Type 2 diabetes risks and determinants in 2nd generation migrants and mixed ethnicity people of South Asian and African Caribbean descent in the UK

Aliki-Eleni Farmaki, PhD, MRC Unit for Lifelong Health and Ageing, Institute of Cardiovascular Science, University College London, 1-19 Torrington Place, London WC1E 7HB, UK

Victoria Garfield, PhD, MRC Unit for Lifelong Health and Ageing, Institute of Cardiovascular Science, University College London, 1-19 Torrington Place, London WC1E 7HB, UK

Sophie V. Eastwood, MRCGP, MRC Unit for Lifelong Health and Ageing, Institute of Cardiovascular Science, University College London, 1-19 Torrington Place, London WC1E 7HB, UK

Ruth E. Farmer, PhD, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Rohini Mathur, PhD, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Praveetha Patalay, PhD, MRC Unit for Lifelong Health and Ageing; Centre for Longitudinal Studies, University College London, London, UK

Naveed Sattar, FMedSci, Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK

Alun Hughes, PhD, MRC Unit for Lifelong Health and Ageing, Institute of Cardiovascular Science, University College London, London, UK

Krishnan Bhaskaran, PhD, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Liam Smeeth, PhD, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Nish Chaturvedi, MD, MRC Unit for Lifelong Health and Ageing, Institute of Cardiovascular Science, University College London, 1-19 Torrington Place, London WC1E 7HB, UK

Abstract

Background: Type 2 diabetes mellitus (T2DM) risk is markedly higher in UK South Asians (SA) and African Caribbeans (AC) compared to Europeans. Explanations for this excess are unclear. We therefore compared risks and determinants of T2DM in first- and second-generation (born in the UK) migrants, and in those of mixed ethnicity populations.

Methods: Data from the UK Biobank, a large population-based cohort of volunteers aged 40-69, were used. T2DM was assigned using self-report and glycated haemoglobin. Ethnicity was self-assigned. Using logistic regression and mediation analysis, we compared T2DM between first- and second-generation migrants, and between mixed European/South Asians (MixESA), or mixed European/African Caribbeans (MixEAC) with both Europeans and SA or AC respectively.

Results: T2DM prevalence was three to five times higher in SA and AC compared with Europeans [OR (95%CI): 4·80(3·60,6·40) and 3·30(2·70,4·10) respectively]. T2DM was 20-30% lower in second versus first generation SA and AC migrants [0·78(0·60,1·01) and 0·71(0·57,0·87) respectively]. T2DM in mixed populations was lower than comparator ethnic minority groups [MixESA versus SA 0·29(0·21,0·39), MixEAC versus AC 0·48(0·37,0·62)] and higher than Europeans, in MixESA 1·55(1·11, 2·17), and in MixEAC 2·06 (1·53, 2·78). Improved adiposity patterns in second generation migrants made an important contribution to risk reduction. Greater socioeconomic deprivation accounted for 17% and 42% of the excess risk of T2DM in MixESA and MixEAC compared to Europeans, respectively.

Conclusion: Excess T2DM risks in South Asians and African Caribbeans compared with Europeans in the UK are attenuated by ~20% in second-generation migrants, demonstrating the marked benefits of favourable changes in environmental risk factors. T2DM prevalence in people of mixed ethnicity was also raised compared with Europeans, but considerably less than in the ethnic minority group; persistent socioeconomic disadvantage accounted for some of the residual excess.

Introduction

Type 2 diabetes mellitus (T2DM) is estimated to affect 693 million people worldwide by 2045 (1). People of African Caribbean (AC) and South Asian (SA) descent have some of the highest rates of T2DM in the world, often three to four times greater, respectively, than those of European ancestry when compared in the same setting (2). Explanations for this excess risk remain unclear.

Studies of migrant offspring, and people of mixed ethnicity, where distribution and inter-relations of genetic and environmental explanatory factors differ, may offer fresh insights. Previous studies suggest second and subsequent generations of migrants are at persistently higher risk(3). In contrast, partial European ancestry has been associated with decreased, and non-European ancestry with increased T2DM risk in admixed populations of Hispanic, African American, Asian, and European descent (4,5).

To date, no study has combined mixed ethnicity comparisons and inter-generational analysis in the same setting to triangulate evidence and understand the impact of mutable environmental risk factors in determining risk. We hypothesised that while offspring of migrants of either SA or AC ethnicity would retain the excess risks of T2DM of their parents, people of mixed European/South Asian (MixESA), or mixed European/African Caribbean (MixEAC) ethnicity would have risks of T2DM intermediate between each of the parental ethnic groups. For the latter, we further hypothesised that T2DM differences across self-reported ethnicities would largely be explained by differences in adiposity and other socioeconomic factors.

Methods

Study design

We used data from UK Biobank, a large population-based cohort of over 500,000 men and women aged 40-69 years recruited from primary care lists in the UK between 2006-2010 (6).

The following data were collected by self-completion or nurse administered questionnaires: self-defined ethnicity using the UK census classification (7), year of migration to the UK to assign generational status, health behaviours including smoking (ever smoked), physical activity (number of days/week of moderate physical activity more than ten minutes) and diet (data from the touchscreen questionnaire on the reported frequency of intake of a range of common food and drink items), and sociodemographic variables such as education and Townsend deprivation score assigned by residential postcode (8). Height, weight and body circumferences were measured directly, and bio-impedance was used to assess fat mass and fat percentage (%). Participants were asked to recall birth weight.

A blood sample was taken for the measurement of biochemical markers in serum. HbA1c (mmol/mol) was measured from blood samples taken at baseline, as outlined in the UK Biobank protocol. Values above 195 mmol/mol (n=5) were considered outliers and excluded from the analysis.

"Known T2DM" at recruitment was defined according to an algorithm based on self-report data and medication; this algorithm has been validated against primary care records (9). "All T2DM" included those with "Known T2DM" plus all those with an HbA1c > 47 mmol/mol.

Migration status (first or second generation) was defined based on the reported year of migration. Self-assigned MixESA and MixEAC were the ethnic groups of interest, with European, SA, and AC ethnicities for comparison.

Matching procedure

Sex and age matching was needed as ethnic minority populations in the UK are younger than the general European origin population; in addition we wished to compare first (born abroad – migrated to the UK) and second (born to two ethnic minority parents resident in the UK) generation migrants. The reference group was the mixed or the second-generation group, depending on the comparison. As the reference groups were the smallest, we matched to optimise power employing 1:4 matching where possible, and 1:2 where the sample was insufficient. Matching was performed at random within sex and five year age bands. Each

matching procedure was performed independently to create unique datasets for each analysis:

- MixESA (n=831) – SA – Europeans (1:4:4)
- MixEAC (n=1,045) – AC – Europeans (1:4:4)
- Second generation SA (n=1,115) – First generation SA – Europeans (1:1:2)
- Second generation AC (n=2,200) – First generation AC – Europeans (1:1:2)

Details of matching procedure and frequency distributions for each of the derived datasets are shown in Supplement S1.

Dietary patterns

Crude food frequencies were captured by touchscreen questionnaire at baseline. Using a principal component approach, we derived a 'healthy' diet variable characterised by fruit, vegetable, fish and water consumption, and an 'unhealthy' diet variable, which included red meat, processed meat and coffee drinking (further details in Supplement S2).

Statistical analyses

The mediators considered as possibly contributing to the association between ethnicity and T2DM included: smoking; Townsend deprivation score as a proxy for socioeconomic status; height, birth weight; years of education derived from qualifications based on the International Standard Classification of Education (ISCED) coding (10); and adiposity measures. We selected waist to hip ratio (WHR) as our key measure for adiposity in the SA analyses, and body mass index (BMI) for the AC analysis, as these measures best accounted for ethnic differences in T2DM in a previous population cohort analysis (2). Sensitivity analyses using BMI in the SA and WHR for the AC were also performed (Supplement S3).

The extent to which adiposity pattern, deprivation, smoking, height and education mediated the relationship between ethnicity and T2DM was explored in path models (the sample sizes of these analyses are shown in Supplement S3). The total effect was the observed effect of ethnicity on T2DM without adjustment; the direct effect was the remaining (independent) effect of ethnicity on T2DM after adjustment for all variables depicted in the Directed Acyclic

Graph (DAG) (Figure 1). The difference between the two, the indirect effect, is attributed to each of the mediators singly or jointly following the DAG defined pathways. The indirect effect can therefore be interpreted as the percentage of the total effect mediated by these explanatory variables. All models were adjusted for age and sex (11).

Our initial DAG included physical activity and diet (Suppl. Figure s3.1). However these behaviours were crudely assessed (physical activity was captured as number of days per week doing more than 10 minutes moderate physical activity, dietary data relied on a 20 item food frequency questionnaire). For this reason, we did not include these measures in subsequent mediation analysis. Similarly, although birth weight was considered an important mediator, only half of the sample had these data, severely diminishing analytical precision. Birth weight was also therefore excluded from the final DAG (Figure 1), we did however perform sensitivity analysis on complete case data (Supplement S4). Sensitivity analyses were also performed replacing T2DM with HbA1c as the outcome (Supplement S3).

Statistical analyses comparing recruitment characteristics were performed in Stata 15. Mediation analysis testing path models was performed with Mplus with MLR estimator and Monte Carlo integration at 10,000.

Results

1. South Asians

1.1. South Asians vs Europeans

T2DM prevalence was almost five times higher in SA (odds ratio (OR): 4·8 (3·6,6·4), compared to Europeans (Figure 2). WHR in SA (0·89) was higher than in Europeans (0·86). Proportion resident in the lowest quintile of deprivation was near double in SA (41%), compared to Europeans (21%) (Table 1).

1.2. Second vs first generation South Asians

T2DM prevalence was a fifth (22%) lower in second (OR: 0·78 (0·60,1·01)) versus first generation SA (Figure 2). WHR and proportion resident in the most deprived quintile were marginally lower in second-generation migrants, and years of education as well (Table 1).

1.2.1. Mediation

The mediating variables we tested accounted for more than a third of the 22% lower risk of T2DM in second versus first generation SA, with WHR making the strongest contribution (30%) (Table 5, Suppl. figure s3.2C).

1.3. MixESA

T2DM prevalence was 55% higher in those of MixESA (OR: 1·55 (1·11,2·17)) compared to Europeans, (Figure 2). WHR was similar in MixESA and Europeans, and markedly lower than in SA. Residence in the most deprived quintile of deprivation in MixESA was intermediate (28%) between that of SA (37%) and Europeans (19%), though MixESA had the most years in education (Table 2). Heights were also intermediate.

1.3.1. Mediation

The analysed data could account for just over a quarter (28%) of the 55% excess risk of T2DM in MixESA versus Europeans (Table 5, Suppl. Figures s3.2A and s3.3). Deprivation mediated most of this association (17%), followed by WHR (9%). In contrast the higher educational status in MixESA compared to Europeans partially mitigated their predisposition to T2DM (-7%). A similar proportion of the markedly (71%) lower risk of T2DM in MixESA versus SA (Figure 2) was accounted for by the above mediators, dominated by WHR (15%) (Table 5, Suppl. figures s3.2B and s3.4).

2. African Caribbeans

2.1. African Caribbeans vs Europeans

Similar inter-ethnic differences were observed when comparing people of AC and European descent (Figure 2, Table 3). Specifically, T2DM risk in AC was three times that of Europeans

(OR: 3·3 (2·7,4·1)). BMI in AC (28·2kg/m^2 and 30·8kg/m^2 in men and women, respectively) was higher than in Europeans (27·7kg/m^2 in men, 26·7kg/m^2 in women). Ever smoking was less prevalent in AC (16%) than in Europeans (30%). Residence in the bottom Townsend quintile was over three times greater in AC (70%) than Europeans (21%).

2.2. Second vs first generation African Caribbeans

Risk of T2DM was 29% lower in second (OR: 0·71 (0·57,0·87) compared to first generation AC (Figure 2), also with lower BMI (29·1 versus 29·7kg/m^2 respectively) (Table 3). Residence in the lowest quintile of deprivation was marginally lower in second, versus first generation migrants, (60 vs. 70%). The proportion of ever smokers in second generation AC was double that of first generation migrants (30% vs. 16%).

2.2.1. Mediation

About a fifth of the 29% lower risk of T2DM in second versus first generation AC was accounted for, with equal contributions from deprivation and BMI (Table 5, Suppl. figure s3.6C).

2.3. MixEAC

T2DM risk in MixEAC was double that of Europeans (OR: 2·06 (1·53,2·78)), but half that of AC (OR: 0·48 (0·37,0·62)) (Figure 2). BMI in MixEAC (28kg/m^2 in both sexes) was close to that of Europeans (27·8kg/m^2 in men, 26·8kg/m^2 in women) and lower than in AC (28·3kg/m^2 and 30·3kg/m^2 in men and women, respectively) (Table 4). Ever smoking was markedly more prevalent in MixEAC (42%) than in AC (20%) and Europeans (30%). Residence in the bottom Townsend quintile was highest in AC (63%), intermediate in MixEAC (48%) and lowest in Europeans (19%). Educational attainment was lowest in MixEAC.

2.3.1. Mediation

About two thirds of the doubling in T2DM prevalence in MixEAC compared to Europeans could be accounted for, largely by deprivation (42%), and BMI (11%) (Table 5, Suppl. figure

s3.6A). In contrast, about a third of the halving in risk of T2DM in MixEAC versus AC could be accounted for, mostly by BMI (16%) (Table 5, Suppl. Figure s3.6B).

3. *Sensitivity analyses*

Using BMI in the SA (instead of WHR) and WHR for the AC (instead of BMI), each accounted for a lower proportion of the observed difference in T2DM risk than the originally selected adiposity measure (Suppl. figures s3.5 and s3.7).

Associations and mediation patterns were similar when HbA1c replaced T2DM as the outcome (Suppl. figures s3.8 and s3.9).

Sensitivity analysis on complete case data (Suppl. figures s4.2 and s4.4) for birthweight showed that although it made some contribution to T2DM risk, the effect was not as strong as other mediators included in our main models, though reduced numbers made estimates more imprecise and restricted detailed analysis.

Figure 1: Directed acyclic graph of ethnicity on type 2 diabetes in its finalised form for mediation analysis. Individual paths are labelled with lower case letters. Abbreviations- T2DM: type 2 diabetes mellitus.

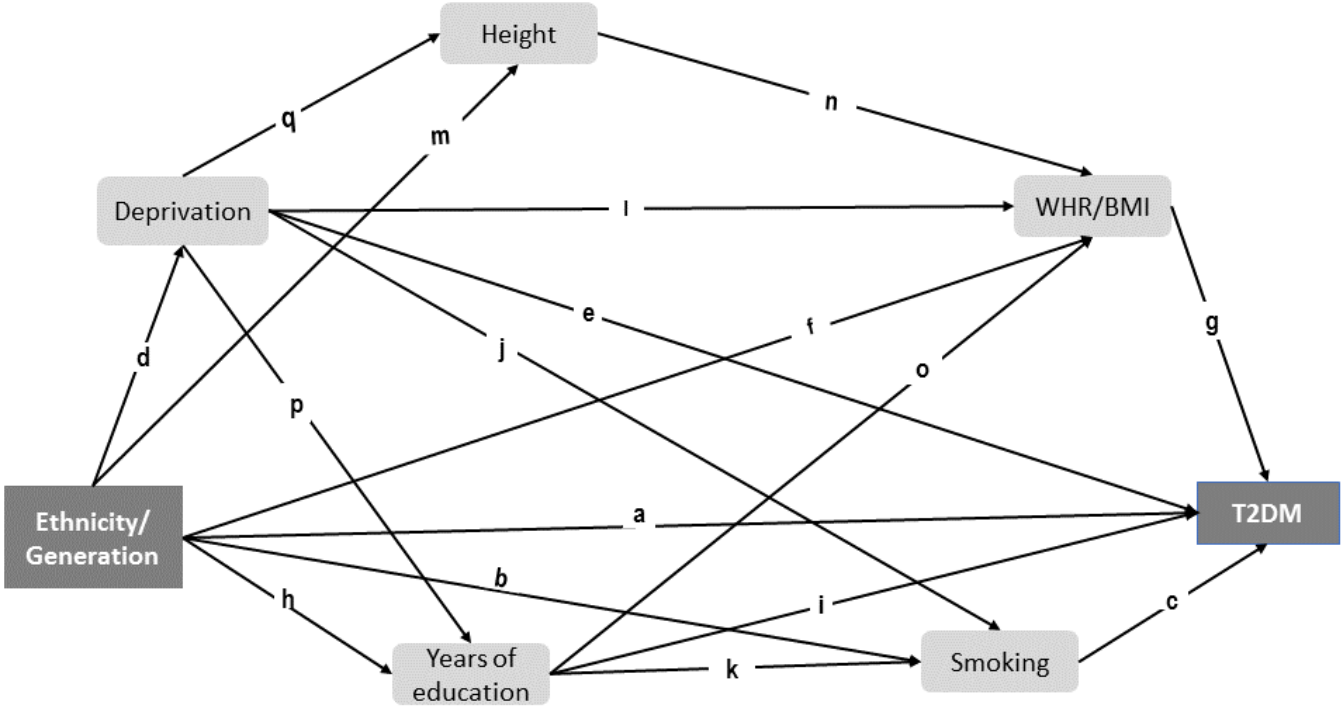

Figure 2: Forest plot of odds ratios (OR) with 95% confidence intervals (95% CI) for type 2 diabetes mellitus, age and sex adjusted (where appropriate). Abbreviations- SA: South Asians, EUR: Europeans, gen: generation, MixESA: mixed European/South Asians, MixEAC: mixed European/African Caribbeans.

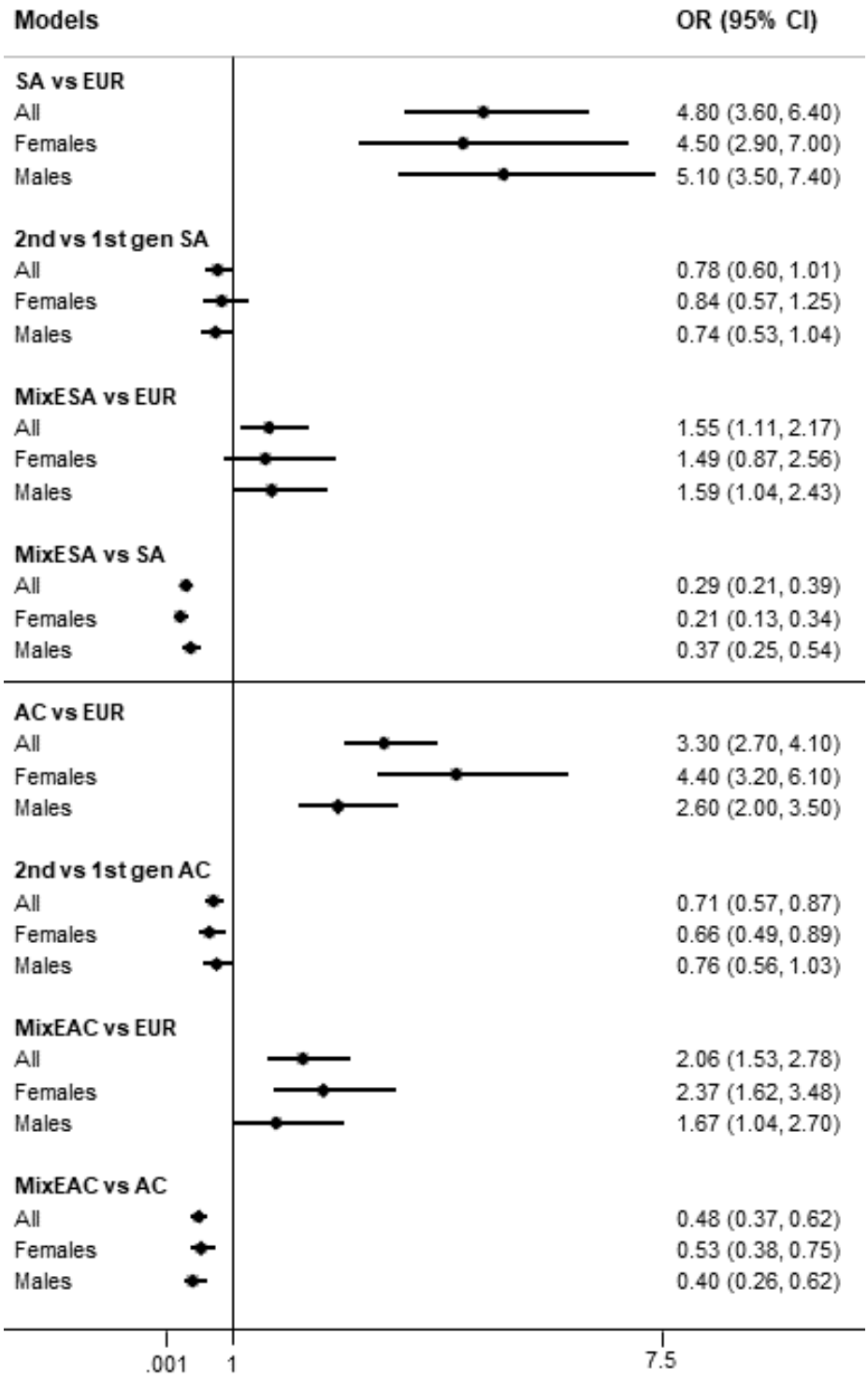

Table 1. Baseline characteristics of UK Biobank participants by ethnicity; European and South Asian origin groups. Data are n (%) and mean (standard deviation). First/ second generation assigned by year of migration. European and South Asian first and second generation groups are age and sex-matched (2:1:1).

	European			South Asian second generation			South Asian first generation		
	All	Males	Females	All	Males	Females	All	Males	Females
n (%)	2230	1072 (48)	1158 (52)	1115	536 (48)	579 (52)	1115	536 (48)	579 (52)
Age, yrs	47.1±6.7	46.6±6.4	47.5±6.9	46.6±6.7	46.2±6.5	47.0±6.9	46.9±6.7	46.5±6.5	47.3±6.9
Ever smoked, n (%)	667 (30)	362 (34)	305 (26)	197 (18)	140 (26)	57 (10)	155 (14)	133 (25)	22 (4)
Most deprived Townsend quintile, n (%)	457 (21)	229 (21)	228 (20)	425 (38)	209 (39)	216 (37)	450 (41)	244 (46)	206 (36)
Townsend index	-1.19 (0.25)	-1.10 (0.22)	-1.28 (0.25)	0.32 (0.40)	0.36 (0.49)	0.29 (0.29)	0.62 (0.42)	0.96 (0.21)	0.31 (0.31)
Years of education-derived	15.8±0.7	15.8± 0.3	15.8 ± 1.0	15.0 ± 1.3	15.1 ± 1.4	14.8 ± 1.2	15.3± 0.8	15.8±0.5	14.8± 0.7
Physical activity >10 mins, days/ wk	3.5±2.3	3.5±2.3	3.5±2.3	3.4±2.3	3.3±2.4	3.5±2.3	3.4±2.4	3.4±2.3	3.4±2.4
Waist/ hip ratio	0.86±0.09	0.92±0.07	0.80±0.07	0.88±0.09	0.93±0.06	0.83±0.08	0.89±0.09	0.94±0.06	0.84±0.08
Weight, kg	79±17	87±15	71±15	75±16	82±15	69±14	73±14	80±13	68±13
Height, cm	170±10	177±7	164±7	165±9	172±7	159±6	164±9	171±6	158±6
BMI, kg/m^2	27±5	28±4	26±5	27±5	28±4	27±6	27±5	27±4	27±5
Fat mass, kg	23.8±9.8	21.7±8.4	25.7±10.6	24.0±9.4	21.4±7.8	26.3±10.2	23.5±8.5	20.7±6.9	26.0±7.1
Fat %	29.7±8.5	24.1 ±5.9	35.0±7.2	31.5±8.4	25.5±5.3	37.1±6.8	31.7±8.2	25.5±5.0	37.5±6.0
Birth weight, kg	3.3±0.6	3.4±0.6	3.3±0.6	3.1±0.6	3.2±0.7	3.0±0.6	3.0±0.7	3.1±0.7	3.0±0.7
Diet									
"Healthy", n (%)	520 (23)	169 (16)	351 (30)	367 (33)	134 (25)	233 (40)	479 (43)	204 (38)	275 (48)
"Unhealthy", n (%)	582 (26)	364 (34)	218 (19)	134 (12)	90 (17)	44 (8)	81 (7)	44 (8)	37 (6)
Known type 2 diabetes, n (%)	48 (2)	30 (3)	18 (2)	91 (8)	53 (10)	38 (7)	115 (10)	65 (12)	50 (9)
All type 2 diabetes, n (%) *	75 (3)	43 (4)	32 (3)	124 (11)	71 (13)	53 (9)	155 (14)	92 (17)	63 (11)
HbA1c mmol/mol	34.5 (1.4)	35.1 (0.7)	34.0 (1.6)	38.2 (2.5)	39.0 (1.9)	37.5 (2.8)	39.1 (2.3)	39.8 (1.9)	38.4 (2.5)
Glucose mmol/l	4.97 (0.16)	5.06 (0.12)	4.89 (0.15)	5.18 (0.27)	5.29 (0.20)	5.08 (0.29)	5.23 (0.32)	5.26 (0.31)	5.21 (0.33)

*Includes known (algorithmically defined) diabetes and HbA1c > 47 mmol/mol

Table 2. Baseline characteristics of UK Biobank participants by ethnicity; European and South Asian origin groups. Data are n(%) and mean(standard deviation). European, Mixed European/South Asian and South Asian groups are age and sex-matched (4:1:4).

	European			Mixed European/ South Asian			South Asian		
	All	Males	Females	All	Males	Females	All	Males	Females
n (%)	3324	1392 (41.9)	1932 (58.1)	831	348 (41.9)	483 (58.1)	3317	1392(41.9)	1925(58.1)
Age, yrs	52.3±8.5	51.7±8.4	52.6±8.5	52.2±8.5	51.5±8.4	52.6±8.5	52.1±8.5	51.6±8.5	52.5±8.5
Ever smoked, n (%)	1092 (33)	516 (37)	576 (30)	309 (37)	141 (40.5)	168 (35)	472 (14)	367 (26.5)	105 (5.5)
Most deprived Townsend quintile, n (%)	622 (19)	248 (18)	374 (19)	259 (31)	115 (33)	144 (30)	1239 (37)	570 (41)	669 (35)
Townsend index	-1.47 (0.33)	-1.49 (0.32)	-1.45 (0.33)	-0.26 (0.37)	-0.15 (0.46)	-0.34 (0.26)	0.28 (0.37)	0.48 (0.39)	0.14 (0.28)
Years of education-derived	15.2 ± 1.0	15.3± 0. 9	15.2 ± 1.2	15.9 ± 1.4	16.0 ± 1.4	15.9 ± 1.4	14.9± 0.7	15.4 ±0.4	14.6 ± 0.6
Physical activity >10 mins, days/ wk	3.6±2.3	3.6±2.3	3.5±2.4	3.5±2.4	3.5±2.4	3.6±2.4	3.5±2.4	3.4±2.3	3.6±2.4
Waist/ hip ratio	0.86±0.09	0.93±0.06	0.81±0.07	0.86±0.09	0.93±0.07	0.82±0.07	0.89±0.09	0.95±0.06	0.85±0.07
Weight, kg	77.6±16.3	85.9±14.3	71.7±15.1	74.6±15.7	82.7±14.8	68.7±13.7	72.6±14.2	78.9±13.3	68.1±13.0
Height, cm	169±9	176±7	163±6	167±9	174±7	161±6	163±9	171±7	157±6
BMI, kg/m^2	27.2±4.9	27.6±4.3	26.9±5.3	26.8±4.9	27.2±4.4	26.5±5.3	27.4±4.7	27.1±4.1	27.7±5.0
Fat mass, kg	25±10	22±8	27±11	24±10	21±9	26±10	24±9	21±7	27±9
Fat %	31±9	24±6	36±7	31±9	24.5±6	36±7	33±8	26±5	38±6
Birth weight, kg	3.3±0.6	3.4±0.7	3.2±0.6	3.2±0.7	3.2±0.7	3.1±0.6	3.0±0.7	3.2±0.7	3.0±0.7
Diet									
"Healthy", n (%)	785 (24)	208 (15)	577 (30)	240 (29)	68 (20)	172 (36)	1531 (46)	515 (37)	1016 (53)
"Unhealthy", n (%)	811 (24)	460 (33)	351 (18)	183 (22)	106 (31)	77 (16)	212 (6)	125 (9)	87 (5)
Known type 2 diabetes, n (%)	107 (3)	74 (5)	33 (2)	39 (5)	25 (7)	14 (3)	477 (14)	243 (18)	234 (12)
All type 2 diabetes, n (%) *	140 (4)	88 (6)	52 (3)	52 (6)	33 (10)	19 (3)	599 (18)	299 (22)	300 (16)
HbA1c mmol/mol	35.1 (1.7)	35.5 (1.5)	34.8 (1.7)	36.6 (2.2)	37.8 (2.1)	35.8 (2.0)	40.2 (2.6)	41.0 (2.5)	39.6 (2.5)
Glucose mmol/l	5.01 (0.14)	5.07 (0.13)	4.96 (0.14)	5.10 (0.29)	5.31 (0.32)	4.95 (0.14)	5.34 (0.30)	5.47 (0.34)	5.25 (0.23)

*Includes known (algorithmically defined) diabetes and HbA1c > 47 mmol/mol

Table 3. Baseline characteristics of UK Biobank participants by ethnicity; European and African Caribbean origin groups. Data are n(%) and mean (standard deviation). First/ second generation assigned by year of migration. European and African Caribbean first and second generation groups are age and sex-matched (2:1:1).

	European			African Caribbean second generation			African Caribbean first generation		
	All	Males	Females	All	Males	Females	All	Males	Females
n (%)	4400	1886 (43)	2514 (57)	2200	943 (43)	1257 (57)	2200	943 (43)	1257 (57)
Age, yrs	47.7±5.8	47.3±5.8	48.0±5.8	47.5±5.7	47.0±5.6	47.8±5.7	47.7±5.8	47.3±5.7	48.1±5.9
Ever smoked, n (%)	1322 (30)	598 (32)	724 (29)	655 (30)	329 (35)	326 (26)	344 (16)	223 (24)	121 (10)
Most deprived Townsend quintile, n (%)	905 (21)	406 (22)	499 (20)	1323 (60)	554 (59)	769 (61)	1525 (70)	665 (71)	860 (69)
Townsend index	-1.2 (0.27)	-1.12 (0.26)	-1.26 (0.26)	2.28 (0.27)	2.23 (0.19)	2.31 (0.32)	3.14 (0.48)	3.36 (0.47)	2.97 (0.41)
Years of education-derived	15.9 ± 0.6	15.9 ± 0.5	15.9 ± 0.7	15.3 ± 0.9	14.8 ± 0.8	15.6 ± 0.7	15.9± 0.7	16.2 ± 0.7	15.8 ± 0.7
Physical activity >10 mins, days/ wk	3.5±2.3	3.5±2.3	3.5±2.3	3.6±2.3	3.6±2.3	3.6±2.3	3.5±2.3	3.5±2.3	3.5±2.2
Waist/ hip ratio	0.85±0.09	0.92±0.06	0.81±0.07	0.86±0.08	0.90±0.07	0.84±0.07	0.87±0.07	0.91±0.06	0.84±0.07
Weight, kg	78±17	87±15	72±15	84±18	89±16	80±18	83±16	85±14	81±17
Height, cm	169.5±9.4	177.4±6.8	163.6±6.2	169.2±8.9	176.2±6.8	164±6.3	167.4±8.4	173.8±6.6	162.6±6.0
BMI, kg/m^2	27.1±5.0	27.7±4.3	26.7±5.5	29.1±5.6	28.5±4.7	29.6±6.1	29.7±5.4	28.2±4.1	30.8±5.9
Fat mass, kg	24±10	22±8	26±11	28±12	22±9	32±13	29±12	22±8	34±12
Fat %	30.5±8.8	24.0±5.8	35.3±7.3	32.3±9.6	24.3±5.9	38.3±7.2	33.9±9.5	25.4±5.4	40.3±6.4
Birth weight, kg	3.3±0.6	3.4±0.6	3.3±0.6	3.2±0.7	3.4±0.7	3.1±0.6	3.4±0.7	3.5±0.7	3.3±0.7
Diet									
"Healthy", n (%)	1038 (24)	267 (14)	771 (31)	610 (28)	175 (19)	435 (35)	593 (27)	152 (16)	441 (35)
"Unhealthy", n (%)	1213 (28)	701 (37)	512 (20)	391 (18)	210 (22)	181 (14)	236 (11)	138 (15)	98 (8)
Known type 2 diabetes, n (%)	112 (3)	69 (4)	43 (2)	113 (5)	57 (6)	56 (5)	184 (8)	86 (9)	98 (8)
All type 2 diabetes, n (%) *	150 (3)	92 (5)	58 (3)	163 (7)	85 (9)	78 (6)	227 (10)	110 (12)	117 (9)
HbA1c mmol/mol	34.5 (1.4)	35.1 (0.7)	34.0 (1.6)	38.2 (2.5)	39.0 (1.9)	37.5 (2.8)	39.1 (2.3)	39.8 (1.9)	38.4 (2.5)
Glucose mmol/l	4.97 (0.16)	5.06 (0.12)	4.89 (0.15)	5.18 (0.27)	5.29 (0.21)	5.08 (0.29)	5.23 (0.32)	5.26 (0.31)	5.21 (0.33)

*Includes known (algorithmically defined) diabetes and HbA1c > 47 mmol/mol

Table 4. Baseline characteristics of UK Biobank participants by ethnicity; European and African Caribbean origin groups. Data are n (%) and mean(standard deviation). European, Mixed European/African Caribbean and African Caribbean groups are age and sex-matched (4:1:4).

	European			Mixed European/ African Caribbean			African Caribbean		
	All	Males	Females	All	Males	Females	All	Males	Females
n (%)	4180	1436(34.4)	2744(65.7)	1045	359(34.4)	686(65.7)	4180	1436(34.4)	2744(65.7)
Age, yrs	51.1±7.8	51.2±8.0	51.1±7.6	50.9±7.8	51.0±8.0	50.8±7.7	51.0±7.8	51.1±8.1	50.9±7.6
Ever smoked, n (%)	1289 (31)	515 (36)	774 (28)	439 (42)	160 (45)	279 (41)	844 (20)	412 (29)	432 (16)
Most deprived Townsend quintile, n (%)	829 (20)	296 (21)	533 (19)	498 (48)	161 (45)	337 (49)	2634 (63)	906 (63)	1728 (63)
Townsend index	-1.31 (0.25)	-1.22 (0.31)	-1.36 (0.19)	1.13 (0.52)	0.88 (0.44)	1.27 (0.51)	2.60 (0.26)	2.77 (0.34)	2.52 (0.15)
Years of education-derived	15.5 ± 0.9	15.7± 0.5	15.5 ± 1.0	15.3 ± 0.8	15.2 ±1.1	15.4 ±0.7	15.4± 0.8	15.4 ± 1.1	15.5± 0.7
Physical activity >10 mins, days/ wk	3.5±2.3	3.5±2.3	3.5±2.3	3.7±2.3	3.8±2.3	3.6±2.3	3.6±2.2	3.6±2.2	3.6±2.2
Waist/ hip ratio	0.85±0.09	0.93±0.06	0.81±0.07	0.85±0.09	0.92±0.07	0.82±0.08	0.87±0.08	0.91±0.07	0.84±0.07
Weight, kg	76.8±16.2	86.514.3	71.7±14.7	79.0±16.5	87.4±14.6	74.7±15.8	82.216.1	86.0±14.9	80.2±16.3
Height, cm	168±9	176±7	164±6	168±9	177±7	163±7	167±9	174±7	163±6
BMI, kg/m^2	27.1±5.0	27.8±4.2	26.8±5.3	28.0±5.3	28.0±4.3	28.0±5.8	29.6±5.5	28.3±4.3	30.3±5.9
Fat mass, kg	25±10	22±8	27±11	27±11	22±8	29±11	29±12	22±8	33±12
Fat %	32±9	24.4±5.9	35.6±7.3	32.9±9.2	24.5±5.9	37.3±7.4	34.7±9.4	25.3±5.8	39.6±6.8
Birth weight, kg	3.3±0.6	3.4±0.7	3.3±0.6	3.2±0.7	3.3±0.7	3.2±0.7	3.3±0.7	3.4±0.7	3.2±0.7
Diet									
"Healthy", n (%)	1024 (25)	213 (15)	811 (30)	297 (28)	58 (16)	239 (35)	1339 (32)	294 (21)	1045 (38)
"Unhealthy", n (%)	1052 (25)	504 (35)	548 (20)	239 (23)	117 (33)	122 (18)	455 (11)	210 (15)	245 (9)
Known type 2 diabetes, n (%)	109 (3)	51 (4)	58 (2)	51 (5)	22 (6)	29 (4)	416 (10)	183 (13)	233 (9)
All type 2 diabetes, n (%) *	144 (4)	65 (5)	79 (3)	70 (7)	26 (7)	44 (6)	534 (13)	227 (16)	307 (11)
HbA1c mmol/mol	34.8 (1.4)	35.1 (1.2)	34.6 (1.4)	36.5 (2.1)	37.2 (1.6)	36.2 (2.3)	39.2 (2.2)	40.0 (2.2)	38.7 (2.1)
Glucose mmol/l	4.98 (0.14)	5.04 (0.09)	4.95 (0.15)	4.95 (0.16)	4.97 (0.17)	4.94 (0.15)	5.10 (0.25)	5.23 (0.28)	5.04 (0.20)

*Includes known (algorithmically defined) diabetes and HbA1c > 47 mmol/mol

Table 5. Proportion of T2DM risk mediated by individual and joint effects of environmental risk factors comparing between generations of ethnic groups, and between those of mixed and non-mixed ethnicity. Paths are shown in Figure 1. Total and direct effects are presented as log odds ratios (age and sex adjusted). The mediated percentages shown are rounded to the nearest integer and for this reason they might not add up to the total mediated (*). Abbreviations- SA: South Asians, EUR: Europeans, gen: generation, MixESA: mixed European/South Asians, MixEAC: mixed European/African Caribbeans, WHR: waist to hip ratio, BMI: body mass index.

South Asians	**Total effect**	**Direct effect**	**% Mediated by**													
			Smoking	**Deprivation**	**WHR**	**Education**	Deprivation Smoking	Education Smoking	Deprivation WHR	Height WHR	Education WHR	Deprivation Education	Deprivation Education Smoking	Deprivation Height WHR	Deprivation Education WHR	**Total**
Pathway		(a)	(b-c)	(d-e)	(f-g)	(h-i)	(d-h-c)	(i-j-c)	(d-l-g)	(m-n-g)	(h-o-g)	(d-p-i)	(d-p-k-c)	(d-q-n-g)	(d-p-o-g)	
Generation																
2vs1 SA	-0.070	-0.046	-1	7	30	-4	0	0	3	1	-2	1	0	0	1	**34***
Ethnicity																
MixESA vs EUR	0.076	0.055	1	17	9	-7	1	0	6	3	-3	2	0	0	1	**28***
MixESA vs SA	-0.254	-0.201	-2	2	15	3	0	0	1	1	1	0	0	0	0	**21**

African Caribbeans	**Total effect**	**Direct effect**	**% Mediated by**													
			Smoking	**Deprivation**	**BMI**	**Education**	Deprivation Smoking	Education Smoking	Deprivation BMI	Height BMI	Education BMI	Deprivation Education	Deprivation Education Smoking	Deprivation Height BMI	Deprivation Education BMI	**Total**
Generation																
2vs1 AC	-0.100	-0.078	0	11	11	-3	0	0	2	1	0	0	0	0	0	**22**
Ethnicity																
MixEAC vs EUR	0.141	0.055	1	42	11	-1	1	0	5	0	-1	2	0	1	1	**61***
MixEAC vs AC	-0.154	-0.109	1	9	16	0	0	0	2	1	0	0	0	0	0	**29**

Discussion

We show that T2DM prevalence was substantially higher in SA and AC compared with Europeans. Second generation migrants, i.e. those born in the UK, of either SA or AC origin retain high rates of T2DM, but importantly, these are reduced by about a fifth compared to the first generation. In contrast, T2DM risks in people of MixESA, and of MixEAC ethnicities approach those of Europeans.

The ~20% reduction in T2DM risk in second versus first generation migrants is comparable to the long term effects of either lifestyle (27%) or metformin (18%) intervention on T2DM risk over 15 years in the Diabetes Prevention Programme (12). In mediation analyses, lower WHR in second-generation SA migrants appeared to account for a third of their lower risk of T2DM. In contrast, only a quarter of the reduced T2DM risk in second versus first generation AC migrants was accounted for by socioeconomic status and lower BMI. The impact of relatively modest differences in adiposity measures (0.01 for WHR in SA, and 0.6 kg/m^2 for BMI in AC), is striking. Unexplained differences in risk could be due to imprecise measurement of adiposity status, for example ectopic depots in liver and elsewhere, and socioeconomic deprivation, lack of such measures across the whole of life, or failure to account for other unknown exposures which influence T2DM risk. For example, physical activity and diet, the effects of which may not be wholly mediated by adiposity, although collected in UK Biobank, lacked precision and were not included in our models. We performed a sensitivity analysis on the subsample with self-reported birthweight, to capture early life determinants of T2DM, and observed little impact, though we acknowledge the limitations both of self-report and the reduced sample size. Importantly, these findings indicate the potential impact of environmental risk factor modification in addressing the higher rates of T2DM in these ethnic minority groups, confirming that these risks are not immutable.

In contrast to comparisons between first and second-generation migrants, who differ mainly in terms of environmental exposures, when comparing mixed ethnic groups, both genetic backgrounds and environmental exposures are likely to differ. T2DM risks in MixESA are about 70% lower than SA, and about 50% lower in MixEAC than AC. In both mixed populations, risks approached those of Europeans (1.5 fold excess in MixESA (versus 4.80-

fold in SAs), and 2.0 fold excess in MixEAC vs 3.30-fold in AC). Much of this residual excess could be accounted for by continued socioeconomic disadvantage; 28% of MixESA, and 50% of MixEAC people in this study resided in the most deprived neighbourhoods, compared to ~20% of Europeans. Socioeconomic deprivation contributed to much of the 1/3 and 2/3 excess of T2DM in MixESA and MixEAC people respectively, with a smaller direct contribution from adiposity measures (WHR and BMI, respectively). Interestingly, the greater levels of education in MixESA mitigated somewhat against the potential excess risk of T2DM. In contrast, markedly more favourable adiposity patterns or levels in the mixed ethnicity samples, approaching those of Europeans, played a greater part in accounting for the 70% lower risk of T2DM in MixESA versus SA, and for the halving in risk of T2DM in MixEAC versus AC.

Previous studies for African ancestry, all from the US, report conflicting findings. While some claim that only environmental factors (largely socioeconomic status and BMI) determine ethnic differences in T2DM (13), others aver that genetic factors play a stronger (14) or major role. There are no comparable studies of people of South Asian ethnicity.

People from ethnic minority groups, be it first generation migrants, their offspring, or those of mixed ethnicity, continue to experience significant social disadvantage compared to their white European counterparts. We show how this markedly increases T2DM risk, but importantly also show how improvement in social circumstances can reduce the burden of T2DM, to an extent similar to that observed in response to active lifestyle intervention or medication.

There are limitations of this analysis. Response rates to UK Biobank were <5%. Responders are likely to be healthier, and of higher socioeconomic status than non-responders, and this bias may differ by ethnicity. However, ethnic differences in diabetes prevalence in our study accord with those of previous, representative population cohorts (2). Further, we observed marked ethnic differences in socioeconomic deprivation, of similar magnitude to that anticipated from previous population studies (15). Years of education were derived from qualification(s) using the ISCED coding, which may not be comparable across countries due to differential access to educational opportunities, particularly for women. While UKB is

large, numbers of mixed ethnicity, and those of ethnic minority groups who could readily be matched by age and sex to people of mixed ethnicity, were modest, reducing the power of analysis. However, the inclusion of those of mixed ethnicity is unique. Measures of diet and exercise were deemed too imprecise to be included as mediators of the association between ethnicity and T2DM. There might be unmeasured confounding that was not taken into account and unquantifiable uncertainty in the percentages explained by mediation analysis. Although birthweight was not included in the final analysis because of the reduced power caused by missing data, sensitivity analysis showed the effect of this potential mediator was unlikely to be strong. Performing a mixed ethnicity and inter-generational analysis, using self-reported ethnicity, is a strength, as it enables the employment of different approaches to address the same question.

While ethnic specific genetic variants for hyperglycaemia/diabetes have been reported (16–18) and different effects of known variants observed (19), these currently are insufficient to account for the observed marked ethnic differences in diabetes risk. It could be that variants that account for adiposity measures rather than hyperglycaemia/diabetes should be studied, as the former appears key to accounting for differences in diabetes risk.

Our mixed ethnicity and inter-generational analyses show that environmental risk factors, largely related to socioeconomic status and obesity, make a strong contribution to ethnic differences in T2DM risk. Relatively modest differences in risk factors in second-generation migrants and in those of mixed ethnicity are associated with appreciably lower risks of T2DM. This effect is similar to that achieved by effective pharmaceutical interventions and provides evidence that the ethnic minority predisposition to diabetes is highly mutable.

References

1. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract. 2018;

2. Tillin T, Hughes AD, Godsland IF, Whincup P, Forouhi NG, Welsh P, et al. Insulin resistance and truncal obesity as important determinants of the greater incidence of diabetes in Indian Asians and African Caribbeans compared with Europeans: The Southall and Brent Revisited (SABRE) cohort. Diabetes Care. 2013;36(2):383–93.

3. Liem SS, Oemrawsingh P V., Cannegieter SC, Le Cessie S, Schreur J, Rosendaal FR, et al. Cardiovascular risk in young apparently healthy descendents from Asian Indian migrants in the Netherlands: The SHIVA study. Netherlands Hear J. 2009;

4. Gertraud Maskarinec, Yukiko Morimoto, Simone Jacobs, Andrew Grandinetti, Marjorie K. Mau LNK. Ethnic Admixture Affects Diabetes Risk in Native Hawaiians: The Multiethnic Cohort. Eur J Cardiovasc Prev Rehabil. 2016;70(9):1022–7.

5. E J Parra, C J Hoggart, C Bonilla, J M Norris, J A Marshall, R F Hamman, R E Ferrell, P M McKeigue MDS. Relation of type 2 diabetes to individual admixture and candidate gene polymorphisms in the Hispanic American population of San Luis Valley, Colorado. J Med Genet. 2004;41(e116):1–9.

6. Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK Biobank: Current status and what it means for epidemiology. Heal Policy Technol [Internet]. 2012 Sep 1 [cited 2019 Feb 20];1(3):123–6. Available from: https://www.sciencedirect.com/science/article/pii/S2211883712000597?via%3Dihub

7. Vickers D, Rees P. Creating the UK National Statistics 2001 output area classification. J R Stat Soc Ser A Stat Soc. 2007;

8. Townsend P. Townsend deprivation index. National database for primary care groups and trusts. 1998.

9. Eastwood S V., Mathur R, Atkinson M, Brophy S, Sudlow C, Flaig R, et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. PLoS One. 2016;

10. ISCED. International Standard Classification of EducationI S C E D 1997 [Internet].

ISCED. 1997. p. 1. Available from: http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm

11. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. Int J Epidemiol. 2013;

12. Nathan DM, Barrett-Connor E, Crandall JP, Edelstein SL, Goldberg RB, Horton ES, et al. Long-term effects of lifestyle intervention or metformin on diabetes development and microvascular complications over 15-year follow-up: The Diabetes Prevention Program Outcomes Study. Lancet Diabetes Endocrinol. 2015;

13. Piccolo RS, Pearce N, Araujo AB, McKinlay JB. The contribution of biogeographical ancestry and socioeconomic status to racial/ethnic disparities in type 2 diabetes mellitus: Results from the Boston Area Community Health Survey. Ann Epidemiol. 2014;

14. Qi L, Nassir R, Hall T, Ave OS, Kosoy R, Hall T, et al. Relationship between diabetes risk and admixture in postmenopausal African-American and Hispanic-American women. Diabetologia. 2012;55(5):1329–37.

15. People living in deprived neighbourhoods [Internet]. 2018 [cited 2019 Nov 8]. Available from: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/people-living-in-deprived-neighbourhoods/latest#people-living-in-the-most-deprived-10-of-neighbourhoods-by-ethnicity

16. Meigs JB, Grant RW, Piccolo R, López L, Florez JC, Porneala B, et al. Association of African genetic ancestry with fasting glucose and HbA 1c levels in non-diabetic individuals: The Boston Area Community Health (BACH) Prediabetes Study. Diabetologia. 2014;

17. Bien SA, Pankow JS, Haessler J, Lu YN, Pankratz N, Rohde RR, et al. Transethnic insight into the genetics of glycaemic traits. Diabetologia [Internet]. 2017;60(12):2384–98. Available from: https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85029429063

&origin=inward

18. Spracklen CN, Horikoshi M, Kim YJ, Lin K, Bragg F, Moon S, et al. Identification of type 2 diabetes loci in 433,540 East Asian individuals. bioRxiv. 2019;

19. Uribe-Salazar JM, Palmer JR, Haddad SA, Rosenberg L, Ruiz-Narváez EA. Admixture mapping and fine-mapping of type 2 diabetes susceptibility loci in African American women. J Hum Genet. 2018;