Comparison of body mass index at diagnosis of diabetes in a multiethnic population: a case-control study with matched non-diabetic controls

Sanjoy K. Paul [1,2*], Ebenezer S. Owusu Adjah [1, 3*], Mayukh Samanta[1], Kiran Patel [7, 8], Srikanth Bellary[6], Wasim Hanif [5], Kamlesh Khunti [4]

[1]Clinical Trials and Biostatistics Unit, QIMR Berghofer Medical Research Institute, Brisbane, Australia;

[2]Melbourne EpiCentre, University of Melbourne and Melbourne Health, Melbourne, Australia

[3] School of Medicine, The University of Queensland, Brisbane, Australia;

[4] Department of Health Sciences, University of Leicester, Leicester, UK

[5] Department of Diabetes, University Hospital Birmingham, UK

[6] Aston Research Centre for Healthy Ageing, Aston University, Birmingham, UK

[7] University of Warwick, Warwick, UK

[8] Heart of England NHS Trust, UK

[*]Joint first authors

Correspondence to:

Prof. Sanjoy K. Paul
[1]Clinical Trials and Biostatistics Unit, QIMR Berghofer Medical Research Institute
300 Herston Road, Brisbane, Australia
Email: Sanjoy.Paul@mh.org.au

Number of Words in Abstract: 252
Number of Words in-Text: 3226
Number of Tables: 2
Number of Figures: 2
Number of References: 28

ABSTRACT

Aims: To investigate the probability of developing type 2 diabetes mellitus (T2DM) at different body mass index compared to matched non-diabetic controls in a multi-ethnic population.

Materials and Methods: Case-control study of 90,367 patients with incident diabetes and 362,548 age-sex-ethnicity matched controls from UK primary care. Probability of developing T2DM was estimated.

Results: Case and control patients were 56 years old at index and 56% were male. Patients with T2DM had significantly higher mean BMI level by about 5 kg/m^2 at diagnosis (32.2 kg/m^2) ,compared to the matched controls (27.4 kg/m^2). White European (n=79,270), African-Caribbean (n=4,115) and South Asians (n=7,252) were 58, 48, and 46 years old with mean BMI of 32.5, 31.1, 29.2 kg/m^2 respectively at diagnosis. More South Asians developed T2DM at BMI below 30 kg/m^2 (38%) than White Europeans (26%) and African-Caribbeans (29%), (all $p<0.01$). Within the 18-70 year age range, South Asian male and female had significantly higher probability of developing diabetes in the continuously measured BMI range of 18-30 kg/m^2, compared to White Europeans and African-Caribbeans. Across all age groups < 70 years, South Asians and African-Caribbeans had significantly higher probability of developing T2DM in the normal weight and overweight categories, compared to White Europeans. However, this risk patterns of developing diabetes was reversed amongst the obese at all age groups.

Conclusion: Risk patterns of developing diabetes at different levels of obesity varies between ethnic groups across all age groups, while South Asians and African-Caribbeans carry the highest risk at younger age and at lower adiposity burden.

Key Words: Body Mass Index, Probability of Developing Type 2 Diabetes, Matched Case-control Study, Multi-ethnic Comparisons

INTRODUCTION

Obesity [body mass index (BMI) ≥ 30 kg/m^2] is a worldwide epidemic affecting people of all ages, and is a major risk factor for type 2 diabetes mellitus (T2DM) and cardiovascular diseases (CVD) [1]. Current epidemiological indices of obesity have doubled since 1980, with about 13% of the world's population being obese as of 2014 [1, 2]. Some population-based studies have been conducted to assess the impact of BMI classification, including overweight and various grades of obesity, on the risk of T2DM [3-5]. The BMI cut-points of 25 kg/m^2 and 30 kg/m^2 were defined as the basis for identifying overweight and obesity based on epidemiological studies investigating association with mortality and morbidity, primarily in White population. However, there is increasing evidence that levels of risk associated with the classification of overweight and obesity vary across ethnic groups [6-11].

Propensity to develop T2DM varies considerably between ethnic groups and identifying the BMI cut points within specific ethnic groups at which the risk of T2DM increases is useful to inform public health policy. Some studies have evaluated the ethnicity-specific diabetes incidence rates in association with prior BMI levels using population level and primary care data [7, 9, 12, 13]. These studies were limited by the number of subjects in the non-white ethnic groups. The pooled analysis of survey data from various countries conducted by the DECODE-DECODA study group in 2003 evaluated the association of ethnicity, BMI and prevalence of T2DM [13]. However, the BMI measurements were not consistently taken at the time of diagnosis of diabetes. Only one previous study has compared the distribution of BMI at diagnosis of T2DM with non-diabetic controls [5]. Ganz and colleagues defined BMI at diagnosis as the last measurement of BMI taken within one year prior to diagnosis of T2DM, and randomly matched controls to cases [5]. While this study reported increased risk of

developing T2DM with higher BMI levels, the differential aspects of ethnicity in the relationship between BMI and risk were not addressed.

No study has compared the distribution of BMI at diagnosis of T2DM by ethnicity with non-diabetic controls. In addition, we are not aware of any population-based study evaluating differences in the risk of developing T2DM in men and women in different age levels between different ethnic groups over the whole spectrum of BMI distribution at diagnosis. Using a large cohort of incident T2DM patients, and an age-sex-ethnicity matched non-diabetic control cohort from United Kingdom primary care, the aims of this study were to evaluate for each ethnic group (1) the distribution of BMI, glycaemic, and vascular risk factors at diagnosis of T2DM, and (2) the probability of developing T2DM over the entire spectrum of BMI and age.

MATERIALS AND METHODS

Data source

Data for this study were obtained from The Health Improvement Network (THIN) database, a large anonymised longitudinal dataset derived from a network of more than 600 primary care providers across the United Kingdom. With longitudinal data on approximately 11 million individuals registered with the primary care system, the THIN database has been extensively used for academic research in various disciplines [14]. The accuracy and completeness of this database has been previously described [15, 16]. Notably, the database has similar distribution of major chronic diseases including diabetes, heart failure and obesity when compared to UK national statistics [16, 17]. Clinically diagnosed diseases are recorded using Read codes [18] and with each diagnosis, an event date is entered. THIN database provides comprehensive patient-level longitudinal information on demographic, anthropometric, clinical and laboratory

measures, clinical diagnosis of diseases / events, along with complete information on prescriptions for medications with dates and doses. Formal access to the database has been obtained and the Scientific Review Committee of the THIN database, UK (reference number: 15THIN030).

Identification of T2DM cases

Patients with T2DM were identified through various steps of clinically guided iterative processes. Specifically, the T2DM cases were selected if:

(i) Patient had a record of Read code related to T2DM,

(ii) Patient from step (i) above had received at least one prescription for an antidiabetic drug in addition to the clinical diagnosis, or

(iii) Patient in step (i) above had received a lifestyle modification intervention.

A set of 345,013 patients with newly diagnosed T2DM (from January 1990 to September 2014) was identified, who had complete information on age at diagnosis (≥18 years) and sex. Of these patients only 90,754 patients had their ethnicity identified as White European, African-Caribbean or South Asian, (Supplementary Figure 1). South Asians were defined as patients with Indian, Pakistan, Sri Lanka, and Bangladesh origin while African-Caribbeans were defined as patients with Black-African and/or Caribbean origin. White Europeans were patients with self-reported ethnicity as White, European, Caucasian, and/or New Zealand European.

Development of control subjects

A control pool of patients without T2DM was obtained by selecting individuals who had no diagnosis of any type of diabetes and had never received an antidiabetic prescription. Exact matches based on ethnicity, age, and sex were obtained from these pool of potential controls without replacement. To the 90,637 eligible T2DM cases with identified three ethnic groups, 362,548 controls were successfully matched in a 1:4 ratio. Index date for controls was defined as the date of the diabetes diagnosis for their matched cases.

The following information on index date was extracted for all patients where available: smoking status, deprivation score, weight, BMI, glycated haemoglobin (HbA_{1c}), systolic blood pressure (SBP), diastolic blood pressure (DBP), low density lipoproteins (LDL-C), high density lipoproteins (HDL-C), and triglycerides. All available measures on or within 3 months prior to the index date were considered as the baseline measures. Anti-glycaemic agents, anti-hypertensive agents, cardio-protective medications (CPM), weight lowering drugs and anti-depressants were also obtained along with dates of prescription. The CPMs were defined as use of statins or angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers or beta blockers on or before diagnosis. BMI categories were defined following WHO established criteria [19] as follows: normal weight (18.5-24.99 kg/m^2), overweight (25-29.99 kg/m^2), Grade 1 obese (30-34.99 kg/m^2), Grade 2 obese (35-39.99 kg/m^2) and Grade 3 obese ($\geq$ 40 kg/m^2).In addition, records of cardiovascular diseases (CVDs), renal diseases and cancer on or before index date were also obtained. A composite variable for CVD (any CVD) was defined as the occurrence of angina or myocardial infarction or coronary artery disease (including bypass surgery and angioplasty) or heart failure or stroke before diagnosis.

Statistical analysis

Basic characteristics of incident T2DM patients and their matched non-diabetic control population, separately for ethnic group, were summarized using number (%), means ± SD or median (first quartile, third quartile) as appropriate. Differences between patients with T2DM and their matched controls were evaluated using rank sum test for continuous variables and the chi-squared for binary data. The distributions of BMI at diagnosis of T2DM in three different ethnic groups were compared using analysis of variance models, separately for different age groups at diagnosis.

To explore the association of BMI at diagnosis with the risk of developing T2DM, in interaction with different ethnic groups, multivariate logistic regression models were fitted. The covariates for adjustments were age, sex, smoking status, deprivation score, and the history of CVD, cancer, and chronic kidney disease (CKD) on or prior to the index date. The probability of developing T2DM over the whole distribution of BMI in different ethnic groups was evaluated, using both continuous measure of BMI and the World Health Organisation defined categories of BMI. To explore the possible differences in the patterns of association of BMI with the risk of developing T2DM in different ethnic groups for male and female, and also over different age groups at index date, separate adjusted models were fitted. The differences in predicted probabilities between ethnic groups were calculated using the methodology described by King and colleagues (2000) [20] and Zelner (2009) [21]. The estimated probabilities and the 95% confidence intervals were presented as appropriate. Sensitivity analyses to support the above analyses include (1) an extended model incorporating measures of SBP, LDL-C, HDL-C and triglyceride at index, the use of CPMs, weight lowering drugs, anti-hypertensives, and anti-depressants before diagnosis, and (2) comparison of distribution of BMI and HbA1c at diagnosis for patients diagnosed after 01 Jan 2006.

RESULTS

The basic demographic and clinical profiles of 90,367 patients with T2DM and 362,548 age-sex-ethnicity matched controls, separately for ethnic groups, are shown in Table 1. Case and control patients were 56 years old at index and 56% were male. Patients with T2DM had significantly higher mean BMI level by about 5 kg/m^2 at diagnosis (32.2 kg/m^2) , compared to the matched controls (27.4 kg/m^2). However, this difference was smaller in the African-Caribbean (2.4 kg/m^2) and South Asians (2.8 kg/m^2). Furthermore, cases were more likely to receive anti-hypertensives, cardio-protective medications, and anti-depressants and were more likely to have any CVD before diagnoses than controls. South Asians were more likely to develop diabetes at a significantly lower age (mean age 46 years, 31% below 40 years) compared to the White Europeans (mean age 58 years, 9% below 40 years) and African-Caribbeans (mean age 48 years, 23% below 40 years). Compared to male patients, females developed T2DM at a significantly higher BMI level across all ethnic groups. More South Asians developed T2DM at BMI below 30 kg/m^2 (38%) than White Europeans (26%) and African-Caribbeans (29%), (all $p<0.01$). Those not included in the study because of non-availability of ethnicity data were older (mean age 61 years compared to 56 years in the study cohort), but had similar distribution of sex, BMI, current smokers, ex-smokers, and never smokers.

The average SBP in South Asians at the time of diagnosis of T2DM (132 mmHg) was significantly lower with only 20% having SBP $\geq$ 140 mmHg, compared to the two other ethnic groups ($p<0.01$). Among those who developed T2DM, only 28% of South Asians were current or ex-smokers at index date, compared to 60% and 33% in the White and African-Caribbean ethnic groups respectively. African-Caribbean and South Asians had significantly higher LDL-C levels at diagnosis of T2DM (LDL-C $\geq$ 100 mg/dl: 34% and 28% respectively) compared to the White Europeans (LDL-C $\geq$ 100 mg/dl: 23%, Table 1). African-Caribbean patients had

significantly higher mean HbA_{1c} level at diagnosis [9.1% (76 mmol/mol), 30% with $HbA_{1c} \geq$ 7.5 %(58 mmol/mol)] compared to South Asians [8.5% (69mmol/mol), 28% with $HbA_{1c} \geq$ 7.5% (58 mmol/mol)] and White Europeans [8.2% (66 mmol/mol), 22% with $HbA_{1c} \geq 7.5\%$ (58 mmol/mol].

The distributions of BMI at diagnosis of T2DM in different ethnic groups, by age groups at diagnosis, are presented in Table 2. South Asians and African-Caribbeans aged 18-70 years at diagnosis developed T2DM at significantly lower BMI than White Europeans. Over the whole distribution of BMI level, the probability of developing T2DM in South Asians compared with other ethnic groups, separately for male and female and by different age groups, are presented in Figure 1 and Figure 2 respectively. When analysed with continuous measure of BMI, compared to both White Europeans and African-Caribbeans, South Asians had a significantly higher probability of developing T2DM within the range of BMI from 18 kg/m^2 to about 30 kg/m^2, for both male and female (Figure 1 A-D). The adjusted probability (95% CI) of developing T2DM at different BMI category levels compared between the three ethnic groups, separately for different age groups, are presented in Supplementary Figure 2. Across all age groups within the age of 70 years, South Asians and African-Caribbeans had significantly higher probability of developing T2DM in the normal weight and overweight categories, compared to White Europeans. Sensitivity analysis with extended list of covariates revealed similar results.

Given the different patterns of risk of developing T2DM among obese patients between ethnic groups across different age groups (Supplementary Figure 2), we evaluated the odds of developing T2DM in African-Caribbeans and South Asians, compared to White Europeans, separately for each age group. Within each age group, the probability of developing T2DM

was greater amongst South Asians at lower BMI. This relationship was reversed at higher BMI levels and compared to White Europeans, the South Asians had 22% (95% CI of OR: 0.65,0.95), 30% (95% CI of OR: 0.62,0.81), 24% (95% CI of OR: 0.65,0.88) and 39% (95% CI of OR: 0.48,0.77) lower odds (adjusted) of developing T2DM in the age groups ≤ 40, 41-50, 51-60 and 61-70 years respectively. African-Caribbean patients had 43% (95% CI of OR: 0.45, 0.73), 43% % (95% CI of OR: 0.50, 0.66), 33% (95% CI of OR: 0.57, 0.78) and 35% (95% CI of OR: 0.52, 0.82) lower adjusted odds of developing T2DM in the respective age groups. However, these odds were not statistically significantly different between African-Caribbeans and South Asians (Supplementary Table 2).

DISCUSSION

This case-control study with a large number of White European, South Asian and African-Caribbean individuals from a nationally representative primary care database reveals significantly different (1) distributions of body weight, BMI and other cardiovascular risk factors at the time of diagnosis of T2DM and (2) probability of developing T2DM over the whole spectrum of BMI, in interaction with age and sex. This study also reveals that the risk patterns of developing diabetes at different levels of obesity varies between ethnic groups across all age groups. To the best of our knowledge, this is the first study exploring the variations in T2DM risk over the whole distribution of BMI at the time of diagnosis across the South Asian, African-Caribbean and White European populations.

Our findings confirm the association of increased risk of T2DM with increasing BMI. More importantly, it adds to the evidence that for any given age, South Asians have a greater risk of T2DM at lower BMI. Typically, African-Caribbeans and South Asians in our study were significantly younger and had a distinct metabolic risk profile compared to White Europeans

characterised by lower body weight, systolic blood pressure, and lower rates of smoking but significantly higher HbA_{1c} levels. The observed higher HbA_{1c} level in South Asians and African-Caribbeans is in line with earlier findings [22]. While different possible reasons, including ethnic differences in pre- and post-prandial glycaemia and glycation rate of haemoglobin have been postulated, no confirmatory mechanistic study has yet been reported on this aspect. At population level, evaluation of longitudinal patterns of pre- and post-prandial glucose changes along with the measures of insulin deficiency from pre-diabetes state may reflect some light on this issue. Based on a US population with 12,179 T2DM patients and 25,177 controls, Ganz and colleagues (2014) reported a mean age and BMI of 55 years and 35 kg/m^2 respectively at diagnosis [5]. With similar age at diagnosis of T2DM, our UK study cohort had significantly lower BMI level (32 kg/m^2). However, this distribution of BMI at diagnosis is consistent with other studies reporting BMI at diagnosis of T2DM in UK population [23-25] . Similarly, our finding that women developed T2DM at significantly higher BMI level compared to men across ethnicity is consistent with earlier reports [26, 27]. Earlier studies have shown that the onset of T2DM occurs up to a decade early amongst South Asians. A Canadian cohort based diabetes incidence study reported that the median age at diagnosis was lowest among South Asians (49 years), followed by African-Caribbeans (57 years), and Whites (58 years) [9]. Our data is consistent with these observations and on average South Asians were 12 years younger at diagnosis compared to their White European counterparts. Factors that influence the predisposition of South Asians to develop T2DM at younger are largely unknown and may be related to a combination of genetic and environmental factors that have not yet been fully characterised [26, 27].

Ganz and colleagues (2014) reported significantly increasing odds of developing T2DM with increasing BMI [5]. While this study also evaluated the risk of developing T2DM in various age

groups using additive models, the interaction of age and body weight in the risk of developing T2DM was not explored. Given that age and obesity are two major risk factors for T2DM, we explored the interaction of age and BMI levels (separately for male and female) in evaluating the risk of developing T2DM across the three ethnic groups. One of the novelties of this study is comparative exploration of the probability of developing T2DM over the whole continuous distribution as well as categories of BMI by ethnic groups. When BMI was analysed as a continuous variable, we found that South Asians aged 40 years and above had a significantly greater probability of T2DM at lower BMI levels (18-30kg/m^2) compared to the other two ethnic groups. When analysed with BMI as a categorical variable, the higher probability of T2DM for South Asians with lower BMI extended from those younger than 40 years to those less than 70 years age. This difference in observed probability of developing T2DM using continuous BMI versus BMI categories is reflective of the fact that there is loss of statistical information when converting a continuous variable to a categorical variable. Interestingly, in both the analyses, we identified a distinct pattern of risk between South Asians and White Europeans with probability of T2DM being greater at lower BMI for South Asians and at higher BMI for White Europeans.

Earlier ethnicity specific studies evaluating association of prior BMI with the incident rates of T2DM reported higher risk in South Asians at a lower BMI level [7, 9, 28]. During a median follow-up of 6 years, Chui and colleagues reported higher T2DM incidence rates T2DM in South Asians at lower ages and BMI compared to the White Europeans [9]. Similar observation was made in another follow-up study reported by Tillin and colleagues [7]. Our study elaborates on the significantly higher likelihood of developing T2DM at lower BMI levels among South Asians, compared to White Europeans and African-Caribbeans. We have also identified a significant change in the risk pattern at higher BMI levels while compared between ethnic

groups at different age levels (Supplementary Figure 2 and Figure 2). Additionally, our study provides detailed information on the contrasting probability of developing T2DM at different age groups and ethnicity across the entire spectrum of BMI. There are several possible explanations for this contrasting effects of obesity on probability of diabetes between ethnic groups. While BMI is an accepted measure of obesity, it does not differentiate between patterns of obesity (visceral v sub cutaneous). It is well known that South Asians have excess visceral adiposity even at lower BMI and that may explain the higher propensity of South Asians to develop T2DM at lower BMI. However, that pattern would be expected to persist even at higher levels of BMI and therefore South Asians would be expected to have a greater risk in comparison with White Europeans for any level of BMI. Our observation that this difference is only evident at lower BMI but not at higher BMI range would suggest that the excess visceral adiposity alone does not explain this variance. An alternative explanation could be that relative contribution of obesity to the risk of T2DM may be greater amongst White Europeans compared to that in South Asians and that the risk of T2DM in South Asians may additionally be determined by underlying beta cell dysfunction. Clearly, this needs to be addressed in future studies.

The strength of this study is that it includes a large number of T2DM patients from a primary care system with a large number of South Asians; a representative age and sex matched non-diabetic control cohort; use of anthropometric, clinical risk factor measures at index date; and a robust analysis approach to explore the potential interactions between age, sex and BMI in different ethnic groups. Patient-level data from electronic health records present challenges in terms of accuracy and completeness of the study variables of interest. The limitations of this study include: (1) availability of ethnicity data on a limited number of patients, (2) missing risk factor data, (3) potential for residual confounding, and (4) inability to draw a causal link

between BMI and T2DM, as with all observational studies. However, ethnicity recording for South Asians and African-Caribbeans in the electronic database used for this study is comparable to the general population of UK [29]. We also attempted to minimize bias introduced by confounders through the use of multivariate models with a detailed list of possible confounders. However, unavailability of information on education, physical activity, diet, and other risk factors may have introduced bias into the risk estimates.

To conclude, the South Asian and African-Caribbean populations have an increased burden of T2DM with its complications. In this large case control data analysis we have demonstrated that T2DM occurs in South Asians at least 12 years earlier with mean age 46 years, when compared to White Europeans, with about a third developing under the age of 40 years. We believe the early presentation of diabetes in this ethnic population contributes to the glycaemic load and the burden of complications. Hence, the early diagnosis of diabetes, recognising the lower age of presentation may help to ameliorate the glycaemic burden and to do this a lower age cut-off for screening in national programmes for South Asians is required. This is the first large cohorts that we are aware of in which it has been demonstrated that South Asians develop diabetes at a mean lower BMI by 5 kg/m^2 when compared to White Europeans. This has implications both in terms of diagnosing obesity in South Asians along with appropriate management interventions.

ACKNOWLEDGEMENTS

S.K.P. and K.K. conceived the idea and was responsible for the primary design of the study. E.S.O.A contributed to the study design, and conducted data extraction and the statistical analyses. E.S.O.A. and S.K.P. developed the first draft of the manuscript. E.S.O.A., S.K.P., M.S., K.P., S.B., and W.H. contributed in the interpretation of the results and to the

finalization of the manuscript. S.K.P. had full access to all the data in the study and is the guarantor, taking responsibility for the integrity of the data and the accuracy of the data analysis. We also acknowledge Mr. Julius Agbeve and Dr. Margaret Haughton for their role in obtaining the original THIN data from IMS Health Inc. E.S.O.A. was supported by QIMR Berghofer International Ph.D. Scholarship and The University of Queensland International Scholarship. QIMR Berghofer gratefully acknowledges the support from the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS) initiative through Therapeutic Innovation Australia and the research project funding from the National Health and Medical Research Council of Australia (Project Number: GNT1063477). K.K. acknowledges support for this article from the UK National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care – East Midlands (NIHR CLAHRC – EM) and the NIHR Diet, Lifestyle & Physical Activity Biomedical Research Unit based at University Hospitals of Leicester and Loughborough University. South Asian Health Foundation (UK) supported the study.

Declaration of interests

S.K.P. has acted as a consultant and/or speaker for Novartis, GI Dynamics, Roche, AstraZeneca, Guangzhou Zhongyi Pharmaceutical and Amylin Pharmaceuticals LLC. He has received grants in support of investigator and investigator initiated clinical studies from Merck, Novo Nordisk, AstraZeneca, Hospira, Amylin Pharmaceuticals, Sanofi-Avensis and Pfizer. K.K. has acted as a consultant, speaker or received funds for research from Novartis, Novo Nordisk, Sanofi-Aventis, Lilly, Merck Sharp & Dohme, Janssen, Astra Zeneca and Boehringer Ingelheim, Roche and Servere. W.H. has received travel grants, research grant and consultancy fees from Novo Nordisk, Eli Lilly, Sanofi, MSD, Jansen, Astra Zeneca and Boehringer Ingelheim. S.B. has received research grants from The Binding Site Ltd and Novo

Nordisk (UK). S.B. has also received honoraria towards speaker fees and participation in advisory boards of Astra Zeneca, Boehringer Ingelheim, Janssen, Sanofi Aventis, Eli Lilly, MSD and Novo Nordisk. E.S.O.A. and M.S. have no conflict of interest to declare.

REFERENCES

[1] World Health Organization. Global status report on noncommunicable diseases 2014. WHO. Geneva: World Health Organization, 2015

[2] Ng M, Fleming T, Robinson M*, et al.* Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*. 2014; **384**: 766-781

[3] Kodama S, Horikawa C, Fujihara K*, et al.* Comparisons of the Strength of Associations With Future Type 2 Diabetes Risk Among Anthropometric Obesity Indicators, Including Waist-to-Height Ratio: A Meta-Analysis. *Am J Epidemiol*. 2012; **176**: 959-969

[4] Garber AJ. Obesity and type 2 diabetes: which patients are at risk? *Diabetes, obesity & metabolism*. 2012; **14**: 399-408

[5] Ganz ML, Wintfeld N, Li Q*, et al.* The association of body mass index with the risk of type 2 diabetes: a case-control study nested in an electronic health records system in the United States. *Diabetol Metab Syndr*. 2014; **6**: 50

[6] WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet (London, England)*. 2004; **363**: 157

[7] Tillin T, Sattar N, Godsland IF*, et al.* Ethnicity-specific obesity cut-points in the development of Type 2 diabetes - a prospective study including three ethnic groups in the United Kingdom. *Diabetic medicine : a journal of the British Diabetic Association*. 2015; **32**: 226-234

[8] Misra A. Ethnic-Specific Criteria for Classification of Body Mass Index: A Perspective for Asian Indians and American Diabetes Association Position Statement. *Diabetes Technol Ther*. 2015; **17**: 667-671

[9] Chiu M, Austin PC, Manuel DG, Shah BR, Tu JV. Deriving Ethnic-Specific BMI Cutoff Points for Assessing Diabetes Risk. *Diabetes care*. 2011; **34**: 1741-1748

[10] Misra A, Vikram NK, Gupta R*, et al.* Waist circumference cutoff points and action levels for Asian Indians for identification of abdominal obesity. *Int J Obes (Lond)*. 2006; **30**: 106-111

[11] Bodicoat DH, Gray LJ, Henson J*, et al.* Body Mass Index and Waist Circumference Cut-Points in Multi-Ethnic Populations from the UK and India: The ADDITION-Leicester, Jaipur Heart Watch and New Delhi Cross-Sectional Studies. *PLoS One*. 2014; **9**: e90813

[12] Shai I, Jiang R, Manson JE*, et al.* Ethnicity, Obesity, and Risk of Type 2 Diabetes in Women: A 20-year follow-up study. *Diabetes care*. 2006; **29**: 1585-1590

[13] Decode-Decoda Study Group. Age, body mass index and type 2 diabetes—associations modified by ethnicity. *Diabetologia*. 2003; **46**: 1063-1070

[14] IMS Health Incorporated. THIN Bibliography. Registered in the UK. 0532066, 2015

[15] Denburg MR, Haynes K, Shults J, Lewis JD, Leonard MB. Validation of The Health Improvement Network (THIN) database for epidemiologic studies of chronic kidney disease. *Pharmacoepidemiol Dug Saf*. 2011; **20**: 1138-1149

[16] Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Inform Primary Care*. 2011; **19**: 251-255

[17] Mulnier HE, Seaman HE, Raleigh VS*, et al.* Mortality in people with type 2 diabetes in the UK. *Diabetic medicine : a journal of the British Diabetic Association*. 2006; **23**: 516-521

[18] Read J. The Read clinical classification (Read codes). *Br Homoeopath J*. 1991; **80**: 14-20

[19] World Health Organization. Obesity: preventing and managing the global epidemic: report of a WHO Consultation on Obesity. Geneva: WHO, 1998:3-5 June 1997

[20] King G, Tomz M, Wittenberg J. Making the most of statistical analyses: Improving interpretation and presentation. *Am J Pol Sci*. 2000: 347-361

[21] Zelner BA. Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management Journal*. 2009; **30**: 1335-1348

[22] James GD, Baker P, Badrick E*, et al.* Ethnic and social disparity in glycaemic control in type 2 diabetes; cohort study in general practice 2004–9. *J R Soc Med*. 2012; **105**: 300-308

[23] Thomas G, Khunti K, Curcin V*, et al.* Obesity paradox in people newly diagnosed with type 2 diabetes with and without prior cardiovascular disease. *Diabetes, obesity & metabolism*. 2014; **16**: 317-325
[24] Paul SK, Klein K, Thorsted BL, Wolden ML, Khunti K. Delay in treatment intensification increases the risks of cardiovascular events in patients with type 2 diabetes. *Cardiovascular diabetology*. 2015; **14**: 1
[25] Paul S, Klein K, Majeed A, Khunti K. Association of smoking and concomitant use of metformin with cardiovascular events and mortality in people newly diagnosed with type 2 diabetes. *J Diabetes*. 2015; **201**:
[26] Paul S, Thomas G, Majeed A, Khunti K, Klein K. Women develop type 2 diabetes at a higher body mass index than men. *Diabetologia*. 2012; **55**: 1556-1557
[27] Peters SA, Huxley RR, Woodward M. Sex differences in body anthropometry and composition in individuals with and without diabetes in the UK Biobank. *BMJ open*. 2016; **6**: e010007
[28] Ntuk UE, Gill JMR, Mackay DF, Sattar N, Pell JP. Ethnic-Specific Obesity Cutoffs for Diabetes Risk: Cross-sectional Study of 490,288 UK Biobank Participants. *Diabetes care*. 2014; **37**: 2500-2507
[29] Mathur R, Bhaskaran K, Chaturvedi N*, et al.* Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health (Oxf)*. 2014; **36**: 684-692

FIGURE LEGENDS

Figure 1: The association between BMI at diagnosis and risk of T2DM (95% CI) compared between three groups of ethnicity. [***A****: Adjusted probability of developing T2DM (95% CI) at different levels of BMI for male South Asians, compared to male White Europeans;* ***B****: Adjusted probability of developing T2DM (95% CI) at different levels of BMI for female South Asians compared to female White Europeans ;* ***C****: Adjusted probability of developing T2DM at different BMI levels for male South Asians compared to male African-Caribbean;* ***D*** *: Adjusted probability of developing T2DM at different BMI levels for female South Asians compared to female White Europeans ; Probability estimated in Figures 1A, 1B, 1C, & 1D are adjusted for age, smoking status, deprivation score, and history of CVD, cancer and CKD on or prior to the index date.*].

Figure 2: Adjusted probability of developing T2DM (95% CI) across levels of BMI for different age groups. [***A****: South Asians compared to White Europeans;* ***B****:South Asians compared to African-Caribbeans ;Estimated are adjusted for sex, smoking status, deprivation score, and history of CVD, cancer and CKD on or prior to the index date; SA: South Asian; WE: White European; AC: African-Caribbean*].

Supplementary Figure 1: The identification of T2DM study cohort and their matched controls from THIN database. [[1]: T2DM Read code + (antidiabetic medication or lifestyle modification intervention, excludes patients with any other type of diabetes (eg. Type 1 diabetes, Gestational diabetes); [2]: Age at index date greater or equal to 18, complete information on gender and ethnicity (White European, African-Caribbean and South Asian only]; [3] Excludes patients who have ever received anti-hyperglycaemic drugs.

Supplementary Figure 2: Adjusted probability of developing T2DM (95% CI) across levels of BMI categories for different age groups. [*Adjusted for sex, smoking status, deprivation score, and history of CVD, cancer, and CKD on or prior to the index date; NW: Normal weight; OW: Overweight; G1O: Grade 1 Obese; G2O: Grade 2 Obese; G3O: Grade 3 Obese]*.

TABLES

Table 1: Distribution of basic characteristics of T2DM patients and their matched controls, stratified by ethnicity

	White European (N=396,350)		**African-Caribbean (N=20,575)**		**South Asian (N=36,260)**		**ALL (N=453,770)**	
	T2DM	Control	T2DM	Control	T2DM	Control	T2DM	Control
Patients*	79,270	317,080	4,115	16,460	7,252	29,008	90,637	362,548
Age at diagnosis (years) †	58 ± 12	58 ± 12	48 ± 12	48 ±12	46 ± 12	46 ± 12	56 ± 13	56 ± 13
Age at diagnosis (years) ‡	58 (49, 67)	58 (49, 67)	48 (40, 56)	48 (40, 56)	45 (38, 54)	45 (38, 54)	57 (47, 66)	57 (47, 66)
Age group *								
≤40	6,724 (9)	26,896 (9)	949 (23)	3,796 (23)	2,204 (30)	8,816 (30)	9,877 (11)	39,508 (11)
41-50	15,614 (20)	62,456 (20)	1,511 (37)	6,044 (37)	2,598 (36)	10,392 (36)	19,723 (22)	78,892 (22)
51-60	22,425 (28)	89,700 (28)	1,007 (25)	4,028 (25)	1,529 (21)	6,116 (21)	24,961 (28)	99,844 (28)
61-70	21,751 (27)	87,004 (27)	491 (12)	1,964 (12)	688 (10)	2,752 (10)	22,930 (25)	91,720 (25)
71+	12,756 (16)	51,024 (16)	157 (4)	628 (4)	233 (3)	932 (3)	13,146 (15)	52,584 (15)
Male *	44,651 (56)	178,604 (56)	2,102 (51)	8,408 (51)	4,005 (55)	16,020 (55)	50,758 (56)	203,032 (56)
Current smokers *	15,581 (20)	59,060 (19)	527 (13)	2,307 (14)	967 (13)	3,460 (12)	17,075 (19)	64,827 (18)
Ex-smokers *	31,966 (40)	114,099 (36)	808 (20)	2,612 (16)	1,058 (15)	3,497 (12)	33,832 (37)	120,208 (33)
Never smokers *	31,427 (40)	138,903 (44)	2,769 (67)	11,130 (68)	5,195 (72)	21,162 (73)	39,391 (44)	171,195 (47)
Highest affluence *	3,391 (4)	15,054 (5)	564 (14)	1,954 (12)	665 (9)	2,525 (9)	4,620 (5)	19,533 (5)
Lowest affluence *	16,923 (21)	56,124 (18)	994 (24)	4,290 (26)	1,852 (26)	7,073 (24)	19,769 (22)	67,487 (19)
HbA_{1c} (%) ,[mmol/mol] §	8.2 ± 2.1 [66 ± 23.0]		9.1 ± 2.7 [76 ± 29.5]		8.5 ± 2.1 [69 ± 23.0]		8.3 ± 2.1 [67 ± 23.0]	
$HbA_{1c} \geq 7.5\%$ *§	17,730 (22)		1,234 (30)		2,021 (28)		20,985 (23)	
Weight (kg) †	92.2 ± 21.0	77.8 ± 16.8	88.3 ± 18.7	81.0 ± 16.2	78.8 ± 17.1	71.3 ± 14.6	91.0 ± 21.0)	77.5 ± 16.8
Weight (kg) ‡	90.0 (78, 104)	76.2 (66, 88)	86.0 (75, 99)	79.7 (70, 90)	76.0 (67, 88)	70.0 (61, 80)	88.9 (76, 103)	76.0 (66, 87)
BMI (kg/m^2) † -- All	32.5 ± 6.8	27.4 ± 5.2	31.1 ± 6.2	28.7 ± 5.6	29.2 ± 5.7	26.5 ± 4.8	32.2 ± 6.8	27.4 ± 5.2
BMI (kg/m^2) † -- Male	31.7 ± 6.0	27.5 ± 4.7	29.3 ± 5.3	27.3 ± 4.6	28.4 ± 5.4	26.2 ± 4.6	31.4 ± 6.0	27.4 ± 4.7
BMI (kg/m^2) † -- Female	33.4 ± 8.0	27.4 ± 5.8	33.0 ± 6.6	30.0 ± 6.2	30.2 ± 5.8	26.6 ± 5.1	31.5 ± 6.8	27.0 ± 5.3
BMI (kg/m^2) ‡	31.4 (28, 36)	26.8 (24, 30)	30.2 (27, 35)	28.0 (25, 32)	28.3 (25, 32)	26.0 (23, 29)	31.1 (28, 36)	26.8 (24, 30)
Normal weight*	4,958 (6)	24,052 (8)	365 (9)	834 (5)	921 (13)	2,374 (8)	6,244 (7)	27,260 (8)
Overweight*	15,439 (20)	30,135 (10)	820 (20)	1,283 (8)	1,830 (25)	2,376 (8)	18,089 (20)	33,794 (9)
Grade 1 obese*	15,592 (20)	13,654 (4)	719 (18)	764 (5)	1,050 (15)	896 (3)	17,361 (19)	15,314 (4)

Grade 2 obese*	8,973 (11)	4,073 (1)	383 (9)	288 (2)	386 (5)	251 (1)	9,742 (11)	4,612 (1)
SBP (mmHg) †	141 ± 19	136 ± 19	137 ± 19	133 ± 19	132 ± 18	128 ± 17	140 ± 19	135 ± 18
SBP ≥ 140 mmHg *	28,971 (37)	52,256 (17)	1,104 (27)	1,742 (11)	1,461 (20)	2,092 (7)	31,536 (35)	56,090 (16)
DBP (mmHg) †	82 ± 11	80 ± 10	83 ± 11	81 ± 11	82 ± 11	79 ± 10	82 ± 11	80 ± 10
LDL-C(mg/dl) †	117 ± 42	120 ± 40	126 ± 39	124 ± 35	119 ± 39	121 ± 35	117 ± 42	120 ± 39
LDL-C ≥ 100 mg/dl*	17,944 (23)	25,328 (8)	1,388 (34)	1,531 (9)	2,003 (28)	2,684 (9)	21,335 (24)	29,543 (8)
HDL-C (mg/dl)†	46 ± 14	56 ±17	48 ± 13	58 ± 17	44 ± 11	51 ± 14	46 ± 13	56 ± 17
HDL-C ≤ 45 mg/dl*	19,024 (24)	11,941 (4)	924 (23)	501 (3)	2,041 (28)	1,452 (5)	21,989 (24)	13,894 (4)
Triglycerides (mg/dl) ‡	159 (122, 213)	115 (87, 159)	115 (81, 159)	82 (62, 115)	151 (115, 204)	115 (89, 168)	159 (115, 213)	115 (84, 159)
Triglyceride ≥ 150 mg/dl*	18,654 (24)	13,323 (4)	601 (15)	260 (2)	1,681 (23)	1,316 (5)	20,936 (23)	14,899 (4)
Complications*****								
CKD (≥ stage 3)	1,752 (2)	4,493 (1)	52 (1)	146 (1)	34 (1)	143 (1)	1,838 (2)	4,782 (1)
Cancer	4,746 (5)	18,793 (6)	100 (2)	278 (2)	68 (1)	339 (1)	4,914 (5)	19,410 (5)
Myocardial Infarction	4,802 (6)	9,120 (3)	34 (1)	85 (1)	181 (3)	355 (1)	5,017 (6)	9,560 (3)
Heart Failure	1,743 (2)	2,635 (1)	29 (1)	38 (0)	41 (1)	76 (<0.1)	1,813 (2)	2,749 (1)
Angina	6,529 (8)	12,975 (4)	48 (1)	93 (1)	196 (3)	440 (2)	6,773 (8)	13,508 (4)
Stroke	4,014 (5)	9,521 (3)	105 (3)	198 (1)	111 (2)	275 (1)	4,230 (5)	9,994 (3)
Any CVD	15,769 (20)	33,155 (11)	225 (6)	443 (3)	519 (7)	1,101 (4)	16,513 (18)	34,699 (10)
Hypertension	33,234 (42)	62,749 (20)	1,419 (35)	2,852 (17)	1,803 (25)	2,941 (10)	36,456 (40)	68,542 (19)
Anti-hyperglycaemic drugs (Ever prescribed)*****								
None	10,767 (14)	317,080 (100)	294 (7)	16,460 (100)	536 (7)	29,008 (100)	11,597 (13)	362,548(100)
Insulin	17,693 (22)		917 (22)		1,381 (19)		19,991 (22)	
Biguanides	62,598 (79)		3,530 (86)		6,324 (87)		72,452 (80)	
Sulphonylureas	38,281 (48)		2,090 (51)		3,739 (52)		44,110 (49)	
Thiazolidinedione	14,481 (18)		622 (15)		1,411 (20)		16,514 (18)	
GLP1-RA	3,769 (5)		117 (3)		225 (3)		4,111 (5)	
DPP-4	11,078 (14)		633 (15)		1,231 (17)		12,942 (14)	
Alpha glucosidase	1,545 (2)		58 (1)		117 (2)		1,720 (2)	
SGLT2	615 (1)		21 (1)		59 (1)		695 (1)	
Metglinides	902 (1)		58 (1)		100 (1)		1,060 (1)	
Other medications (Ever prescribed) *****								

Antihypertensive	3,577 (5)	7,496 (2)	173 (4)	390 (2)	145 (2)	320 (1)	3,895 (4)	8,206 (2)
Diuretics	20,688 (26)	41,226 (13)	585 (14)	1,421 (9)	684 (9)	1,508 (5)	21,957 (24)	44,155 (12)
Beta blockers	18,938 (24)	42,580 (13)	427 (10)	1,103 (7)	758 (11)	1,913 (7)	20,123 (22)	45,596 (13)
Calcium blockers	15,070 (19)	30,059 (10)	724 (18)	1,588 (10)	703 (10)	1,386 (5)	16,497 (18)	33,033 (9)
Statins	19,376 (24)	35,419 (11)	627 (15)	811 (5)	1,219 (17)	1,679 (6)	21,222 (23)	37,909 (11)
Ace inhibitors	17,145 (22)	30,152 (10)	558 (14)	889 (5)	823 (11)	1,370 (5)	18,526 (20)	32,411 (9)
CPMs	34,595 (44)	73,919 (23)	1,272 (31)	2,473 (15)	1,997 (28)	3,671 (13)	37,864 (42)	80,063 (22)
Anti-depressants	17,419 (22)	54,807 (17)	393 (10)	1,413 (9)	926 (13)	3,025 (10)	18,738 (21)	59,245 (16)
Anti-obesity	3,213 (4)	2,940 (1)	121 (3)	171 (1)	200 (3)	229 (1)	3,534 (4)	3,340 (1)

***: n (%);*
†: mean ± SD;
‡: median (Q1, Q3)
***§**: Not presented for control subjects*
BMI: Body mass index;
SPB: Systolic blood pressure;
LDL-C: Low density lipoprotein cholesterol;
HDL-C: High density lipoprotein cholesterol;
CKD: Chronic kidney disease;
Any CVD: Defined as the occurrence of angina or myocardial infarction or coronary artery disease (including bypass surgery and angioplasty) or heart failure or stroke before diagnosis

Table 2: Mean ± SD of BMI in three ethnic groups at the time of diagnosis of T2DM, by different age groups at diagnosis among patients diagnosed from 01 Jan 2006. The p values are estimated to present the significance of differences in the distribution of BMI between each combination of two ethnic groups. The proportions of patients under different obesity grades are presented by number (%).

			Absolute difference in mean (p-value)			BMI categories*		
	n	**BMI†**	**WE Vs SA**	**WE Vs AC**	**AC Vs SA**	**Grade 1 Obese**	**Grade 2 Obese**	**Grade 3 Obese**
Age group: ≤ 40 years								
White European	2051	36.3 ± 8.6	6.6 (<0.001)	5.2(<0.001)	1.4 (<0.022)	509 (24)	468 (23)	615 (30)
African-Caribbean	311	31.1 ± 6.5				89 (29)	43 (14)	28 (9)
South Asian	892	29.8 ± 6.2				212 (24)	90 (10)	64 (7)
Age group: 41-50 years								
White European	5717	35.1 ± 7.4	5.3 (<0.001)	3.6 (<0.001)	2.1 (<0.001)	1635 (29)	1360 (24)	1263 (22)
African-Caribbean	716	31.5 ± 6.3				215 (30)	128 (18)	57 (8)
South Asian	1092	29.4 ± 5.8				278 (26)	109 (10)	44 (4)
Age group: 51-60 years								
White European	7907	33.4 ± 6.7	3.8 (<0.001)	2.2 (<0.001)	1.6 (<0.001)	2537 (32)	1644 (21)	1139 (14)
African-Caribbean	497	31.2 ± 6.2				132 (27)	82 (17)	43 (9)
South Asian	679	29.6 ± 5.7				167 (25)	73 (11)	27 (4)
Age group: 61-70 years								
White European	8487	32.1 ± 6.2	3.5 (<0.001	1.1 (<0.038)	2.3 (<0.001)	2849 (34)	1446 (17)	824 (10)
African-Caribbean	188	30.9 ± 5.4				65 (35)	27 (14)	12 (6)
South Asian	264	28.6 ± 5.4				56 (21)	21 (8)	9 (3)
Age group: 70+ years								
White European	5198	30.2 ± 5.4	2.7 (<0.001)	0.44 (1.000)	2.3 (0.015)	1554 (30)	637 (12)	255 (5)
African-Caribbean	80	29.7 ± 6.6				19 (24)	9 (11)	8 (10)
South Asian	100	27.4 ± 5.5				18 (18)	8 (8)	3 (3)
Female								
White European	12,292	33.9 ± 7.7	3.4 (<0.001)	0.6 (0.051)	2.8 (<0.001)	3414 (28)	2534 (21)	2328 (19)
African-Caribbean	823	33.3 ± 6.4				265 (32)	194 (24)	107 (13)

South Asian	1303	30.5 ± 6.0				344 (26)	193 (15)	86 (7)
Male								
White European	17068	32.3 ± 6.2	3.7 (<0.001)	2.8 (<0.001)	0.9 (0.002)	5,670 (33)	3,012 (16)	1768 (10)
African-Caribbean	969	29.5 ± 5.4				255 (26)	95 (10)	41 (4)
South Asian	1724	28.6 ± 5.6				387 (22)	108 (6)	61 (4)

: n (%);
†: mean ± SD;
WE: White European;
SA: South Asian;
AC: African-Caribbean

FIGURES

Figure 1:

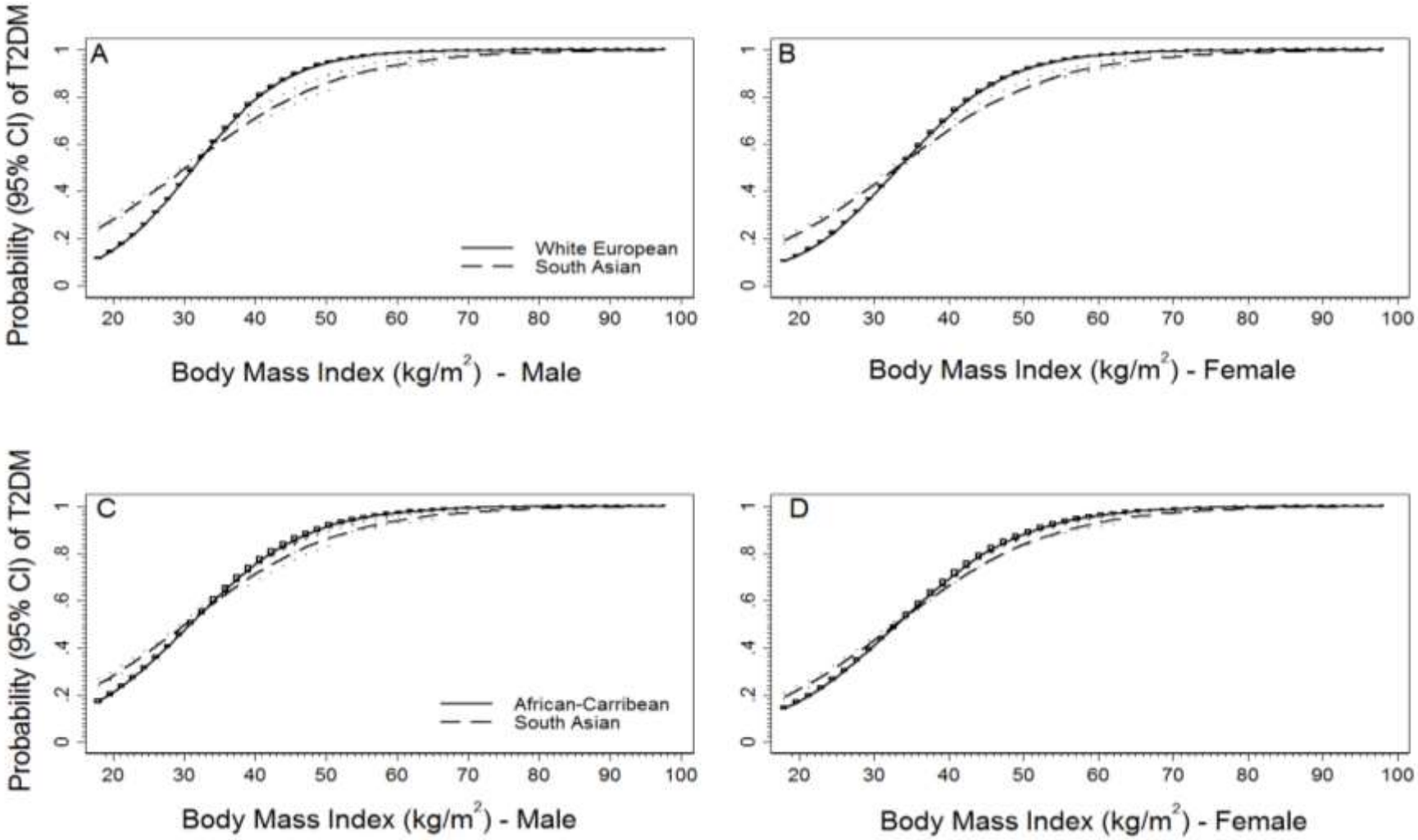
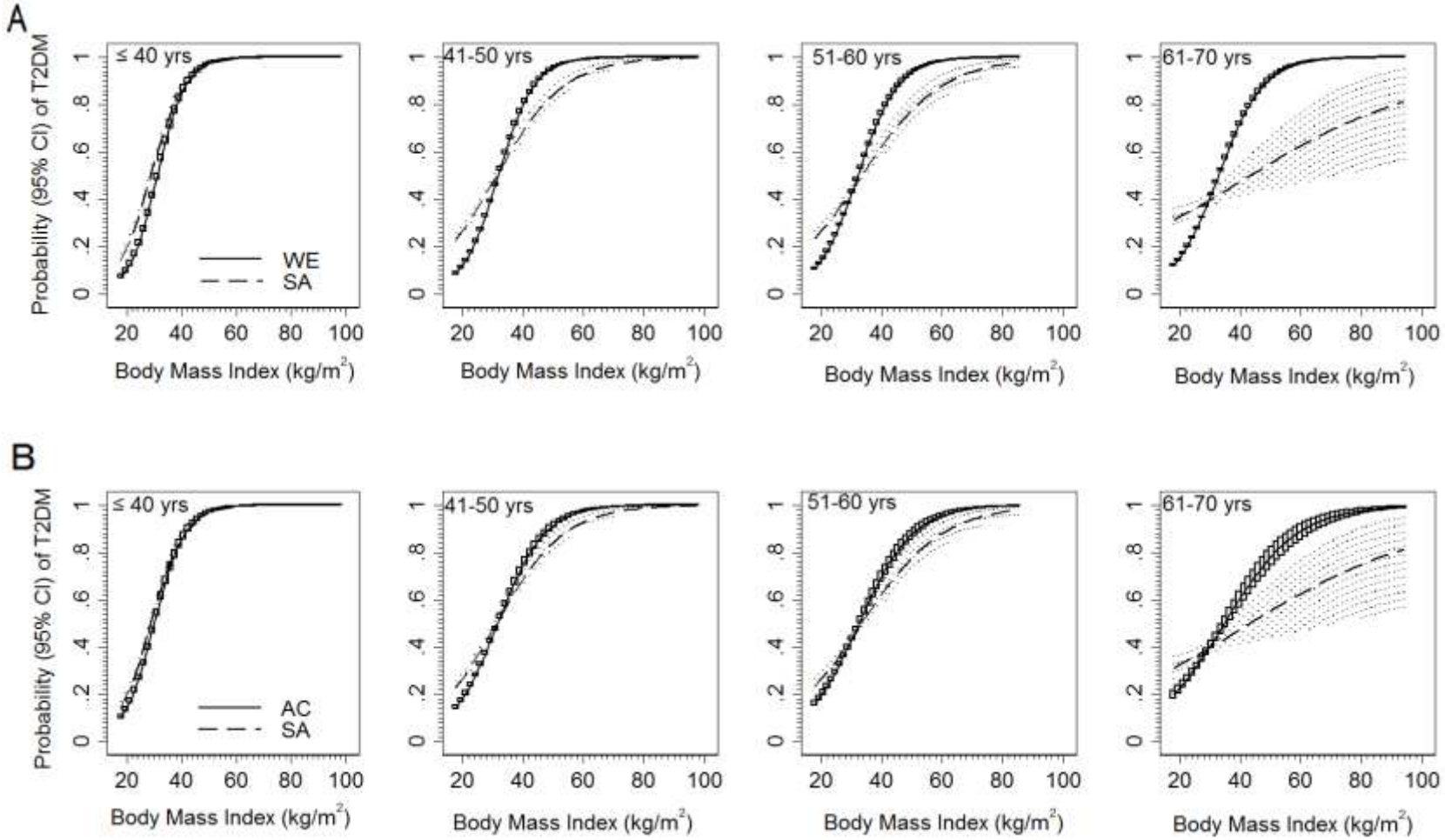


Figure 2:


A
≤ 40 yrs
41-50 yrs
51-60 yrs
61-70 yrs
Probability (95% CI) of T2DM
WE
SA
Body Mass Index (kg/m²)
B
≤ 40 yrs
41-50 yrs
51-60 yrs
61-70 yrs
Probability (95% CI) of T2DM
AC
SA
Body Mass Index (kg/m²)