

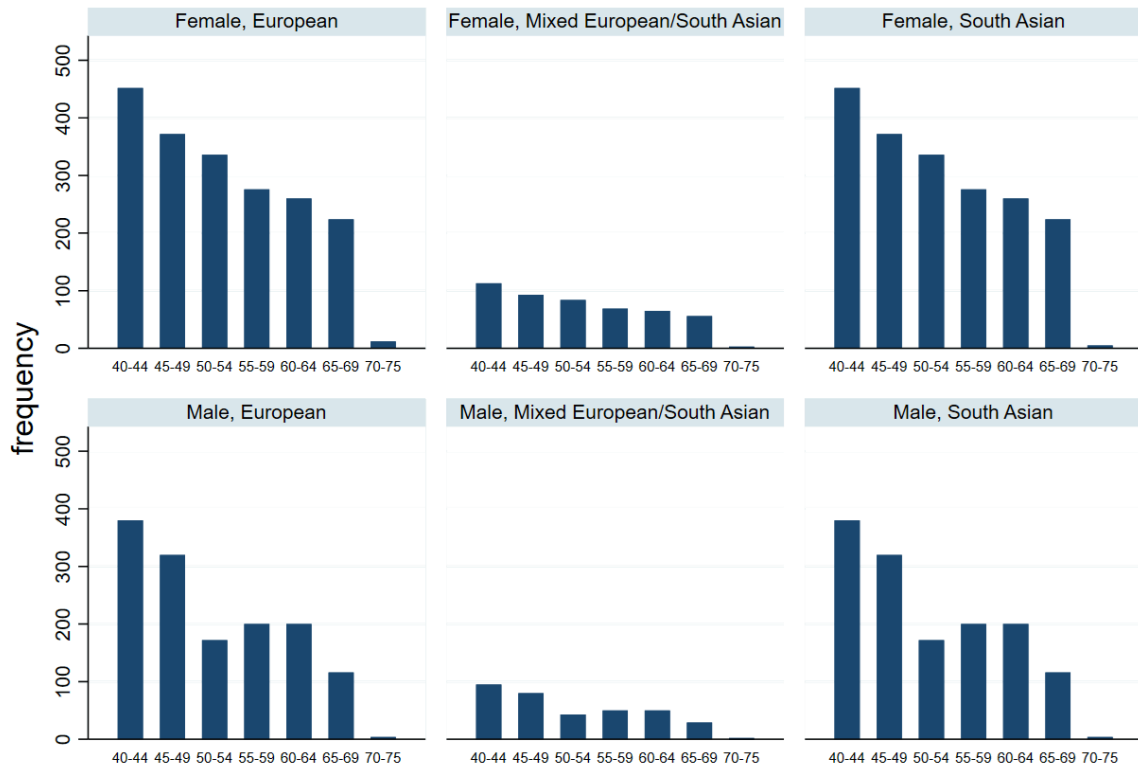
Supplement

S1: Matching procedure

The reference ethnic group is the mixed ethnicity or the second generation groups, depending on the comparison. Matching for each comparison was performed independently for sex and five year age strata.

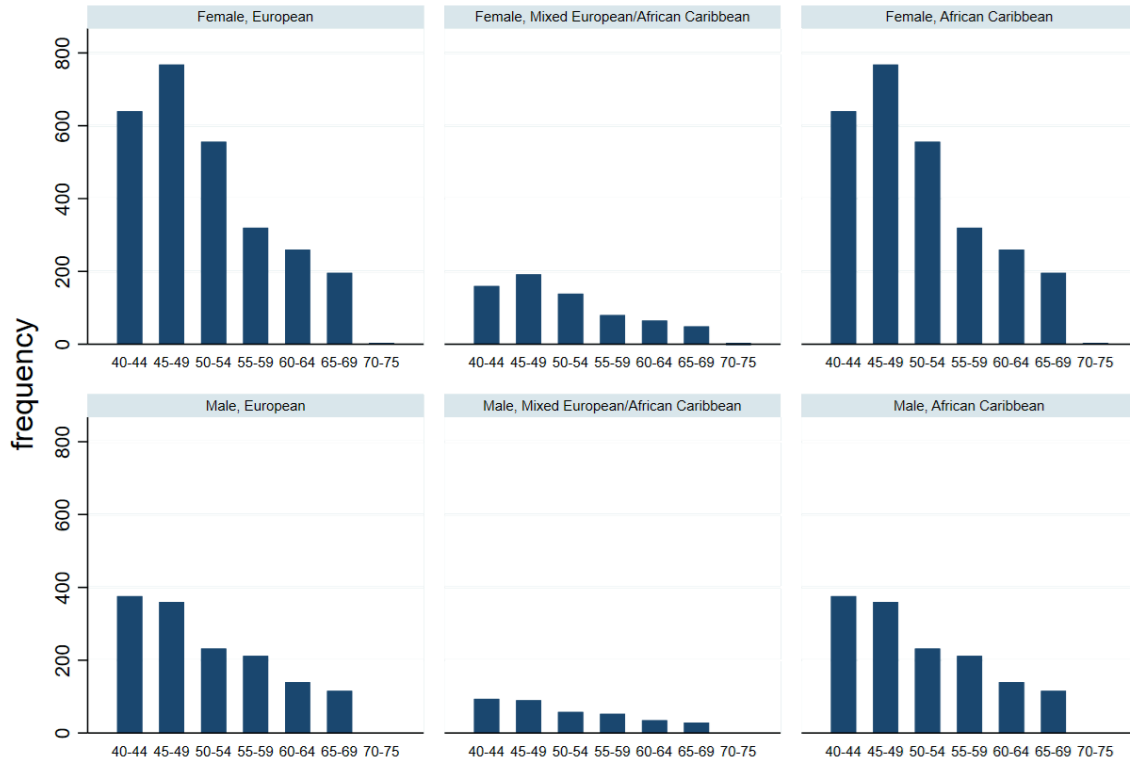
- Mixed European/South Asians (MixESA) (n=831) – South Asians (SA) – Europeans (1:4:4, N=7,479)

Figure s1.1: Frequency distribution by age, sex and European/South Asian ethnicity



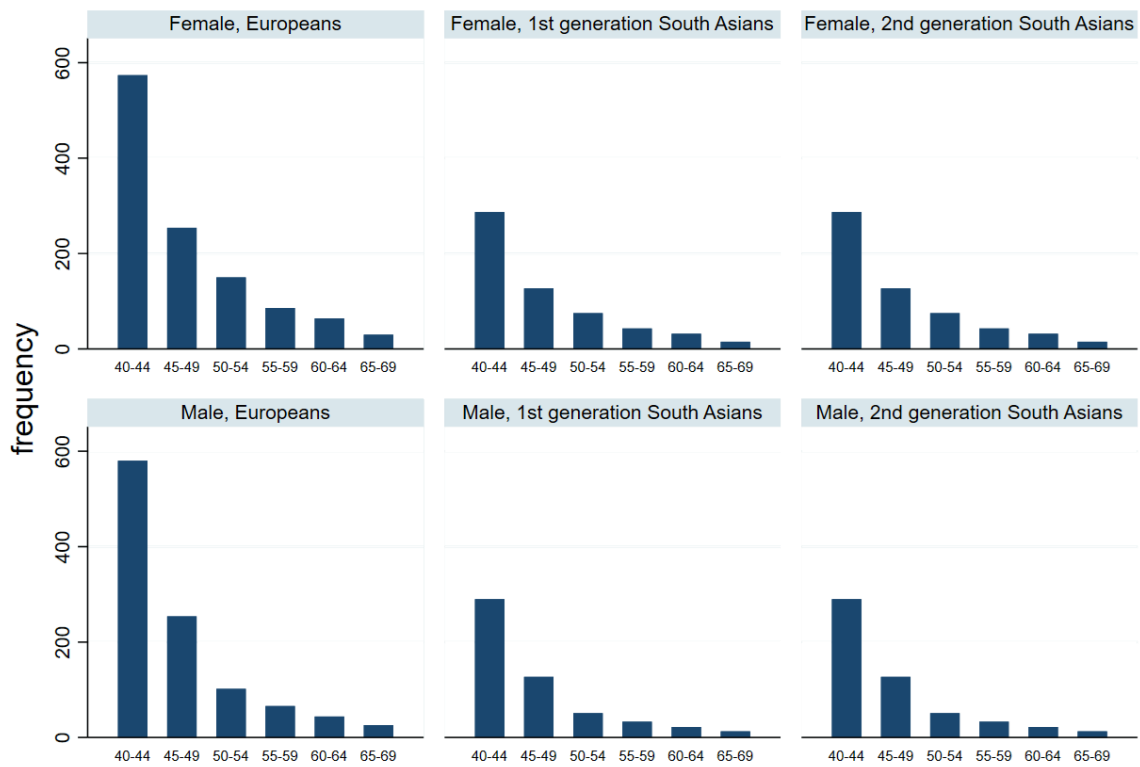
- Mixed European/African Caribbeans (MixEAC) (n=1,045) – African Caribbeans (AC) – Europeans (1:4:4, N=9,405)

Figure s1.2: Frequency distribution by age, sex and European/African Caribbean ethnicity



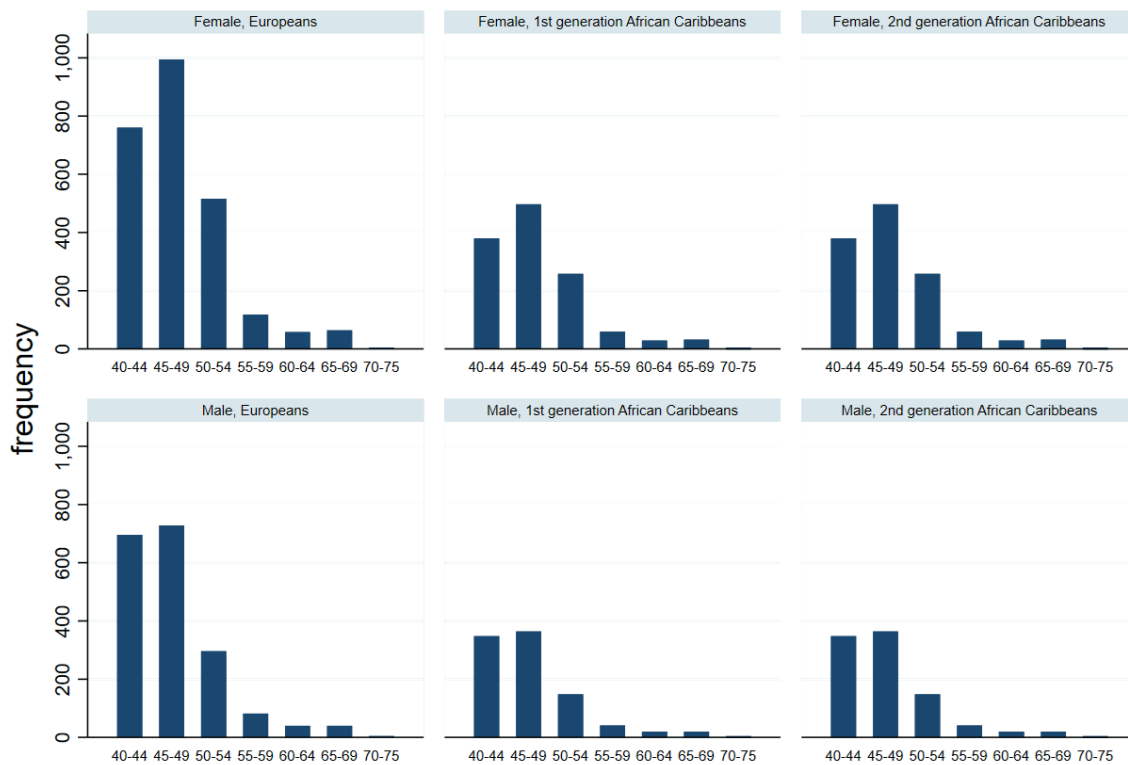
- Second generation SA (n=1,115) – First generation SA – Europeans (1:1:2, N=4,460)

Figure s1.3: Frequency distribution by age, sex and generations of South Asians



- Second generation AC (n=2,200) – First generation AC – Europeans (1:1:2, N=8,800)

Figure s1.4: Frequency distribution by age, sex and generations of African Caribbeans



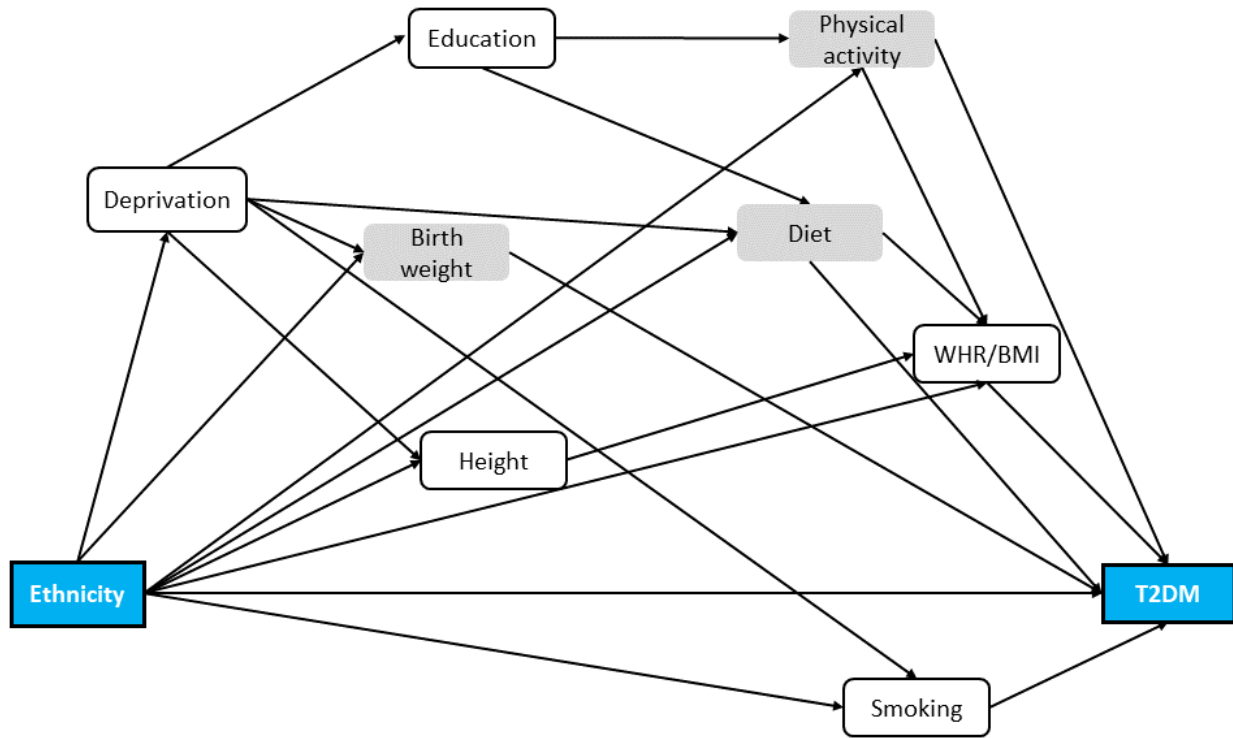
S2: Dietary patterns

We conducted principal component analysis (PCA) to identify underlying dietary patterns. Food frequency dietary data from the touchscreen questionnaire at the baseline assessment were used. All frequencies were quantified on a weekly basis and similar foods incorporated into food groups – fruits: dried and fresh, vegetables: salad, cooked and raw, fish: oily and non-oily, red meat: beef, lamb, mutton and pork.

The Kaiser–Meyer–Olkin test of sphericity and Barlett’s criterion was 0.63, implying high interrelationships between food variables. The orthogonal varimax rotation was used to derive optimal non-correlated components (dietary patterns) and we decided to retain two components. Based on the fact that factor loadings/correlation coefficients represent the correlation of each predicting variable with the dietary pattern score, higher absolute values indicate that the variable contributes more to the construction of this particular pattern. The dietary patterns were named according to scores of the predicting variables that correlated most with the component/pattern. The two main components emerging from the PCA were a ‘healthy’ dietary pattern (component 1) that was characterised by fruit, vegetable, fish and water consumption, and an ‘unhealthy’ diet variable which included red meat, processed meat and coffee drinking. In order to assign each individual to one of the components we created tertiles of scores for healthy and unhealthy patterns. Individuals in the two highest tertiles of the healthy pattern (component 1) and in the lowest of unhealthy pattern (component 2) were grouped as following a healthy diet. Individuals in the two highest tertiles of the unhealthy pattern (component 2) and in the lowest healthy pattern (component 1) were grouped as following an unhealthy diet.

S3: Mediation path analysis

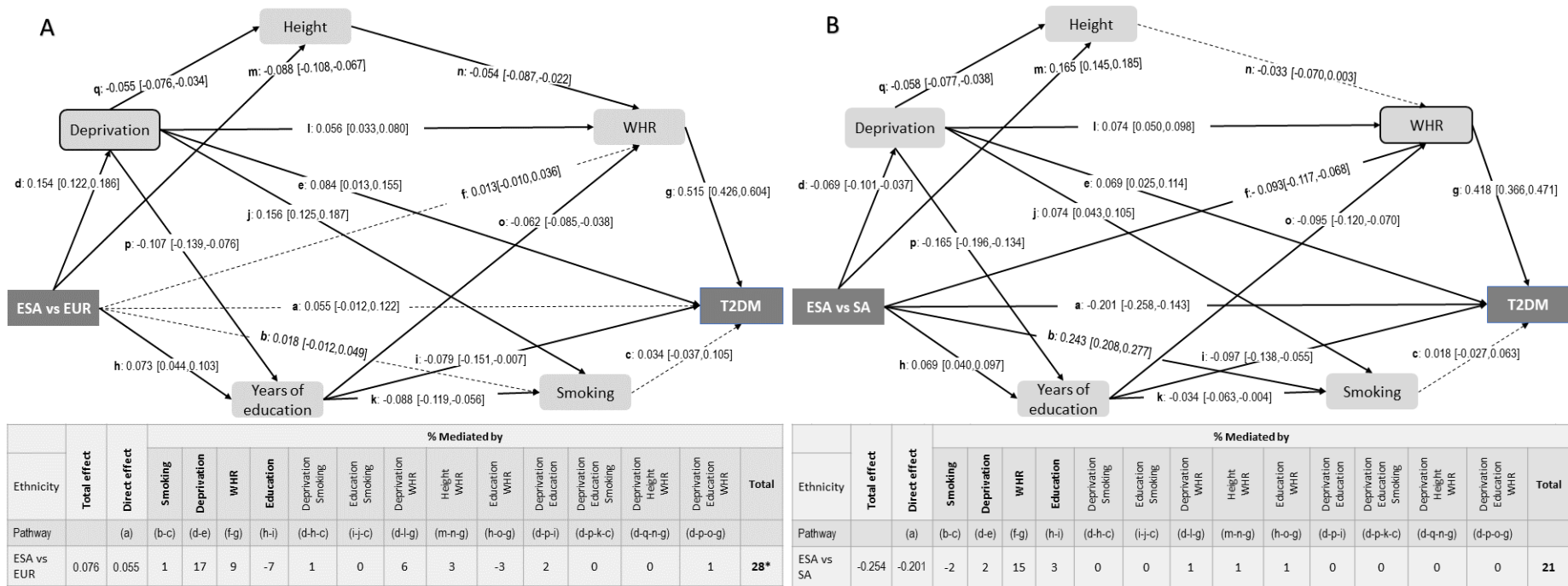
Figure s3.1: Directed acyclic graph of ethnicity on type 2 diabetes, including all the potential determinants of this relationship. (The grey coloured variables have not been carried forward to the subsequent analyses).

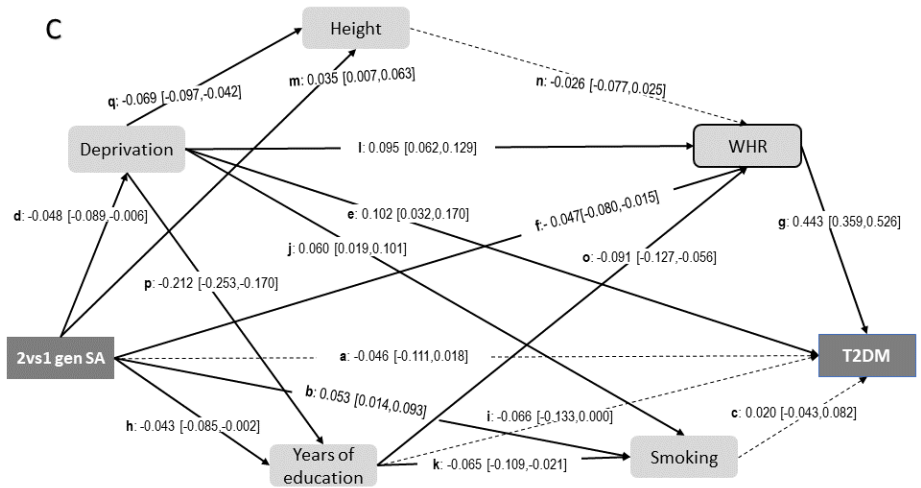


Datasets sample size:

MixESA (n=831) vs EUR	(1:4, N=4,155)
MixESA (n=831) vs SA	(1:4, N=4,148)
2 nd vs 1 st generation SA	(1:1, N=2,230)
MixEAC (n=1,045) vs EUR	(1:4, N=5,225)
MixEAC (n=1,045) vs AC	(1:4, N=5,225)
2 nd vs 1 st generation AC	(1:1, N=4,400)

Figure s3.2: Diagrams of mediational model between South Asian ethnic groups (A, B) and generations (C) and the proportion of the observed difference in type 2 diabetes prevalence, which is explained by five mediators (smoking, deprivation, WHR, height, and years of education) and their interrelationships. The dashed arrows indicate non-significant association and the numbers are standardised estimates, age and sex adjusted. The mediated percentages shown are rounded to the nearest integer and for this reason they might not be added up to the total (*).

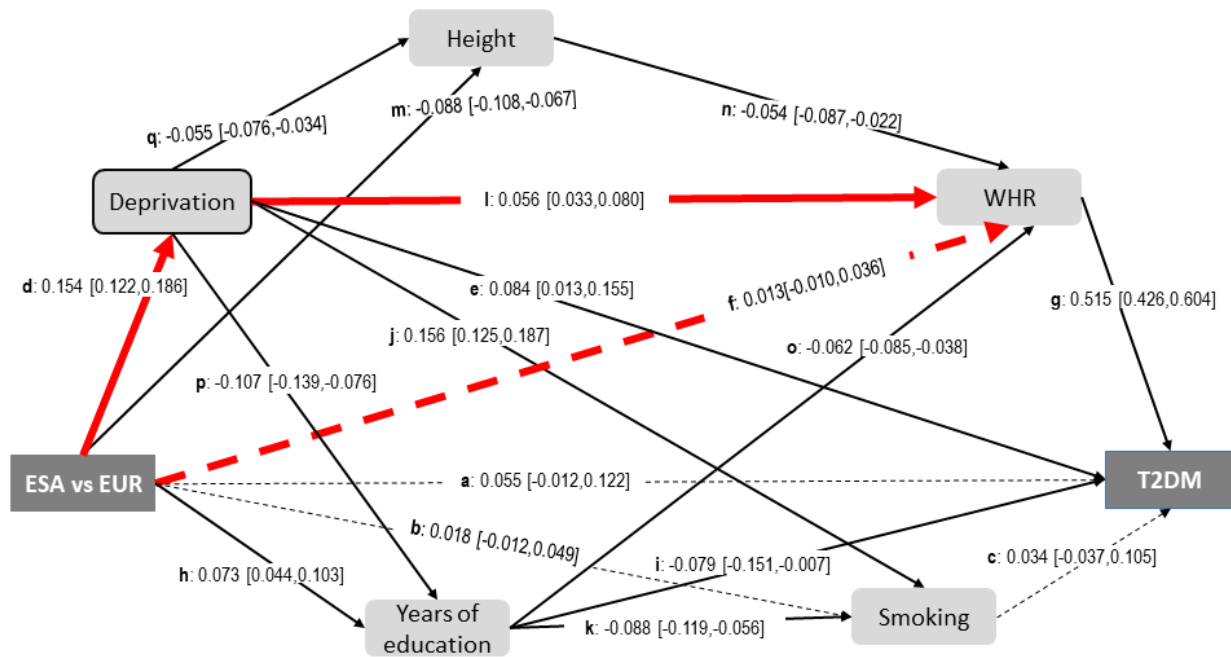




Generation	Total effect	Direct effect	% Mediated by													Total
			Smoking	Deprivation	WHR	Education	Deprivation Smoking	Education Smoking	Deprivation WHR	Height WHR	Education WHR	Deprivation Education	Deprivation Education Smoking	Deprivation Height WHR	Deprivation Education WHR	
Pathway	(a)	(b-c)	(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)		
2vs1 SA	-0.07	-0.046	-1	7	30	-4	0	0	3	1	-2	1	0	0	1	34*

When breaking down the component parts (red arrows Figure s2.3), we observed that ethnicity was strongly associated with deprivation ($\beta_{std}=0.154$, 95% CI 0.122 to 0.186), which in turn had a marked impact on WHR ($\beta_{std}=0.056$, 95% CI 0.033 to 0.080), attenuating any effect of ethnicity itself ($\beta_{std}=0.013$, 95% CI -0.010 to 0.036). Thus, a large part (37%) of the effect of ethnicity on the WHR in MixESA versus Europeans was mediated via deprivation.

Figure s3.3: Diagram of mediational model between Mixed Europeans/South Asian versus South Asians in type 2 diabetes prevalence, using five mediators (smoking, deprivation, WHR, height, and years of education) and their interrelationships, with focus on the pathway ethnicity-deprivation-WHR. The dashed arrows indicate non-significant association and the numbers are standardised estimates, age and sex adjusted.



When breaking down the component parts (red arrows Figure s2.4), we observed that there was a modest association of ethnicity with both deprivation ($\beta_{std}=-0.069$, 95% CI -0.101 to -0.037) and WHR ($\beta_{std}=-0.093$, 95% CI -0.117 to -0.068), with the latter having a greater association on diabetes ($\beta_{std}=0.418$, 95% CI 0.366 to 0.471) than deprivation itself ($\beta_{std}=0.069$, 95% CI 0.025 to 0.114). Thus, part (24%) of the effect of deprivation in accounting for the excess prevalence of T2DM in MixESA versus SA was mediated via WHR.

Figure s3.4: Diagram of mediational model between Mixed Europeans/South Asian versus South Asians in type 2 diabetes prevalence, using five mediators (smoking, deprivation, WHR, height, and years of education) and their interrelationships, with focus on the pathway deprivation-WHR-T2DM. The dashed arrows indicate non-significant association and the numbers are standardised estimates, age and sex adjusted.

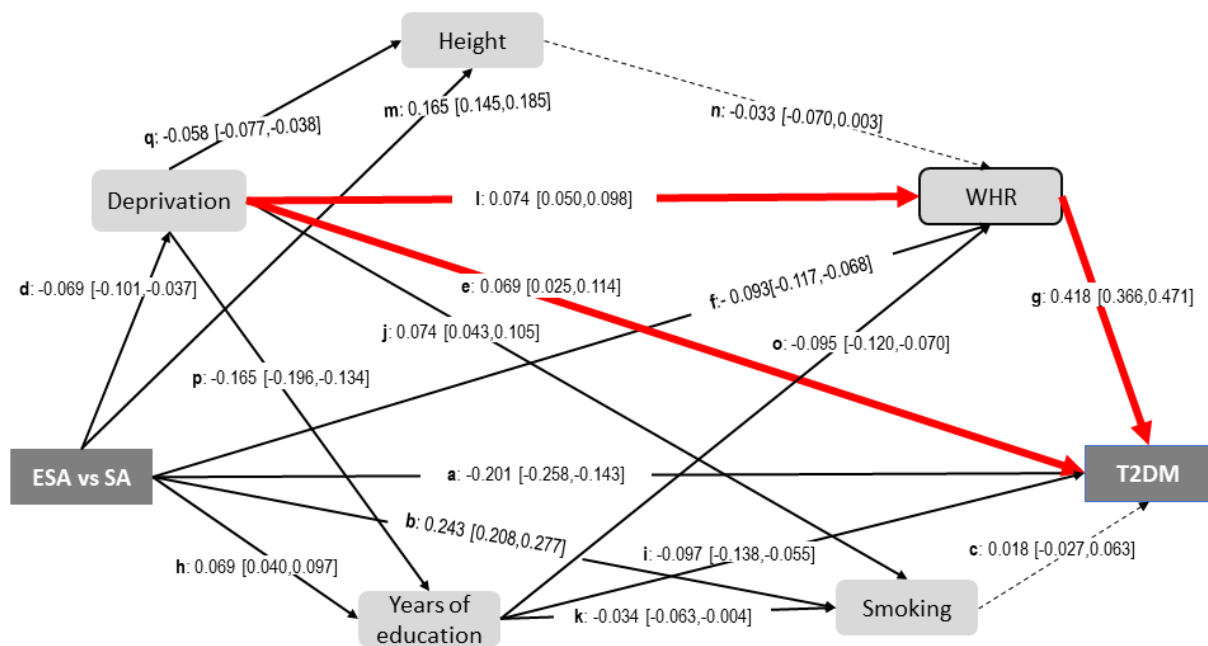
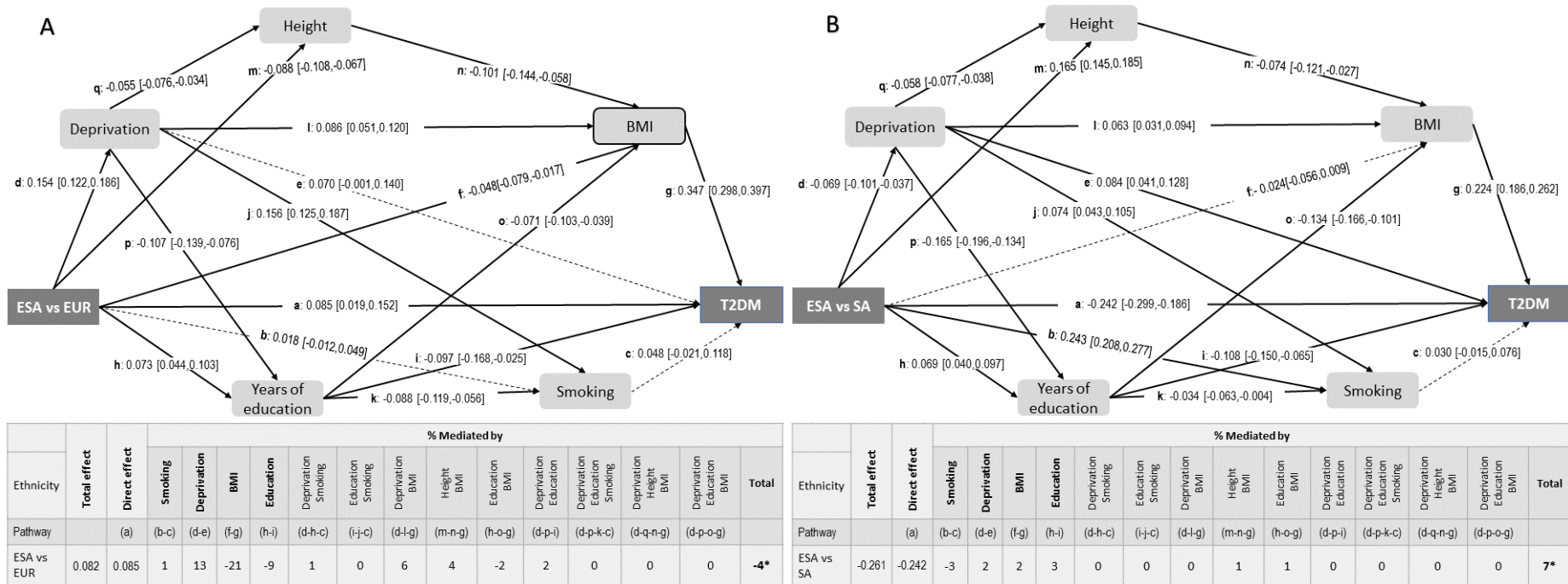
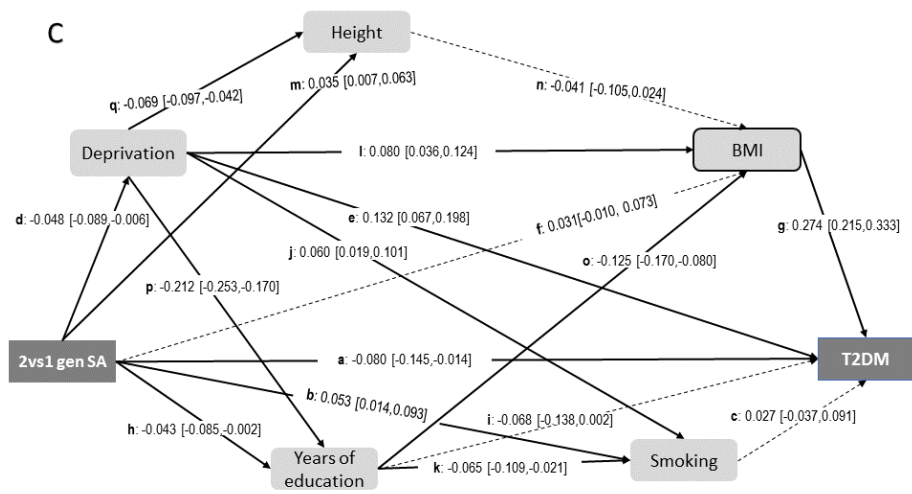


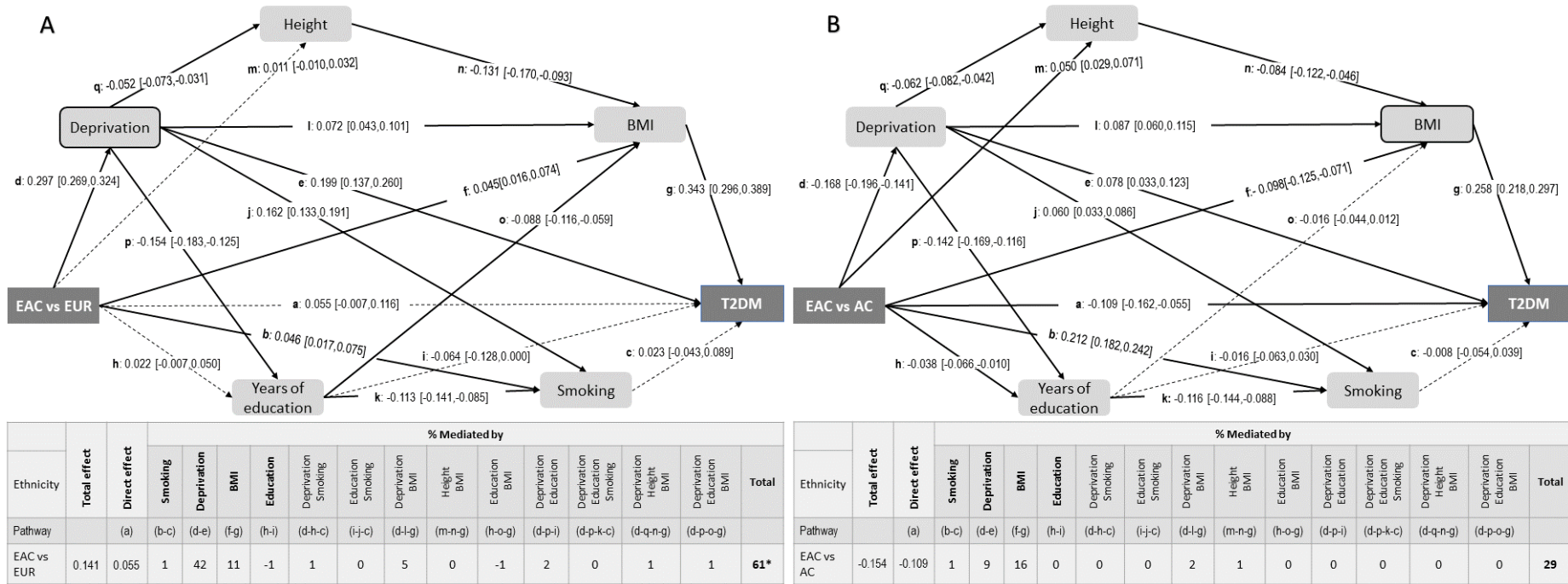
Figure s3.5: Diagrams of mediational model between South Asian ethnic groups (A, B) and generations (C) and the proportion of the observed difference in type 2 diabetes prevalence, which is explained by five mediators (smoking, deprivation, BMI, height, and years of education) and their interrelationships. The dashed arrows indicate non-significant association and the numbers are standardised estimates, age and sex adjusted. The mediated percentages shown are rounded to the nearest integer and for this reason they might not be added up to the total (*).

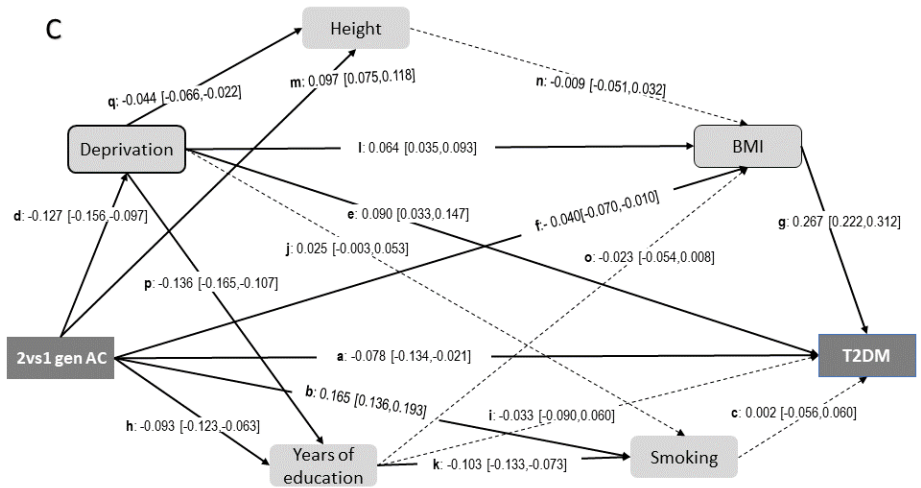




Generation	Total effect	Direct effect	% Mediated by												Total
			Smoking	Deprivation	BMI	Education	Deprivation Smoking	Education Smoking	Deprivation BMI	Height BMI	Education BMI	Deprivation Education	Deprivation Education Smoking	Deprivation Height BMI	
Pathway	(a)	(b-c)	(d-e)	(f-g)	(h-i)	(d-h-c)	(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)	
2vs1 SA	-0.074	-0.080	-1	8	-12	-4	0	0	1	0	-1	1	0	0	-8

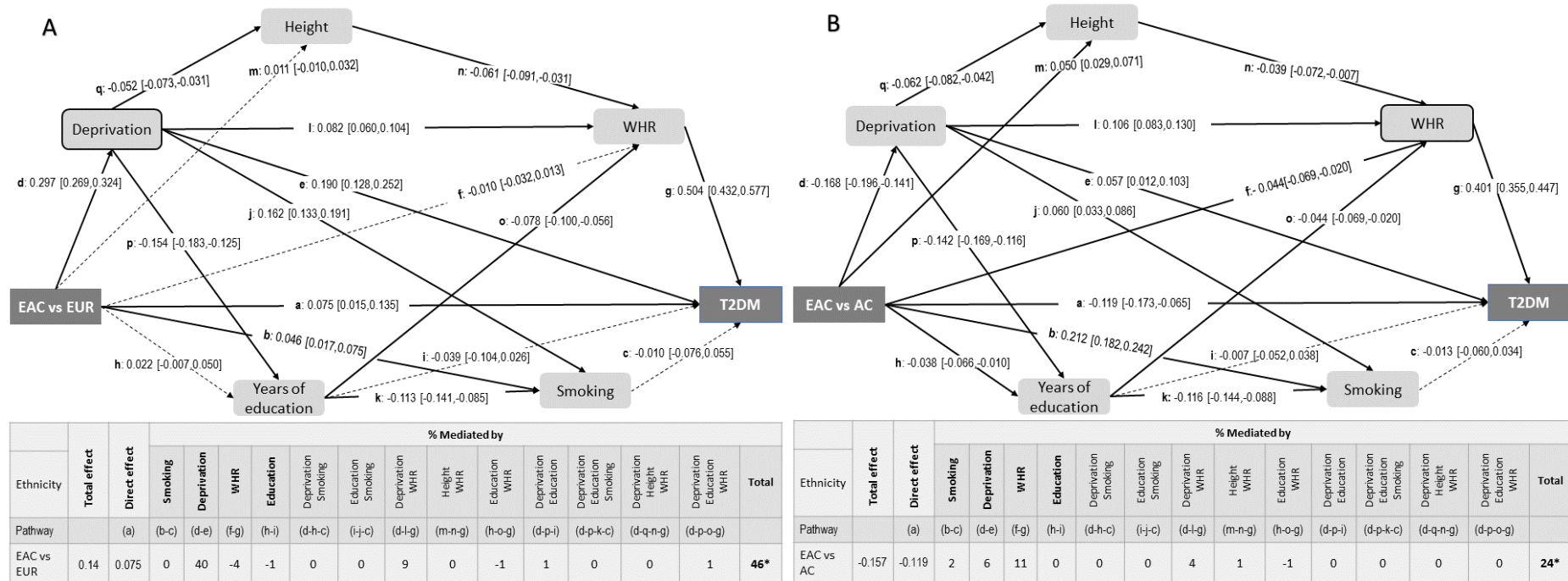
Figure s3.6: Diagrams of mediational model between African Caribbean ethnic groups (A, B) and generations (C) and the proportion of the observed difference in type 2 diabetes prevalence, which is explained by five mediators (smoking, deprivation, BMI, height and years of education) and their interrelationships. The dashed arrows indicate non-significant association and the numbers are standardised estimates, age and sex adjusted. The mediated percentages shown are rounded to the nearest integer and for this reason they might not be added up to the total (*).

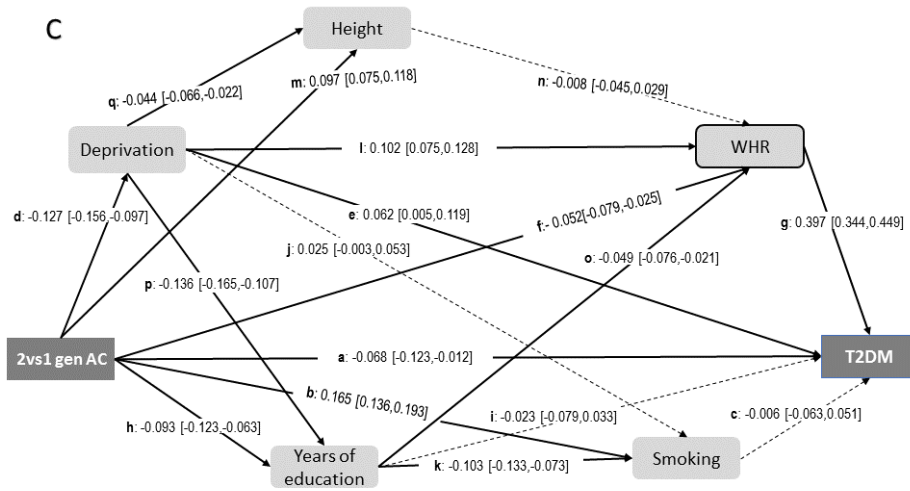




Generation	Total effect	Direct effect	% Mediated by												Total
			Smoking	Deprivation	BMI	Education	Deprivation Smoking	Education Smoking	Deprivation BMI	Height BMI	Education BMI	Deprivation Education	Deprivation Education Smoking	Deprivation Height BMI	
Pathway	(a)	(b-c)	(d-e)	(f-g)	(h-i)	(d-h-c)	(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)	
2vs1 AC	-0.1	-0.078	0	11	11	-3	0	0	2	1	0	0	0	0	22

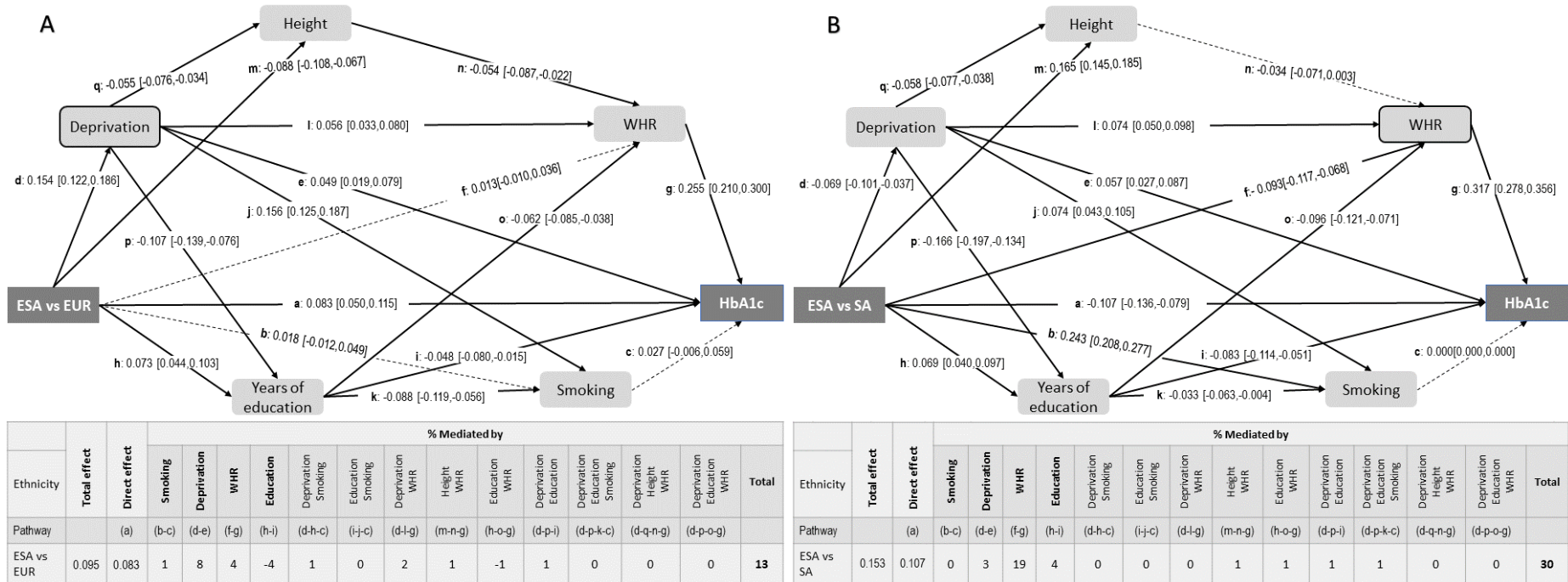
Figure s3.7: Diagrams of mediational model between African Caribbean ethnic groups (A, B) and generations (C) and the proportion of the observed difference in type 2 diabetes prevalence, which is explained by five mediators (smoking, deprivation, WHR, height and years of education) and their interrelationships. The dashed arrows indicate non-significant association and the numbers are standardised estimates, age and sex adjusted. The mediated percentages shown are rounded to the nearest integer and for this reason they might not be added up to the total (*).

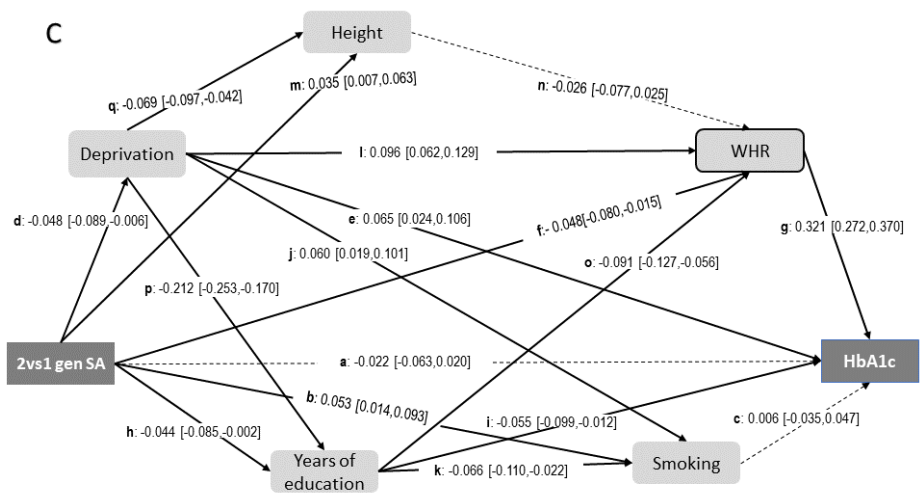




Generation	Total effect	Direct effect	% Mediated by													Total
			Smoking	Deprivation	WHR	Education	Deprivation Smoking	Education Smoking	Deprivation WHR	Height WHR	Education WHR	Deprivation Education	Deprivation Education Smoking	Deprivation Height WHR	Deprivation Education WHR	
Pathway	(a)	(b-c)	(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)		
2vs1 AC	-0.099	-0.068	1	8	21	-2	0	0	5	0	-2	0	0	0	32*	

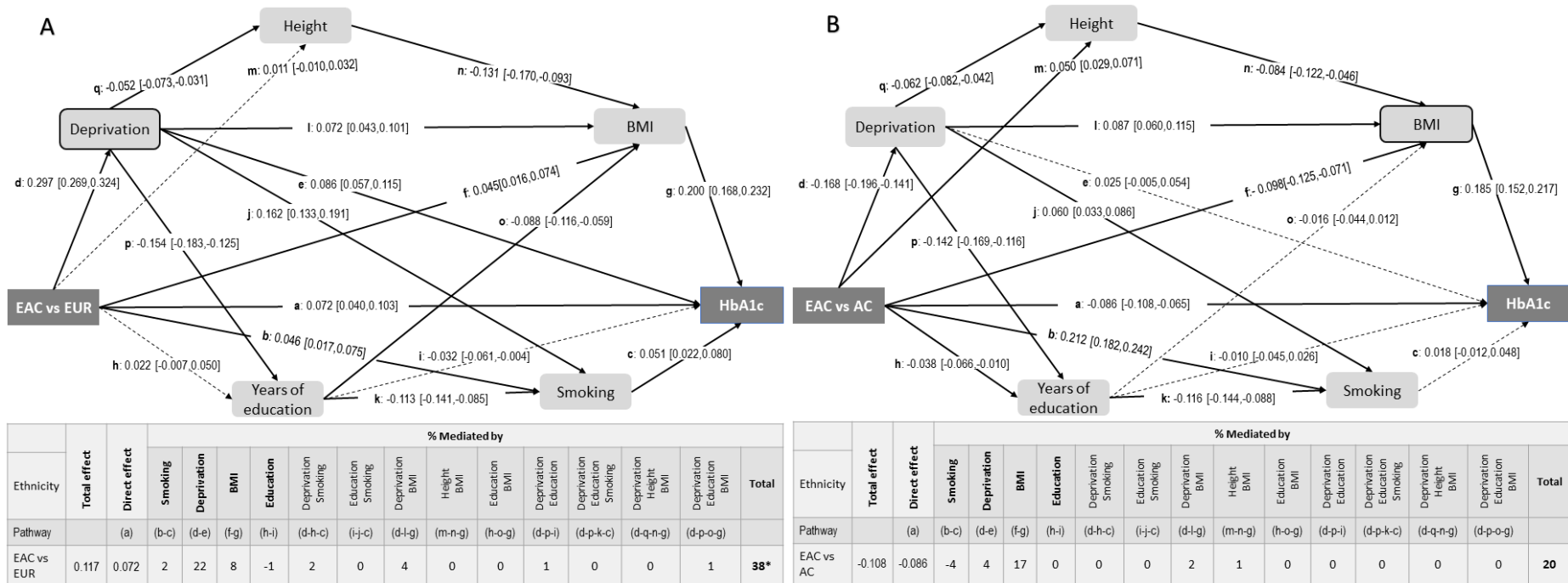
Figure s3.8: Diagrams of mediational model between South Asian ethnic groups (A, B) and generations (C) and the association with HbA1c levels, which is explained by five mediators (smoking, deprivation, WHR, height, and years of education) and their interrelationships. The dashed arrows indicate non-significant association and the numbers are standardised estimates, 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 (*).

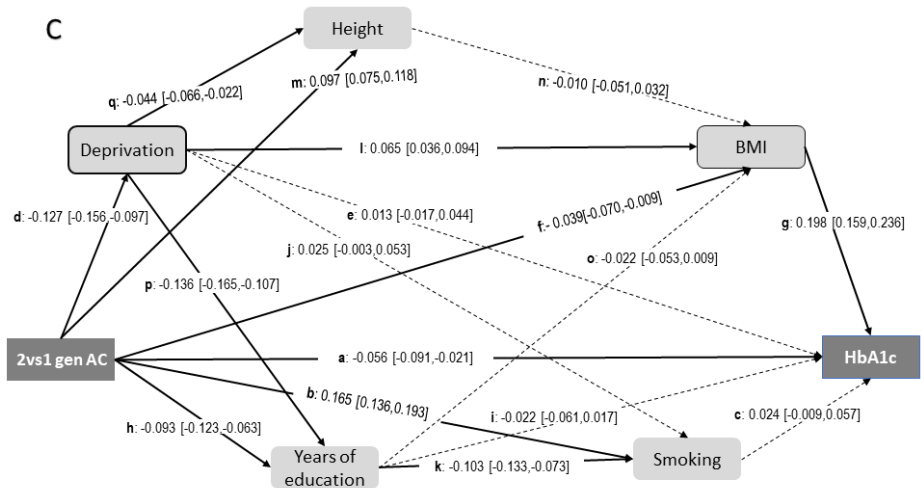




Generation	Total effect	Direct effect	% Mediated by													Total
			Smoking	Deprivation	WHR	Education	Deprivation Smoking	Education Smoking	Deprivation WHR	Height WHR	Education WHR	Deprivation Education	Deprivation Education Smoking	Deprivation Height WHR	Deprivation Education WHR	
Pathway	(a)	(b-c)	(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)		
2vs1 SA	-0.039	-0.022	-1	8	40	-6	0	0	4	1	-3	1	0	0	1	44

Figure s3.9: Diagrams of mediational model between African Caribbean ethnic groups (A, B) and generations (C) and the association with HbA1c levels, which is explained by five mediators (smoking, deprivation, BMI, height and years of education) and their interrelationships. The dashed arrows indicate non-significant association and the numbers are standardised estimates, 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 (*).



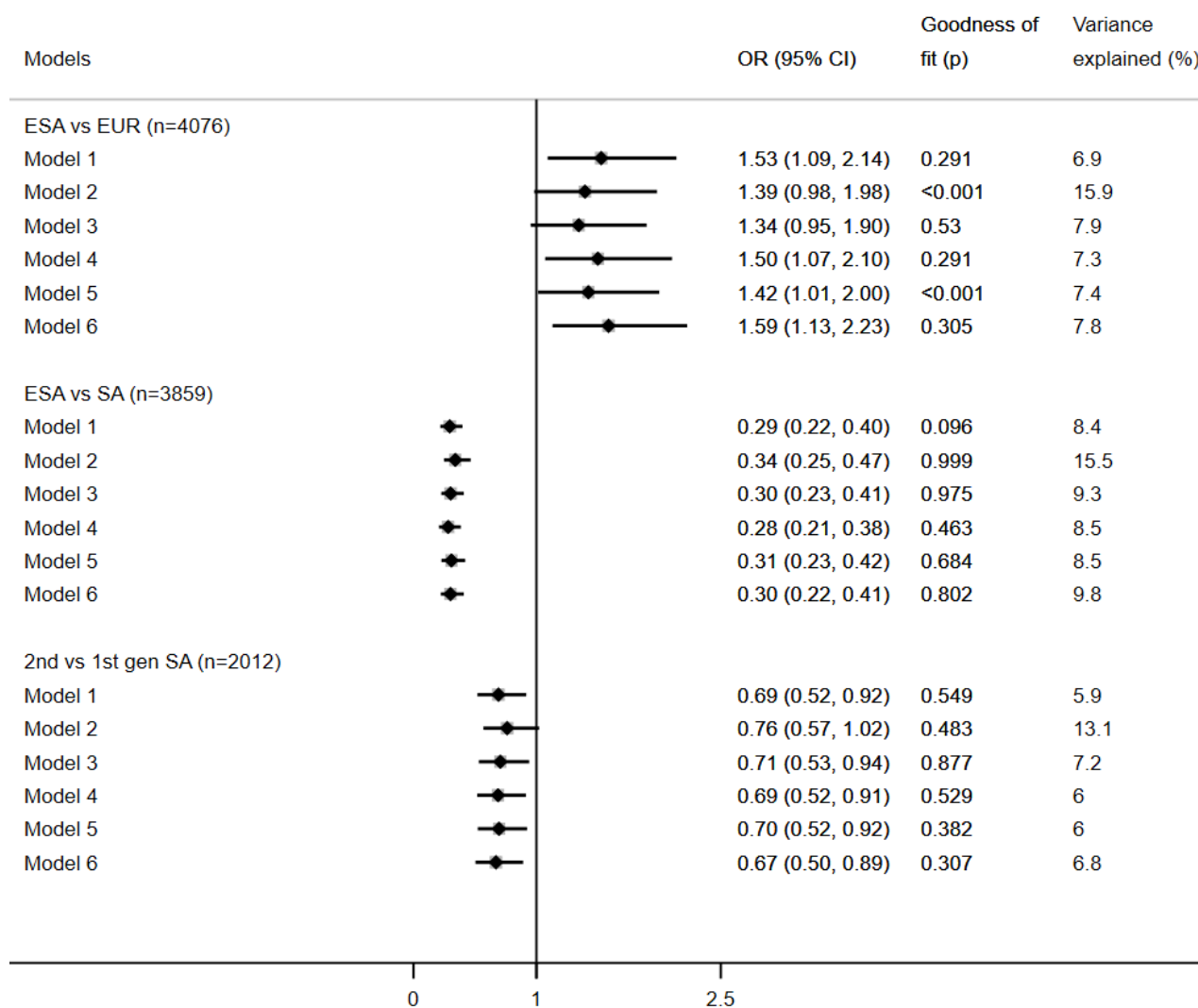


Generation	Total effect	Direct effect	% Mediated by												Total	
			Smoking	Deprivation	BMI	Education	Deprivation Smoking	Education Smoking	Deprivation BMI	Height BMI	Education BMI	Deprivation Education	Deprivation Education Smoking	Deprivation Height BMI		Deprivation Education BMI
Pathway	(a)	(b-c)	(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)		
2vs1 AC	-0.062	-0.056	-6	3	13	-3	0	0	3	0	-1	1	0	0	0	9*

S4: Regression models and sensitivity analysis

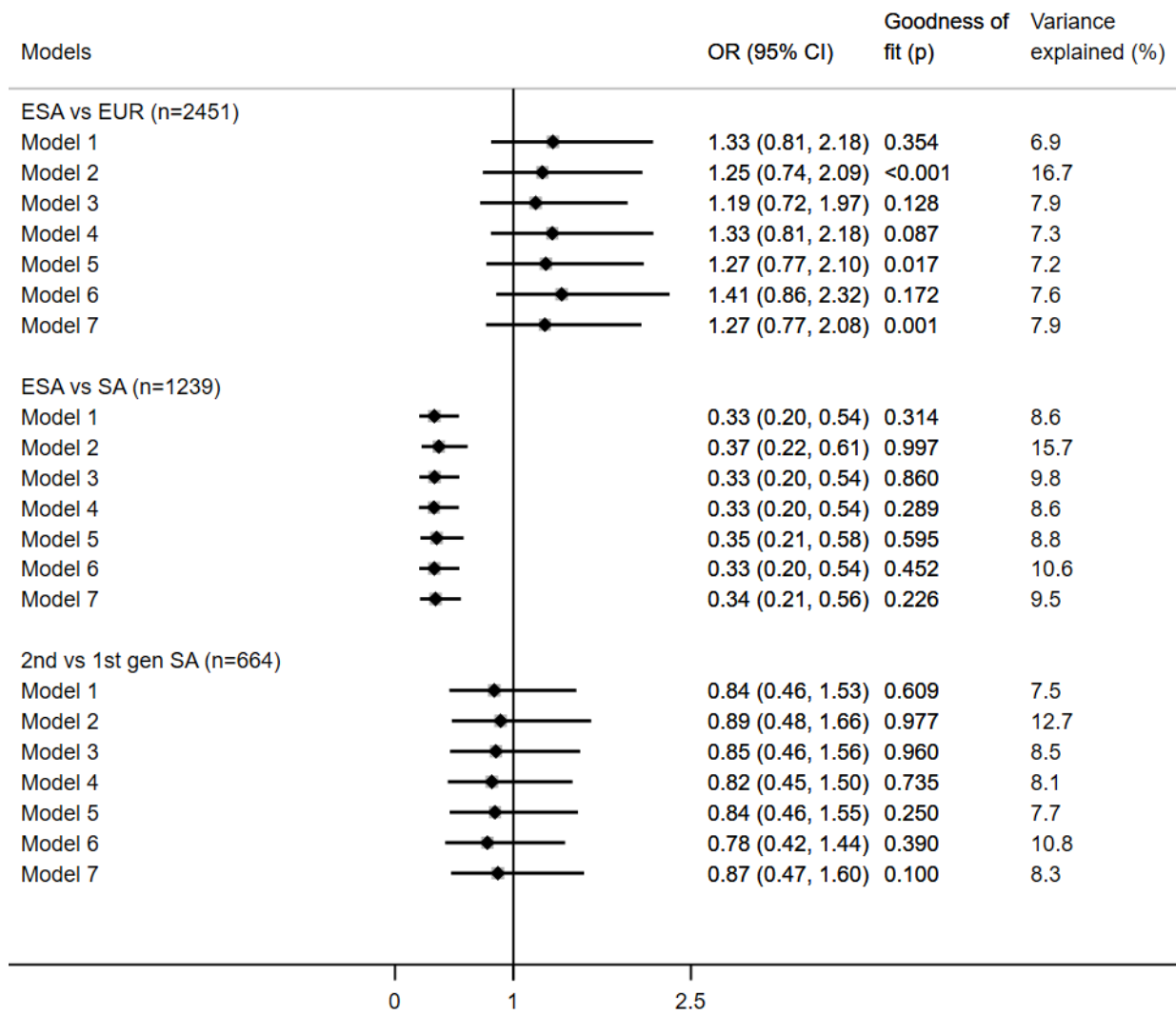
Multivariate logistic regression was used to examine the contribution of each different risk factor /determinant on the association of ethnicity with type 2 diabetes. In particular, six different models were applied: Model 1, adjusted for age and sex; Model 2, as per Model 1 plus WHR (for SA) or BMI (for AC); Model 3, as per Model 1 plus deprivation index; Model 4, as per Model 1 plus ever smoking; Model 5, as per Model 1 plus height; Model 6, as per Model 1 plus years of education. The results are shown below in forests plots (figures s6 and s8). Sensitivity analysis was conducted on full case data applying on this data subset Models 1-6 and Model 7, as per Model 1 plus birth weight (figures s7 and s9).

Figure s4.1: Forest plots of multivariate regression models for diabetes in South Asians



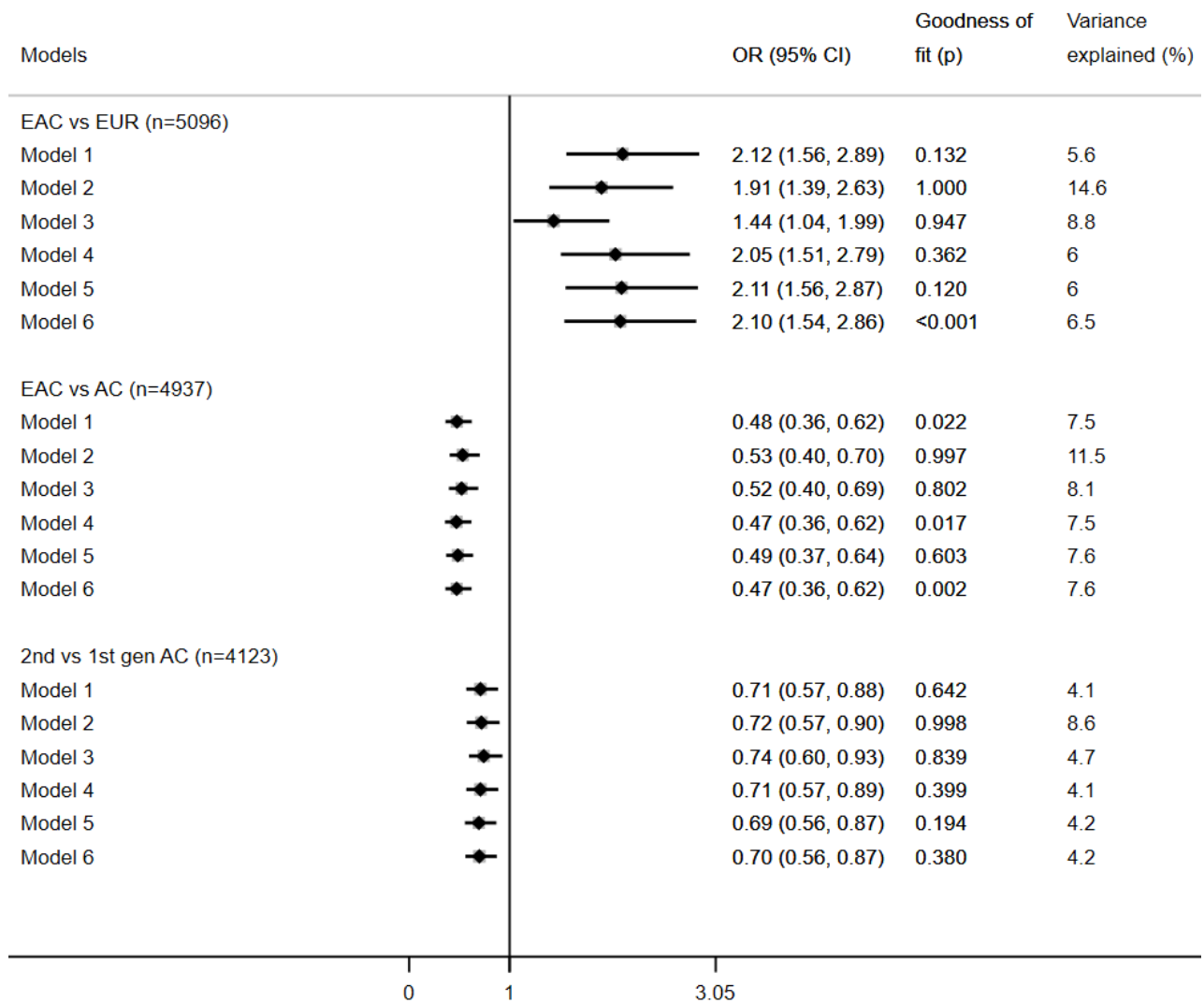
Model 1: Ethnicity/generations+ age +sex,
 Model 2: Model 1 +WHR,
 Model 3: Model 1 +deprivation,
 Model 4: Model 1 +smoking,
 Model 5: Model 1 +height,
 Model 6: Model 1 +years of education

Figure s4.2: Forest plots of multivariate regression models for diabetes in South Asians - sensitivity analysis



Model 1: Ethnicity/generations+ age +sex
 Model 2: Model 1 +WHR
 Model 3: Model 1 +deprivation
 Model 4: Model 1 +smoking
 Model 5: Model 1 +height
 Model 6: Model 1 +years of education
 Model 7: Model 1 +birth weight

Figure s4.3: Forest plots of multivariate regression models for diabetes in African Caribbeans



Model 1: Ethnicity/generations+ age +sex
 Model 2: Model 1 +WHR
 Model 3: Model 1 +deprivation
 Model 4: Model 1 +smoking
 Model 5: Model 1 +height
 Model 6: Model 1 +years of education

Figure s4.4: Forest plots of multivariate regression models for diabetes in African Caribbeans - sensitivity analysis

