 42]. One of the largest studies directed at the elucidation of the impact of ethnicity on the association of age and BMI with T2D was conducted by the Diabetes Epidemiology Collaborative analysis of Diagnostic criteria in Europe-Diabetes Epidemiology Collaborative analysis of Diagnostic criteria in Asia (DECODE-DECODA) study [21]. Nakagami et al. [21] analyzed studies/surveys representing different ethnic groups, including European, Maltese, Indian, Chinese, and Japanese populations. They found that the

Table 4 Countries that fall as outliers of the positive association of obesity and insufficient physical activity (IPA) with diabetes risk in males and females.

Country	Region	Obesity	IPA	Diabetes
Males				
Bhutan	South Asia	Low	Low	High
Bangladesh	South Asia	Low	Low	High
Nepal	South Asia	Low	Low	High
Seychelles	Sub-Saharan Africa	Low	Low	High
Cabo Verde	Sub-Saharan Africa	Low	Low	Moderate
Comoros	Sub-Saharan Africa	Low	Low	Moderate
Gambia	Sub-Saharan Africa	Low	Low	Moderate
China	East and South East Asia	Low	Low	Moderate
Lao PDR	East and South East Asia	Low	Low	Moderate
Germany	High-income Western Countries	High	High	Low
Greece	High-income Western Countries	High	High	Low
Belgium	High-income Western Countries	High	High	Low
Norway	High-income Western Countries	High	High	Low
Ireland	High-income Western Countries	High	High	Low
United Kingdom	High-income Western Countries	High	High	Low
Romania	Central and Eastern Europe	High	High	Low
Females				
Nepal	South Asia	Low	Low	Moderate
Togo	Sub-Saharan Africa	Low	Low	Moderate
Cambodia	East and South East Asia	Low	Low	Moderate
China	East and South East Asia	Low	Low	Moderate
Myanmar	East and South East Asia	Low	Low	Moderate
Malta	High-income Western Countries	High	High	Low
United Kingdom	High-income Western Countries	High	High	Low
United States	High-income Western Countries	High	High	Low
New Zealand	High-income Western Countries	High	High	Low

Age-standardized prevalence rates of obesity, IPA, and diabetes rates in 163 countries were each split into tertiles. For each of obesity and insufficient physical activity rates, countries that fell in the first, second, and third tertiles were color-coded as low (green), moderate (yellow), and high (red), respectively. For diabetes rates, countries that fell in the first tertile (with the lowest rates of diabetes) were color-coded as low risk (green), and those that fell in the second and third tertiles, were color-coded as moderate (yellow) and high risk (red), respectively. Code for interpreting the table: Bhutan in South Asia has low obesity (green) and low IPA (green), but high diabetes risk (red).

IPA, Insufficient physical activity.

variation in T2D prevalence across ethnic groups cannot be explained by differences in BMI between populations alone, and that ethnicity modified the effect of BMI on age-adjusted prevalence of T2D [21]. These findings support the need for region- and ethnic-specific definitions of obesity.

The heterogeneity in the relationship between obesity (as measured by BMI) and diabetes observed here could also be explained by differences in the contribution of reduced insulin secretion and reduced insulin sensitivity to the development of diabetes between various populations. Patterns of fat distribution differ between ethnic populations [19, 43–45], and, thus, two populations could have identical BMI distributions, but still show large differences with respect to the accumulation of intra-abdominal fat. South Asian people, for example, are more insulin resistant and more prone to abdominal obesity and low muscle mass than other ethnic groups [19, 43–45]. This suggests that excessive insulin resistance may be related to intra-abdominal fat deposition and low muscle mass among South Asians: this could explain their increased risk of developing diabetes, despite the absence of significant obesity as assessed by BMI. This is in accordance with the data analyzed here. Obesity and physical inactivity in South Asia did not correlate with diabetes in the region overall: countries including Bhutan, Bangladesh, and Nepal were estimated to have some of the highest diabetes rates in the world despite their low obesity and activity rates.

Sex differences were observed across all our analyses. Previous studies, largely based on Europeans, showed that diabetes is more frequently diagnosed at a lower age and BMI in males, despite the fact that obesity is more common

in females [46]. This trend was also observed here, not only in European populations (Central and Eastern Europe and High-income Western countries regions), but across all regions. Diversities in culture, lifestyle, and socioeconomic status are some of the factors that may impact differences between males and females in predisposition, development, and clinical presentation of diabetes [46]. Besides obesity, other risk factors, including genetic effects and epigenetic mechanisms, nutritional factors, and sedentary behaviors affect disease risk and complications differently in both sexes [46]. Indeed, in this study, females had higher insufficient physical activity rates compared to males, yet they still had similar rates of diabetes.

The secondary analyses presented here were based on two of the most well-established risk factors for T2D and their contribution to differences in disease risk across populations. It is important to note, however, that the development of T2D involves a complex interplay of adverse lifestyle and genetic exposures. This means that other factors could be confounding our findings, particularly those that are also associated with cultural and faith/religious groups: smoking, alcohol consumption, and nutrition, for example. WHO produces yearly estimates (2007–2018) of cigarette and tobacco smoking prevalence rates in more than 190 countries. Secondary analyses that account for the effect of factors like smoking, diet and nutrition, and socioeconomic status on our findings are urgently required.

Current efforts are largely focused on reducing the global health and economic burden of diabetes and on developing strategies to help prevent this "epidemic" [47–51]. These have provided guidelines as to who should be targeted in the preventive process. Our data showing that there are differences across sexes and different populations on the consequences of obesity and physical inactivity suggest that target groups for intervention strategies may have to be defined for each sex and ethnic group separately, and future studies should be aimed at defining these high risk groups.

Acknowledgements The authors had full access to all the data and had final responsibility for the decision to submit for publication. The work presented in this paper is part of a PhD project by the first author (Budour Alkaf); internal funding by Imperial College London Diabetes Centre is gratefully acknowledged. The authors are also grateful for The Medical Research Council, UK (grant numbers: MR/M013138/1, MRC/BBSRC MR/S03658X/1 (JPI HDHL H2020)) for their financial support during the authors time when conducting this piece of work and writing this paper.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

References

1. King H, Rewers M. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. WHO Ad Hoc Diabetes Reporting Group. Diabetes Care. 1993;16:157–77.
2. Joslin EP. The prevention of diabetes mellitus. J of the Am Med Assoc. 1921;76:79–84.
3. Colditz GA, Willett WC, Stampfer MJ, Manson JE, Hennekens CH, Arky RA, et al. Weight as a risk factor for clinical diabetes in females. Am J Epidemiol. 1990;132:501–13.
4. Cassano PA, Rosner B, Vokonas PS, Weiss ST. Obesity and body fat distribution in relation to the incidence of non-insulin-dependent diabetes mellitus. A prospective cohort study ofmales in the normative aging study. Am J Epidemiol. 1992;136:1474–86.
5. Chan JM, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in males. Diabetes Care. 1994;17:961–9.
6. Colditz GA, Willett WC, Rotnitzky A, Manson JE. Weight gain as a risk factor for clinical diabetes mellitus in females. Ann Intern Med. 1995;122:481–6.
7. Helmrich SP, Ragland DR, Leung RW, Paffenbarger RS Jr. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. N Engl J Med. 1991;325:147–52.
8. Manson JE, Rimm EB, Stampfer MJ, Colditz GA, Willett WC, Krolewski AS, et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in females. Lancet. 1991;338:774–8.
9. Manson JE, Nathan DM, Krolewski AS, Stampfer MJ, Willett WC, Hennekens CH. A prospective study of exercise and incidence of diabetes among US male physicians. JAMA. 1992;268:63–7.
10. Burchfiel CM, Sharp DS, Curb JD, Rodriguez BL, Hwang LJ, Marcus EB, et al. Physical activity and incidence of diabetes: the Honolulu Heart Program. Am J Epidemiol. 1995;141:360–8.
11. Perry IJ, Wannamethee SG, Walker MK, Thomson AG, Whincup PH, Shaper AG. Prospective study of risk factors for development of non-insulin dependent diabetes in middle aged Britishmales. BMJ. 1995;310:560–4.
12. Lynch J, Helmrich SP, Lakka TA, Kaplan GA, Cohen RD, Salonen R, et al. Moderately intense physical activities and high levels of cardiorespiratory fitness reduce the risk of non-insulin-dependent diabetes mellitus in middle-agedmales. Arch Intern Med. 1996;156:1307–14.
13. Hu FB, Sigal RJ, Rich-Edwards JW, Colditz GA, Solomon CG, Willett WC, et al. Walking compared with vigorous physical activity and risk of type 2 diabetes in females: a prospective study. Jama. 1999;282:1433–9.
14. Hu FB, Li TY, Colditz GA, Willett WC, Manson JE. Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in females. Jama. 2003;289:1785–91.

15. Hu G, Qiao Q, Silventoinen K, Eriksson JG, Jousilahti P, Lindström J, et al. Occupational, commuting, and leisure-time physical activity in relation to risk for Type 2 diabetes in middle-aged Finnishmales and females. Diabetologia. 2003;46:322–9.
16. International Diabetes Federation. IDF Diabetes Atlas. 9th edn. Brussels, Belgium: International Diabetes Federation; 2017.
17. Abate N, Chandalia M. The impact of ethnicity on type 2 diabetes. J Diabetes Complications. 2003;17:39–58.
18. Dagogo-Jack S. Ethnic disparities in type 2 diabetes: pathophysiology and implications for prevention and management. J Natl Med Assoc. 2003;95:774.
19. Abate N, Chandalia M. Ethnicity and type 2 diabetes: focus on Asian Indians. J Diabetes Complications. 2001;15:320–7.
20. Banerji MA, Faridi N, Atluri R, Chaiken RL, Lebovitz HE. Body composition, visceral fat, leptin, and insulin resistance in Asian Indianmales. The. Journal of clinical endocrinology and metabolism. 1999;84:137–44.
21. Nakagami T, Qiao Q, Carstensen B, Nhr-Hansen C, Hu G, Tuomilehto J, et al. Age, body mass index and Type 2 diabetes-associations modified by ethnicity. Diabetologia. 2003;46:1063–70.
22. Després JP, Nadeau A, Tremblay A, Ferland M, Moorjani S, Lupien PJ, et al. Role of deep abdominal fat in the association between regional adipose tissue distribution and glucose tolerance in obese females. Diabetes. 1989;38:304–9.
23. Abate N, Garg A, Peshock RM, Stray-Gundersen J, Grundy SM. Relationships of generalized and regional adiposity to insulin sensitivity in males. J Clin Invest. 1995;96:88–98.
24. Després JP, Moorjani S, Lupien PJ, Tremblay A, Nadeau A, Bouchard C. Regional distribution of body fat, plasma lipoproteins, and cardiovascular disease. Arteriosclerosis. 1990;10:497–511.
25. Després JP. Is visceral obesity the cause of the metabolic syndrome? Ann Med. 2006;38:52–63.
26. Golan R, Shelef I, Rudich A, Gepner Y, Shemesh E, Chassidim Y, et al. Abdominal superficial subcutaneous fat: a putative distinct protective fat subdepot in type 2 diabetes. Diabetes Care. 2012;35:640–7.
27. Shai I, Jiang R, Manson JE, Stampfer MJ, Willett WC, Colditz GA, et al. Ethnicity, obesity, and risk of type 2 diabetes in females: a 20-year follow-up study. Diabetes Care. 2006;29:1585–90.
28. Ramachandran A, Snehalatha C, Viswanathan V, Viswanathan M, Haffner SM. Risk of noninsulin dependent diabetes mellitus conferred by obesity and central adiposity in different ethnic groups: a comparative analysis between Asian Indians, Mexican Americans and Whites. Diabetes Res Clin Pract. 1997;36:121–5.
29. Ujcic-Voortman JK, Schram MT, Jacobs-van der Bruggen MA, Verhoeff AP, Baan CA. Diabetes prevalence and risk factors among ethnic minorities. Eur J Public Health. 2009;19:511–5.
30. Admiraal WM, van Valkengoed IG, LdM JS, Stronks K, Hoekstra JB, Holleman F. The association of physical inactivity with Type 2 diabetes among different ethnic groups. Diabet Med. 2011;28:668–72.
31. Jenum AK, Holme I, Graff-Iversen S, Birkeland KI. Ethnicity and sex are strong determinants of diabetes in an urban Western society: implications for prevention. Diabetologia. 2005;48:435–9.
32. Joseph JJ, Echouffo-Tcheugui JB, Golden SH, Chen H, Jenny NS, Carnethon MR, et al. Physical activity, sedentary behaviors and the incidence of type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA). BMJ Open Diabetes Res Care. 2016;4:e000185.
33. NCD Risk Collaboration (NCD-RisC). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387:1513–30.
34. NCD Risk Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet. 2017;390:2627–42.
35. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health. 2018;6:e1077–e86.
36. NCD Risk Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19·2 million participants. Lancet. 2016;387:1377–96.
37. NCD Risk Collaboration (NCD-RisC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. Lancet. 2017;389:37–55.
38. Finucane MM, Paciorek CJ, Danaei G, Ezzati M. Bayesian estimation of population-level trends in measures of health status. Statistical Science. 2014;29:18–25.
39. WHO. Global recommendations on physical activity for health. Geneva: World Health Organization; 2010.
40. Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. Jama. 1995;273:402–7.
41. Norgan NG. Population differences in body composition in relation to the body mass index. Eur J Clin Nutr. 1994;48:S10–25. discussion S6-7
42. Deurenberg P, Deurenberg Yap M, Wang J, Lin FP, Schmidt G. The impact of body build on the relationship between body mass index and percent body fat. Int J Obes Relat Metab Disord. 1999;23:537–42.
43. McKeigue PM, Pierpoint T, Ferrie JE, Marmot MG. Relationship of glucose intolerance and hyperinsulinaemia to body fat pattern in south Asians and Europeans. Diabetologia. 1992;35:785–91.
44. Karter AJ, Mayer-Davis EJ, Selby JV, D'Agostino RB Jr, Haffner SM, Sholinsky P, et al. Insulin sensitivity and abdominal obesity in African-American, Hispanic, and non-Hispanic whitemales and females. The insulin resistance and atherosclerosis study. Diabetes. 1996;45:1547–55.
45. Wang J, Thornton JC, Russell M, Burastero S, Heymsfield S, Pierson RN Jr. Asians have lower body mass index (BMI) but higher percent body fat than do whites: comparisons of anthropometric measurements. Am J Clin Nutr. 1994;60:23–8.
46. Kautzky-Willer A, Harreiter J, Pacini G. Sex and gender differences in risk, pathophysiology and complications of type 2 diabetes mellitus. Endocr Rev. 2016;37:278–316.
47. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med. 2002;346:393–403.
48. Gregg EW, Albright AL. The public health response to diabetes–two steps forward, one step back. Jama. 2009;301:1596–8.
49. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med. 2001;344:1343–50.
50. Ramachandran A, Snehalatha C, Mary S, Mukesh B, Bhaskar AD, Vijay V. The indian diabetes prevention programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). Diabetologia. 2006;49:289–97.
51. Saaristo T, Moilanen L, Korpi-Hyövälti E, Vanhala M, Saltevo J, Niskanen L, et al. Lifestyle intervention for prevention of type 2 diabetes in primary health care: one-year follow-up of the Finnish National Diabetes Prevention Program (FIN-D2D). Diabetes Care. 2010;33:2146–51.