Behavioural and Environmental Obesogenic Variables in the UK Biobank

We selected 12 measures that predominantly represented diet and activity based measures previously studied in gene x BMI variant publications[1-7]. In addition we selected sun protection use as a negative control. The measures are described in more detail below:

Dietary information

Participants completed a generic diet questionnaire during recruitment and a subset of 46,526 individuals completed up to five 24-hour food frequency questionnaires (FFQ). The FFQ focussed on the consumption of approximately 200 commonly consumed food and drinks (http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=118240). For each participant completing the food frequency questionnaire nutrient intakes were estimated by multiplying the quantity consumed by the nutrient composition of the food or beverage, as taken from the UK food composition database [8]. The 46,526 participants with genetic data completing at least one standard (i.e. normal diet) FFQ were included in this study.

Fizzy drink and fried food intake

Fizzy drink consumption was determined from the FFQ and represented number of glasses of fizzy drink consumed on an average day. Fried food intake was determined from the FFQ and combined the reported intake of fried chicken and fried potato.

Percentage fat and protein

Fat and protein (in grams) consumed were taken from the UK Biobank derived nutrients information in the FFQ. The variables were then divided by total energy intake (in KJ).

Western diet

The generic diet questionnaire was used to calculate the average consumption of fruit, vegetables, fish (oily and non-oily), meat (processed, poultry, beef, lamb and pork), cheese, milk, bread, cereal, tea, coffee and water. To condense this information we performed a principal component factor analysis. Seven eigenvalues were greater than 1, factor 1 was considered to represent a “Westernised” diet (this was higher in calories and processed foods), factor 2 representing a prudent diet and factor 3 representing a healthy diet. Here, the “Westernised” diet was investigated further. This information was available for 94,040 individuals of white origin with genetic data available.

Self-reported physical activity

International Physical Activity Questionnaire

The UK Biobank asked a range of questions about physical activity questions to all participants. We derived the total metabolic equivalent of task (MET) minutes of exercise per week (based on the International Physical Activity Questionnaire (IPAQ)) using the IPAQ guidelines[9].

Sedentary behaviour

The UK Biobank asked all participants about the hours per day they spent a) driving, b) using a computer and c) watching television. These three variables were summed to provide hours per day participants spent sat down. Values greater than 24 hours per day were excluded. Those reporting over 16 hours were recoded to 16 hours. Sedentary time was available for 119,688 individuals with genetic data available.

TV watching

Participants in the UK Biobank were asked to report how many hours they spent watching TV in a typical day.

Vigorous activity

Minutes spent undertaking vigorous activity each week was calculated. A dichotomous vigorous activity variable was also derived denoting participants who performed more than 1 hour of vigorous activity per week or not.

Measured physical activity with accelerometer data

Daily accelerometer data were available for 19,229 individuals of White British origin with genetic data available for a period of 6 days. A variable was derived from this data representing the mean levels of moderate physical activity per day for each individual.

Composite score of the obesogenic environment and behaviour

Physical activity (as measured by IPAQ), sedentary time, TV watching and westernised diet were available in 86,549 individuals with BMI genetic variants available. We did not use other variables as they were only available in smaller numbers. The obesogenic variables were combined using a principal components factor analysis in STATA. Only one factor had

an eigenvalue of greater than one and this was utilised as a composite score of the obesogenic environment.

Sun protection use

All participants in the UK Biobank were asked "Do you wear sun protection (e.g. sunscreen lotion, hat) when you spend time outdoors in the summer?" with the options: Never, Sometimes, Most of the time, Always, Don't go out in the sun, Don't know and Prefer not to answer. We derived a binary variable comparing those who always or usually use sun protection to those who never or occasionally use sun protection.

All the variables were dichotomised (at the median for continuous variables) to investigate the mean BMI in these groups (Supplementary table 8) and interaction was investigated in the same way as for TDI (using continuous variables where possible in the interaction model). Additionally, the correlation of each variable with TDI was determined.

Generating the simulated variable for TDI

In our analysis of gene x environment interactions for TDI the basic starting equation is as follows:

$$(\boldsymbol{y}|\boldsymbol{C},\boldsymbol{g},\boldsymbol{e}) = \boldsymbol{C}.\boldsymbol{c} + \boldsymbol{\alpha g} + \boldsymbol{\beta e} + \boldsymbol{\varepsilon} \quad (1)$$

where $\mathbf{y}$ is BMI, $\mathbf{e}$ is the environmental variable (TDI), $\mathbf{g}$ is the BMI genetic risk score and $\mathbf{C}$ represents important covariates.

Since $\mathbf{g}$ is a genetic risk score for the trait $\mathbf{y}$, $\boldsymbol{\alpha}$ is non-zero but $\boldsymbol{\beta}$ can be zero. When statistical interaction is tested, the model is changed to

$$(\boldsymbol{y}|\boldsymbol{C},\boldsymbol{g},\boldsymbol{e}) = \boldsymbol{C}.\boldsymbol{c} + \boldsymbol{\alpha g} + \boldsymbol{\beta e} + \boldsymbol{\gamma}(\boldsymbol{g} * \boldsymbol{e}) + \boldsymbol{\varepsilon} \quad (2)$$

where $\mathbf{g} * \mathbf{e}$ refers to the element-wise product of two vectors.

One of the major problems with testing the $\boldsymbol{\gamma}$ parameter in this model is that if the environmental factor $\mathbf{e}$ is correlated with $\mathbf{y}$ and $\mathbf{g}$ the test may yield spurious or biased interaction coefficients for example due to collider bias or biases well-established in secondary trait analysis.

In our simulation analysis rather than running model (2) alone we also perform the following:

$$(\boldsymbol{y}|\boldsymbol{C},\boldsymbol{g},\boldsymbol{f}) = \boldsymbol{C}.\boldsymbol{c} + \boldsymbol{\alpha g} + \boldsymbol{\beta f} + \boldsymbol{\gamma}(\boldsymbol{g} * \boldsymbol{f}) + \boldsymbol{\tau} \quad (3)$$

where $\boldsymbol{f}$ relates to $\boldsymbol{y}$, $\boldsymbol{g}$ and $\boldsymbol{C}$ marginally as does $\boldsymbol{e}$ and has the same conditional distribution. In other words we create an artificial environmental variable that behaves marginally exactly as the real environmental variable $\boldsymbol{e}$.

In practice one can simulate $\boldsymbol{f}$ easily by regressing $\boldsymbol{e}$ on $[\boldsymbol{C},\boldsymbol{g},\boldsymbol{y}]$ and add the fitted values to a random permutation of the residuals. This ensures that $\boldsymbol{f}$ and $\boldsymbol{e}$ have the same conditional expectations and same residual distributions.

Sensitivity analyses to explore additional factors that could affect gene x obesogenic environment interactions

Evidence of interaction when analysing BMI on the kgm^{-2} scale

Our primary analysis was based on forcing the outcome, BMI, into a normal distribution. We used an inverse normalised distribution because skewed distributions and different variances can inflate effect estimates due to heteroscedasticity. Previous studies have not necessarily accounted for heteroscedasticity. As expected, when analysed on the natural BMI scale (kgm^{-2}), the evidence of interaction was stronger than when using BMI values on the inverse normalised scale, but this increase is likely partly artefactual due to the increased variance in BMI in the group living in more deprived areas (Supplementary Figure 3, Supplementary Table 10).

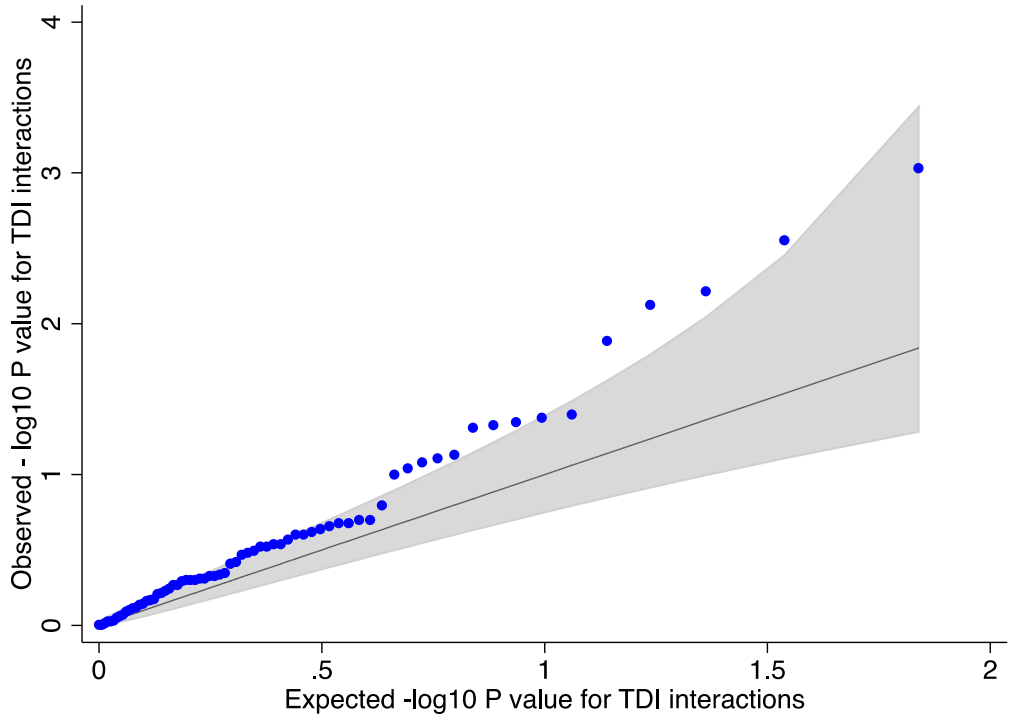
We next tested how the distribution of the environmental interaction term (here TDI) affected the evidence of interaction. Our primary results were based on Townsend deprivation index on its natural scale, which includes a slight right hand skew (Supplementary Figure 2). We therefore tested our results when TDI was inverse normalised. The evidence for interaction remained, with a larger effect of the BMI genetic risk score on BMI for individuals living in more deprived areas (0.025 standard deviations per allele [95%CI: 0.023-0.027]) compared to those in less deprived areas (0.022 [95%CI: 0.020-0.024]) (Table 2), although the statistical confidence was weaker $P_{interaction}$ $7x10^{-4}$ ($P_{interaction}$ $8x10^{-4}$, using robust standard errors).

We next tested how splitting the sample into two groups with different environmental variability can introduce spurious GxE association. Gene environment interaction studies often stratify the population using a high threshold for the environmental variable. This dichotomisation can artificially reduce the environmental variance in one of the groups and hence seemingly increases the observed genetic effect in that stratum. This problem can be reduced by splitting the sample such that the environmental variability is equal in the two groups. We largely avoided this problem

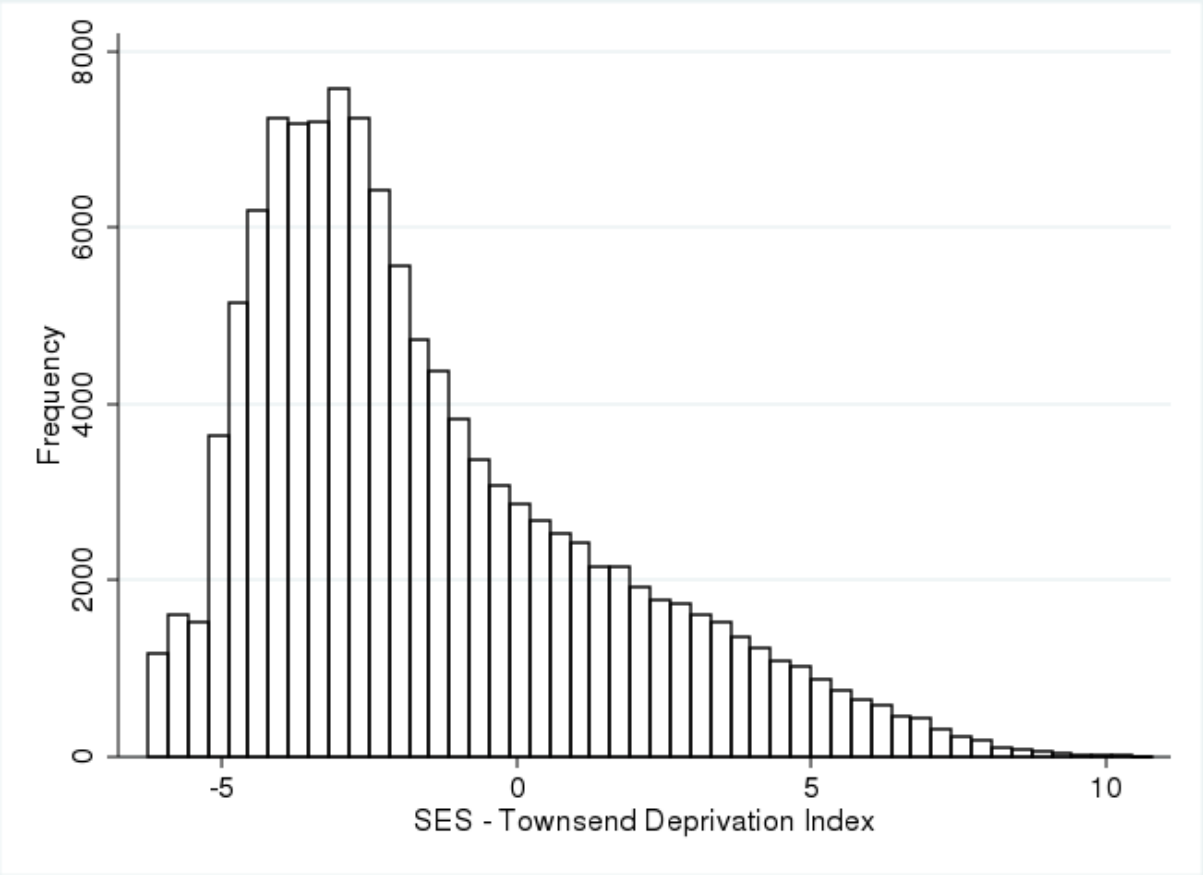
by using a continuous variable as the interaction term but we also tested BMI genetic risk score – BMI associations by dichotomising TDI at different points. We observed similar levels of evidence of interaction when splitting people into a groups based on the 75% most deprived areas and 25% least deprived areas, and vice versa and when splitting 50:50 (Supplementary Table 11).

Finally, we confirmed that the evidence for interaction was similar in both sexes (Supplementary table 12). This analysis was important because the variance in BMI is wider in women, and our previous studies[10] show that BMI is likely to causally influence TDI in women to a greater extent than men.

Supplementary Figure 1: QQ plot showing the observed TDI-BMI genetic interaction p-values from the 69 SNPs against the expected p-values

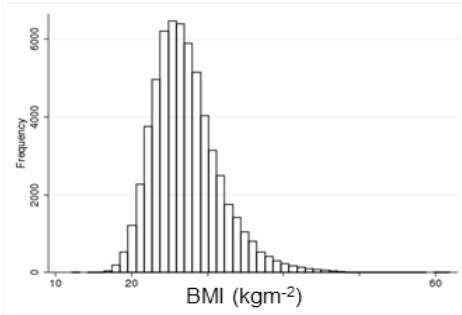


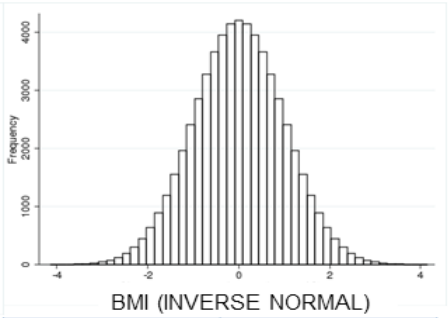
Supplementary Figure 2: Histogram showing the distribution of the Townsend Deprivation Index for the 119,464 individuals in the UK Biobank



Supplementary Figure 3: Histograms representing the distribution of BMI in high SES and low SES groups. BMI represents raw BMI but effect sizes and p values are based on BMI adjusted for age, sex, ancestry principal components, assessment centre location and genotyping chip. BMI (INVERSE NORMAL) transforms the BMI residual variable to the inverse normal scale with a mean of zero and standard deviation of 1. The BMI GRS variance refers to the variance in BMI explained by the BMI genetic risk score of 69 variants weighted by their effects on BMI.

LOW RISK OBESOGENIC ENVIRONMENT – LOW DEPRIVATION

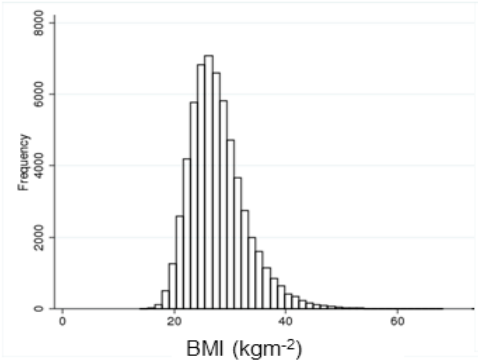


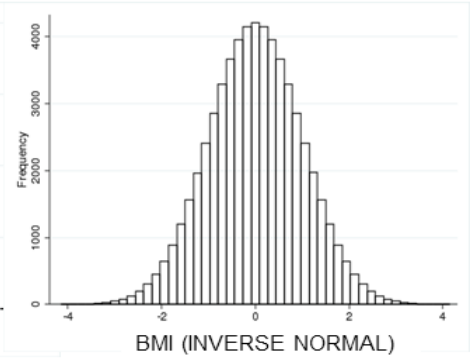


Mean BMI	**27.20**
SD BMI	4.47
BMI GRS beta	0.098
BMI GRS P	$1x10^{-176}$
BMI GRS: variance	1.56
P interaction with high deprivation	$5x10^{-17}$

Mean BMI	**0**
SD BMI	1
BMI GRS beta	0.022
BMI GRS P	$3x10^{-176}$
BMI GRS: variance	1.55
P interaction with high deprivation	$6x10^{-12}$

HIGH RISK OBESOGENIC ENVIRONMENT – HIGH DEPRIVATION





Mean BMI	**27.87**
SD BMI	5.13
BMI GRS beta	0.128
BMI GRS P	$2x10^{-228}$
BMI GRS: variance	2.02

Mean BMI	**0**
SD BMI	1
BMI GRS beta	0.025
BMI GRS P	$3x10^{-225}$
BMI GRS: variance	1.96

Supplementary table 1: Demographics of the 472,279 individuals in the UK Biobank of white origin with Townsend deprivation index available.

	Most deprived	Least deprived	P*
N	236,036	236,243	
Mean age at recruitment (SD)	56.3 (8.2)	57.3 (7.9)	$<1x10^{-15}$
Male, N (%)	107,650 (45.6)	107,408 (45.5)	0.33
Mean Townsend deprivation index (SD)	0.88 (2.52)	-3.76 (0.94)	$<1x10^{-15}$
Mean BMI (SD)	27.7 (5.1)	27.1 (4.4)	$<1x10^{-15}$
Obese, N (%)	57,643 (24.4)	47,576 (20.1)	$<1x10^{-15}$
Current smoker, N (%)	30,276 (12.8)	14,434 (6.1)	$<1x10^{-15}$
Type 2 diabetes, N (%)	8,491 (3.6)	6,032 (2.6)	$<1x10^{-15}$
Coronary artery disease, N (%)	12,370 (5.2)	8,934 (3.8)	$<1x10^{-15}$

Supplementary Table 2: Differences in BMI by BMI genetic risk score decile (kgm^2) and by allele (inverse normalised scale) for a) Townsend deprivation index split at the median and b) Townsend deprivation index split at the UK average deprivation value. Interaction p-values are calculated using the binary TDI variable for both to enable comparison.

Trait	Obesogenic category	N	BMI (SD)	BMI difference in 10% lowest genetic risk	BMI difference in 10% highest genetic risk	Per-allele Beta	SE	P association	P interaction*	P Interaction Robust**
Townsend Deprivation Index (natural scale)	High SES TDI≤-2.294	59,872	27.20 (4.47)			0.022	0.001	$<1x10^{-15}$	**$4x10^{-6}$**	**$5x10^{-6}$**
	Low SES TDI>-2.294	59,861	27.87 (5.13)	+0.35 kgm^{-2}	+0.92 kgm^{-2}	0.025	0.001	$<1x10^{-15}$		
Townsend Deprivation Index (natural scale)	High SES	84,526	27.30 (4.56)			0.022	0.001	$<1x10^{-15}$	**$9x10^{-9}$**	**$6x10^{-8}$**
	Low SES	35,357	28.11 (5.37)	+0.42 kgm^{-2}	+1.06 kgm^{-2}	0.027	0.001	$<1x10^{-15}$		

Supplementary Table 3: BMI genetic risk score association with BMI for different age groups in the UK Biobank. The interaction effect was then investigated for TDI in the three age groups and the P values for normal and robust models are presented.

Age group	N	Beta for BMI GRS against BMI	SE	P	Variance explained (%)	TDI Pinteraction	TDI Pinteraction robust
40-49	25658	0.028	0.001	$<1x10^{-15}$	2	9.00E-05	3.00E-04
50-59	40131	0.025	0.001	$<1x10^{-15}$	1.7	3.00E-05	1.00E-04
60-73	53944	0.020	0.0008	$<1x10^{-15}$	1.2	6.00E-04	1.00E-03

Supplementary table 4: Individual SNP associations with BMI in high and low Townsend deprivation index groups

SNP	Obesogenic category	Beta	SE	P association	P interaction	P interaction robust
rs1000940	Low SES	0.001	0.006	0.90	0.10	0.13
	High SES	0.020	0.006	0.002		
rs10132280	Low SES	0.020	0.006	0.002	0.30	0.33
	High SES	0.021	0.006	$8x10^{-4}$		
rs1016287	Low SES	0.007	0.006	0.30	$9x10^{-4}$	0.002
	High SES	0.031	0.006	$1x10^{-6}$		
rs10182181	Low SES	0.035	0.006	$2x10^{-9}$	0.50	0.53
	High SES	0.032	0.006	$3x10^{-8}$		
rs10733682	Low SES	0.021	0.006	$4x10^{-4}$	0.99	0.99
	High SES	0.017	0.006	0.004		
rs10938397	Low SES	0.024	0.006	$4x10^{-5}$	0.013	0.02
	High SES	0.037	0.006	$4x10^{-10}$		
rs10968576	Low SES	0.024	0.006	$1x10^{-4}$	0.49	0.51
	High SES	0.024	0.006	$1x10^{-4}$		
rs11057405	Low SES	0.031	0.009	$8x10^{-4}$	0.61	0.63
	High SES	0.030	0.009	0.001		
rs11126666	Low SES	0.001	0.007	0.93	0.89	0.90
	High SES	0.004	0.007	0.53		
rs11165643	Low SES	0.017	0.006	0.003	0.47	0.49
	High SES	0.014	0.006	0.014		
rs11191560	Low SES	0.037	0.011	$6x10^{-4}$	0.49	0.71
	High SES	0.019	0.011	0.08		
rs11583200	Low SES	0.020	0.006	0.001	0.54	0.56
	High SES	0.017	0.006	0.005		
rs1167827	Low SES	0.019	0.006	0.001	0.93	0.94
	High SES	0.019	0.006	$9x10^{-4}$		
rs11688816	Low SES	0.019	0.006	0.001	0.59	0.61
	High SES	0.009	0.006	0.13		
rs11727676	Low SES	0.001	0.010	0.95	0.50	0.52
	High SES	-0.008	0.010	0.41		
rs11847697	Low SES	0.007	0.014	0.61	0.25	0.28
	High SES	0.019	0.014	0.17		
rs12286929	Low SES	0.011	0.006	0.07	0.94	0.95
	High SES	0.010	0.006	0.10		
rs12401738	Low SES	0.008	0.006	0.15	0.46	0.49
	High SES	0.015	0.006	0.011		
rs12429545	Low SES	0.029	0.009	$9x10^{-4}$	0.27	0.30
	High SES	0.028	0.009	0.001		
rs12446632	Low SES	0.029	0.008	$5x10^{-4}$	0.45	0.48
	High SES	0.027	0.008	0.001		
rs12566985	Low SES	0.012	0.006	0.044	0.21	0.24
	High SES	0.010	0.006	0.08		
rs12885454	Low SES	0.017	0.006	0.004	0.33	0.35

	High SES	0.014	0.006	0.017		
rs12940622	Low SES	0.021	0.006	$3x10^{-4}$	0.72	0.74
	High SES	0.013	0.006	0.024		
rs13021737	Low SES	0.049	0.008	$1x10^{-10}$	0.20	0.23
	High SES	0.067	0.008	$6x10^{-18}$		
rs13078960	Low SES	0.025	0.007	$4x10^{-4}$	0.99	0.99
	High SES	0.024	0.007	0.001		
rs13191362	Low SES	0.026	0.009	0.003	0.39	0.43
	High SES	0.023	0.009	0.008		
rs1516725	Low SES	0.034	0.008	$5x10^{-5}$	0.67	0.69
	High SES	0.029	0.008	$7x10^{-4}$		
rs1528435	Low SES	0.014	0.006	0.023	0.29	0.32
	High SES	0.014	0.006	0.015		
rs1558902	Low SES	0.072	0.006	$5x10^{-34}$	0.006	0.010
	High SES	0.081	0.006	$3x10^{-43}$		
rs16851483	Low SES	0.010	0.012	0.38	0.22	0.26
	High SES	0.045	0.012	$1x10^{-4}$		
rs16951275	Low SES	0.036	0.007	$1x10^{-7}$	0.23	0.26
	High SES	0.028	0.007	$6x10^{-5}$		
rs17024393	Low SES	0.055	0.018	0.003	0.091	0.11
	High SES	0.091	0.018	$6x10^{-7}$		
rs17094222	Low SES	0.012	0.007	0.09	0.25	0.29
	High SES	0.014	0.007	0.05		
rs17405819	Low SES	0.013	0.006	0.046	0.47	0.50
	High SES	0.018	0.006	0.05		
rs17724992	Low SES	0.017	0.007	0.011	0.68	0.70
	High SES	0.027	0.007	$5x10^{-5}$		
rs1808579	Low SES	0.019	0.006	$9x10^{-4}$	0.21	0.24
	High SES	0.023	0.006	$9x10^{-5}$		
rs1928295	Low SES	0.001	0.006	0.83	0.24	0.27
	High SES	0.018	0.006	0.002		
rs2033732	Low SES	0.011	0.007	0.11	0.07	0.09
	High SES	-0.006	0.007	0.33		
rs205262	Low SES	0.031	0.007	$3x10^{-6}$	0.76	0.77
	High SES	0.025	0.007	$1x10^{-4}$		
rs2112347	Low SES	0.029	0.006	$2x10^{-6}$	0.54	0.57
	High SES	0.024	0.006	$7x10^{-5}$		
rs2121279	Low SES	0.009	0.009	0.33	0.50	0.53
	High SES	0.004	0.009	0.67		
rs2176598	Low SES	0.016	0.007	0.018	0.30	0.33
	High SES	0.028	0.007	$2x10^{-5}$		
rs2207139	Low SES	0.037	0.008	$2x10^{-6}$	0.38	0.41
	High SES	0.042	0.008	$5x10^{-8}$		
rs2245368	Low SES	0.028	0.008	$3x10^{-4}$	0.57	0.60
	High SES	0.016	0.008	0.046		
rs2287019	Low SES	0.028	0.008	$2x10^{-4}$	0.29	0.32
	High SES	0.041	0.008	$5x10^{-8}$		
rs2365389	Low SES	0.035	0.006	$2x10^{-9}$	0.94	0.95

	High SES	0.024	0.006	$4x10^{-5}$		
rs2650492	Low SES	0.020	0.006	0.002	0.97	0.97
	High SES	0.019	0.006	0.004		
rs2820292	Low SES	0.021	0.006	$4x10^{-4}$	0.73	0.75
	High SES	0.019	0.006	0.001		
rs29941	Low SES	0.007	0.006	0.22	0.049	0.06
	High SES	0.026	0.006	$2x10^{-5}$		
rs3101336	Low SES	0.023	0.006	$1x10^{-4}$	0.045	0.06
	High SES	0.031	0.006	$1x10^{-7}$		
rs3736485	Low SES	0.014	0.006	0.014	0.08	0.10
	High SES	0.008	0.006	0.18		
rs3810291	Low SES	0.026	0.006	$3x10^{-5}$	0.047	0.06
	High SES	0.029	0.006	$2x10^{-6}$		
rs3817334	Low SES	0.027	0.006	$3x10^{-6}$	0.040	0.06
	High SES	0.034	0.006	$8x10^{-9}$		
rs3849570	Low SES	0.004	0.006	0.50	0.32	0.35
	High SES	0.016	0.006	0.009		
rs4256980	Low SES	0.017	0.006	0.006	0.20	0.23
	High SES	0.024	0.006	$6x10^{-5}$		
rs4740619	Low SES	0.016	0.006	0.007	0.81	0.82
	High SES	0.016	0.006	0.007		
rs543874	Low SES	0.040	0.007	$2x10^{-8}$	0.008	0.013
	High SES	0.056	0.007	$4x10^{-15}$		
rs6477694	Low SES	0.000	0.006	0.99	0.042	0.06
	High SES	0.014	0.006	0.023		
rs6567160	Low SES	0.046	0.007	$2x10^{-11}$	0.003	0.005
	High SES	0.060	0.007	$1x10^{-18}$		
rs657452	Low SES	0.017	0.006	0.005	0.79	0.80
	High SES	0.012	0.006	0.048		
rs6804842	Low SES	0.008	0.006	0.16	0.77	0.78
	High SES	0.010	0.006	0.08		
rs7138803	Low SES	0.037	0.006	$5x10^{-10}$	0.85	0.86
	High SES	0.030	0.006	$4x10^{-7}$		
rs7141420	Low SES	0.023	0.006	$9x10^{-5}$	0.34	0.37
	High SES	0.014	0.006	0.019		
rs7243357	Low SES	0.023	0.008	0.002	0.16	0.19
	High SES	0.004	0.008	0.59		
rs758747	Low SES	0.007	0.007	0.27	0.62	0.64
	High SES	0.021	0.007	0.001		
rs7599312	Low SES	0.015	0.007	0.025	0.49	0.52
	High SES	0.024	0.007	$3x10^{-4}$		
rs7899106	Low SES	0.033	0.013	0.014	0.87	0.88
	High SES	0.015	0.013	0.27		
rs9400239	Low SES	0.015	0.006	0.020	0.51	0.53
	High SES	0.019	0.006	0.003		
rs9581854	Low SES	0.012	0.007	0.10	0.08	0.10
	High SES	0.016	0.008	0.029		

Supplementary table 5: Differences in BMI by allele (inverse normalised scale) for TDI in the CoLaus Study, occupational status in the 1958 Birth Cohort and the UK Biobank and educational years in the UK Biobank

Study	Obesogenic category	N	BMI (SD)	Per-allele Beta	SE	P association	P interaction*	P Interaction Robust**
CoLaus	High SES based on TDI	2,623	25.53 (4.33)	0.030	0.004	$6x10^{-15}$	0.35	0.34
	Low SES based on TDI	2,614	26.18 (4.78)	0.022	0.004	$1x10^{-8}$		
UK Biobank	High job class	38,942	27.15 (4.57)	0.025	0.001	$<1x10^{-15}$	0.78	0.79
	Low job class	37,374	27.68 (4.89)	0.024	0.001	$<1x10^{-15}$		
1958 Birth Cohort	High job class	2,873	27.17 (4.55)	0.026	0.003	$2x10^{-14}$	0.62	0.62
	Low job class	3,298	27.55 (5.10)	0.024	0.003	$1x10^{-12}$		
UK Biobank	High educational years (19-20)	55,203	27.15 (4.67)	0.024	0.001	$<1x10^{-15}$	0.76	0.76
	Low educational years (<=15)	63,572	27.86 (4.93)	0.023	0.001	$<1x10^{-15}$		

BMI adjusted for age, sex, ancestral principal components and assessment centre location. Models additionally adjusted for genotyping platform

* Interaction p-value

** Interaction p-value accounting for heteroscedasticity using robust standard errors

Supplementary Table 6: Interaction p-values for the 10 self-reported obesogenic variables, measured physical activity and sun protection use.

Trait	Obesogenic category	P interaction*	Adjusted P interaction **
Fizzy drink	None daily ≥1 glass daily	0.86	0.95
Fried food consumption	None daily ≥1 meal daily	0.94	0.98
Percentage fat^	Low risk High risk	0.59	0.63
Percentage protein^	Low risk High risk	0.79	0.98
Western diet^	Low risk High risk	0.07	**0.032**
IPAQ	>1845 MET minutes per week ≤1845 MET minutes per week	**5E-6**	**3E-5**
Measured physical activity^	High activity Low activity	0.11	0.15
Sedentary time	<5 hours daily ≥5 hours daily	**0.030**	0.08
TV watching	<4 hours daily ≥4 hours daily	**7E-5**	**2E-5**
Vigorous activity	>1 hour weekly ≤1 hour weekly	0.10	0.16
Composite score^	Low risk High risk	**2E-4**	**6E-4**
Sun protection use	Usually or always use Never or sometimes use	**1E-4**	**3E-4**

*Robust standard errors utilised to calculate the interaction p-value

**Model includes adjustment for the TDI interaction and robust standard errors

Supplementary table 7: Association of Townsend deprivation index with a range of obesogenic variables. Negative values represent less deprivation.

Obesogenic environment variable	**Beta (95% CI) representing SD change in Townsend deprivation index per unit change in the obesogenic environment variable**	**P**
Dietary factors		
Fat in diet	0.16 (0.11, 0.21)	$2x10^{-11}$
Fizzy drinks	0.036 (0.019, 0.054)	$7x10^{-5}$
Fried food	0.021 (0.006, 0.036)	0.007
Protein in diet	-0.15 (-0.19, -0.11)	$1x10^{-13}$
Westernised diet*	-0.033 (-0.040, -0.027)	$<1x10^{-15}$
Activity measures		
Measured activity*	-0.065 (-0.078, -0.051)	$<1x10^{-15}$
Physical activity (IPAQ)*	0.006 (0.000, 0.012)	0.043
Sedentary time per day (hours)*	0.028 (0.022, 0.034)	$<1x10^{-15}$
TV per day (hours)*	0.119 (0.113, 0.125)	$<1x10^{-15}$
Less than one hour vigorous activity per week	0.077 (0.064, 0.089)	$<1x10^{-15}$
Other factors		
Composite score*	0.020 (0.014, 0.026)	$6x10^{-11}$
More frequent sun protection use	-0.080 (-0.087, -0.074)	$<1x10^{-15}$

*Continuous obesogenic variables single inverse normalised

Supplementary Table 8: Comparison of the high and low risk categories for a range of self-reported obesogenic environmental/behavioural measures, measured physical activity, sun protection use and the composite score.

Environmental factor	Obesogenic category	N	Male, N (%)	Mean BMI	SD BMI	Effect size (95%CI) representing change in BMI (kg/m^2) for people in high risk group compared to the low risk group^	P
Fizzy drink	None daily	39,975	18,327 (45.9)	26.93	4.62	Reference	
	≥1 glass daily	6,393	3,537 (55.3)	27.69	4.91	0.71 (0.58, 0.83)	<1E-15
Fried food intake	None daily	31,821	14,485 (45.5)	26.96	4.66	Reference	
	≥1 meal daily	14,547	7,379 (50.7)	27.20	4.68	0.20 (0.10, 0.29)	0.00002
Percentage fat*	Low risk	23,194	11,080 (47.8)	26.91	4.46	Reference	
	High risk	23,174	10,784 (46.5)	27.16	4.86	0.28 (0.19, 0.36)	1E-10
Percentage protein*	Low risk	23,188	12,137 (52.3)	26.70	4.54	Reference	
	High risk	23,180	9,727 (42.0)	27.37	4.77	0.77 (0.68, 0.85)	<1E-15
Western diet*	Low risk	47,027	19,783 (42.1)	27.06	4.71	Reference	
	High risk	47,013	24,853 (52.9)	28.00	4.79	0.86 (0.80, 0.92)	<1E-15
IPAQ	>1845 MET minutes per week	54,573	27,217 (49.9)	26.86	4.31	Reference	
	≤1845 MET minutes per week	54,569	25,111 (46.0)	27.93	4.99	1.11 (1.06, 1.17)	<1E-15
Measured physical activity*	High activity	9,636	4,038 (41.9)	25.79	3.92	Reference	<1E-15
	Low activity	9,636	4,777 (49.6)	27.79	4.92	1.95 (1.83, 2.09)	
Sedentary time	<5 hours daily	63,343	25,281 (39.9)	26.61	4.47	Reference	
	≥5 hours daily	56,345	31,387 (55.7)	28.56	4.99	1.84 (1.78, 1.89)	<1E-15
TV	<4 hours daily	82,022	38,866 (47.4)	26.98	4.54	Reference	
	≥4 hours daily	36,814	17,496 (47.5)	28.70	5.16	1.69 (1.63, 1.75)	<1E-15
Vigorous activity	>1 hour weekly	35,183	18,637 (53.0)	26.80	4.24	Reference	
	≤1 hour weekly	74,004	33,710 (45.6)	27.68	4.87	0.92 (0.86, 0.98)	<1E-15
Sun protection use	Usually or always use	68,507	25,641 (37.4)	27.32	4.75	Reference	<1E-15
	Never or sometimes use	50,561	30,743 (60.8)	27.81	4.89	0.31 (0.25, 0.37)	

Composite score*	Low risk	43,275	19,768 (45.7)	26.33	4.13	Reference	
	High risk	43,274	21,933 (50.7)	28.46	4.87	2.08 (2.02, 2.14)	<1E-15

^ Adjusted for age, sex and ancestry principal components; * high and low risk taken from median values

Supplementary table 9: Summary of the body mass index (BMI) variants previously identified as associated with those traits at genome wide significance

Trait	Genetic variant	Locus	Exclude from score	Reason for exclusion	Trait raising allele	Trait lowering allele	Directly genotyped or Imputed	Imputation quality	Beta representing SD change in BMI or height for each SNP in UK Biobank data	P value
BMI	rs1000940	*RABEP1*	No	NA	G	A	Imputed	0.99624	0.011 (0.004)	1.60E-02
BMI	rs10132280	*STXBP6*	No	NA	C	A	Imputed	0.97496	0.020 (0.005)	1.10E-05
BMI	rs1016287	*FLJ30838*	No	NA	T	C	Imputed	0.99411	0.019 (0.004)	2.00E-05
BMI	rs10182181	*ADCY3*	No	NA	G	A	Imputed	0.99521	0.033 (0.004)	1.40E-15
BMI	rs10733682	*LMX1B*	No	NA	A	G	Imputed	0.9576	0.019 (0.004)	5.90E-06
BMI	rs10938397	*GNPDA2*	No	NA	G	A	Imputed	1	0.030 (0.004)	5.80E-13
BMI	rs10968576	*LINGO2*	No	NA	G	A	Imputed	1	0.024 (0.004)	6.90E-08
BMI	rs11030104	*BDNF*	Yes	BMI-raising allele also associated with regular smoking (which itself has a causal effect on BMI in opposite direction)	A	G	Imputed	0.99931	NA	NA

BMI	rs11057405	*CLIP1*	No	NA	G	A	Imputed	1	0.030 (0.007)	4.70E-06
BMI	rs11126666	*KCNK3*	No	NA	A	G	Imputed	0.99485	0.002 (0.005)	7.10E-01
BMI	rs11165643	*PTBP2*	No	NA	T	C	Imputed	0.99575	0.016 (0.004)	9.50E-05
BMI	rs11191560	*NT5C2*	No	NA	C	T	Imputed	0.99989	0.026 (0.008)	6.50E-04
BMI	rs11583200	*ELAVL4*	No	NA	C	T	Imputed	0.98728	0.019 (0.004)	7.70E-06
BMI	rs1167827	*HIP1*	No	NA	G	A	Imputed	1	0.020 (0.004)	1.80E-06
BMI	rs11688816	*EHBP1*	No	NA	G	A	Imputed	0.98096	0.014 (0.004)	9.40E-04
BMI	rs11727676	*HHIP*	No	NA	T	C	Imputed	1	-0.003 (0.007)	6.60E-01
BMI	rs11847697	*PRKD1*	No	NA	T	C	Imputed	1	0.014 (0.010)	1.70E-01
BMI	rs12286929	*CADM1*	No	NA	G	A	Imputed	0.99124	0.010 (0.004)	1.20E-02
BMI	rs12401738	*FUBP1*	No	NA	A	G	Imputed	0.99528	0.012 (0.004)	3.30E-03
BMI	rs12429545	*OLFM4*	No	NA	A	G	Imputed	0.97759	0.027 (0.006)	8.00E-06
BMI	rs12446632	*GPRC5B*	No	NA	G	A	Imputed	0.99978	0.028 (0.006)	2.40E-06

BMI	rs12566985	*FPGT-TNNI3K*	No	NA	G	A	Imputed	0.9947	0.011 (0.004)	6.10E-03
BMI	rs12885454	*PRKD1*	No	NA	C	A	Imputed	0.99569	0.015 (0.004)	4.60E-04
BMI	rs12940622	*RPTOR*	No	NA	G	A	Imputed	0.99796	0.017 (0.004)	5.90E-05
BMI	rs13021737	*TMEM18*	No	NA	G	A	Imputed	0.99072	0.059 (0.005)	9.10E-27
BMI	rs13078960	*CADM2*	No	NA	G	T	Imputed	0.9915	0.024 (0.005)	2.50E-06
BMI	rs13107325	*SLC39A8*	Yes	Missense Ala/Thr polymorphism located in exon 7 of SLC39A8, which encodes a zinc transporter that also transports cadmium and manganese. It is also associated with BP and HDL levels, and presumably these and the BMI effect are secondary to the metal ion transport variation.	T	C	Imputed	1	NA	NA
BMI	rs13191362	*PARK2*	No	NA	A	G	Imputed	0.98973	0.026 (0.006)	3.10E-05
BMI	rs1516725	*ETV5*	No	NA	C	T	Imputed	0.99495	0.032 (0.006)	1.00E-07
BMI	rs1528435	*UBE2E3*	No	NA	T	C	Imputed	0.99738	0.014 (0.004)	6.60E-04

BMI	rs1558902	*FTO*	No	NA	A	T	Imputed	0.99914	0.077 (0.004)	1.50E-75
BMI	rs16851483	*RASA2*	No	NA	T	G	Imputed	0.99906	0.028 (0.008)	6.80E-04
BMI	rs16951275	*MAP2K5*	No	NA	T	C	Imputed	0.99819	0.032 (0.005)	4.40E-11
BMI	rs17001654	*SCARB2*	Yes	SNP not in Hardy-Weinberg equilibrium	G	C	Imputed	0.9483	NA	NA
BMI	rs17024393	*GNAT2*	No	NA	C	T	Imputed	0.98934	0.074 (0.013)	1.20E-08
BMI	rs17094222	*HIF1AN*	No	NA	C	T	Imputed	0.96874	0.013 (0.005)	8.50E-03
BMI	rs17405819	*HNF4G*	No	NA	T	C	Imputed	0.99793	0.014 (0.004)	1.30E-03
BMI	rs17724992	*PGPEP1*	No	NA	A	G	Imputed	0.98342	0.023 (0.005)	1.10E-06
BMI	rs1808579	*C18orf8*	No	NA	C	T	Imputed	0.99797	0.022 (0.004)	1.50E-07
BMI	rs1928295	*TLR4*	No	NA	T	C	Imputed	0.99998	0.010 (0.004)	1.60E-02
BMI	rs2033529	*TDRG1*	Yes	SNP not available	G	A	NA	NA	NA	NA
BMI	rs2033732	*RALYL*	No	NA	C	T	Imputed	1	0.002 (0.005)	6.70E-01
BMI	rs205262	*C6orf106*	No	NA	G	A	Imputed	0.9968	0.028 (0.005)	1.10E-09
BMI	rs2075650	*TOMM40*	Yes	SNP not in Hardy-Weinberg equilibrium	A	G	Imputed	0.9865	NA	NA

BMI	rs2112347	*POC5*	No	NA	T	G	Imputed	1	0.026 (0.004)	6.30E-10
BMI	rs2121279	*LRP1B*	No	NA	T	C	Imputed	0.98723	0.006 (0.006)	3.70E-01
BMI	rs2176598	*HSD17B12*	No	NA	T	C	Imputed	1	0.023 (0.005)	1.30E-06
BMI	rs2207139	*TFAP2B*	No	NA	G	A	Imputed	0.9989	0.038 (0.005)	1.80E-12
BMI	rs2245368	*PMS2L11*	No	NA	C	T	Imputed	1	0.022 (0.005)	8.00E-05
BMI	rs2287019	*QPCTL*	No	NA	C	T	Imputed	0.97852	0.035 (0.005)	1.00E-10
BMI	rs2365389	*FHIT*	No	NA	C	T	Imputed	0.99305	0.029 (0.004)	2.70E-12
BMI	rs2650492	*SBK1*	No	NA	A	G	Imputed	0.98144	0.019 (0.005)	3.60E-05
BMI	rs2820292	*NAV1*	No	NA	C	A	Imputed	1	0.019 (0.004)	3.60E-06
BMI	rs29941	*KCTD15*	No	NA	G	A	Imputed	1	0.018 (0.004)	5.00E-05
BMI	rs3101336	*NEGR1*	No	NA	C	T	Imputed	1	0.027 (0.004)	9.50E-11
BMI	rs3736485	*DMXL2*	No	NA	A	G	Imputed	0.98728	0.011 (0.004)	6.40E-03
BMI	rs3810291	*ZC3H4*	No	NA	A	G	Imputed	1	0.028 (0.004)	1.80E-10

BMI	rs3817334	*MTCH2*	No	NA	T	C	Imputed	1	0.031 (0.004)	1.40E-13
BMI	rs3849570	*GBE1*	No	NA	A	C	Imputed	0.99509	0.011 (0.004)	7.80E-03
BMI	rs3888190	*ATP2A1*	Yes	Associated with lots of other traits and is a big haplotype	A	C	Imputed	0.99808	NA	NA
BMI	rs4256980	*TRIM66*	No	NA	G	C	Imputed	0.99283	0.021 (0.004)	1.70E-06
BMI	rs4740619	*C9orf93*	No	NA	T	C	Imputed	0.99762	0.017 (0.004)	5.70E-05
BMI	rs543874	*SEC16B*	No	NA	G	A	Imputed	1	0.049 (0.005)	3.40E-22
BMI	rs6477694	*EPB41L4B*	No	NA	C	T	Imputed	0.99022	0.008 (0.004)	6.70E-02
BMI	rs6567160	*MC4R*	No	NA	C	T	Imputed	0.99663	0.054 (0.005)	9.50E-29
BMI	rs657452	*AGBL4*	No	NA	A	G	Imputed	0.98709	0.014 (0.004)	8.40E-04
BMI	rs6804842	*RARB*	No	NA	G	A	Imputed	0.98778	0.009 (0.004)	3.20E-02
BMI	rs7138803	*BCDIN3D*	No	NA	A	G	Imputed	1	0.034 (0.004)	1.30E-15
BMI	rs7141420	*NRXN3*	No	NA	T	C	Imputed	0.98379	0.019 (0.004)	6.70E-06
BMI	rs7243357	*GRP*	No	NA	T	G	Imputed	0.98998	0.012 (0.005)	2.10E-02

BMI	rs758747	*NLRC3*	No	NA	T	C	Imputed	0.97187	0.014 (0.005)	2.00E-03
BMI	rs7599312	*ERBB4*	No	NA	G	A	Imputed	0.97294	0.019 (0.005)	3.60E-05
BMI	rs7899106	*GRID1*	No	NA	G	A	Imputed	0.98612	0.023 (0.009)	1.40E-02
BMI	rs9400239	*FOXO3*	No	NA	C	T	Imputed	0.99206	0.017 (0.005)	2.30E-04
BMI	rs9581854	*MTIF3*	No	NA	T	C	Imputed	0.98643	0.015 (0.005)	6.20E-03
BMI	rs9925964	*KAT8*	Yes	SNP not in Hardy-Weinberg equilibrium	A	G	Imputed	1	NA	NA

Supplementary Table 10: Change in BMI per allele increase in the BMI genetic risk score when BMI analysed on its natural (kgm^{-2}) scale

Trait	Obesogenic category	N	Beta	SE	P association	P interaction*	P Interaction Robust**
Townsend Derivation Index	High SES TDI≤-2.295	59,872	0.097	0.003	$4x10^{-176}$	**$5x10^{-17}$**	**$7x10^{-14}$**
	Low SES TDI>-2.295	59,861	0.128	0.004	$7x10^{-229}$		

BMI adjusted for age, sex, ancestral principal components and assessment centre location. Models additionally adjusted for genotyping platform

* Interaction p-value

** Interaction p-value accounting for heteroscedasticity using robust standard errors

Supplementary table 11: Change in BMI (single inverse normal scale) per allele increase in the BMI GRS when Townsend deprivation index was dichotomised at approximately 25% low risk, 75% high risk or 75% low risk and 25% high risk.

Townsend Group	Obesogenic category	N	Beta	SE	P association	P interaction*
50% versus 50%	Low risk	59,928	0.022	0.001	$3x10^{-176}$	$6x10^{-6}$
	High risk	59,805	0.025	0.001	$3x10^{-225}$	
25% low risk versus 75% high risk	Low risk	29,946	0.022	0.001	$2x10^{-86}$	$4x10^{-4}$
	High risk	89,787	0.024	0.001	$<1x10^{-15}$	
75% low risk versus 25% high risk	Low risk	89,804	0.022	0.001	$5x10^{-268}$	$1x10^{-8}$
	High risk	29,929	0.027	0.001	$7x10^{-133}$	

* Interaction P-value calculated using the BMI GRS * dichotomous variable. Presented p-values were calculated with robust standard errors

Supplementary table 12: Change in BMI (single inverse normal scale) per allele increase in the BMI GRS when Townsend deprivation index was dichotomised at the median in males and females separately.

Townsend Group	Obesogenic category	N	Beta	SE	P association	P interaction*
Males only	Low risk	28,358	0.023	0.001	$7x10^{-91}$	$3x10^{-5}$
	High risk	28,331	0.025	0.001	$2x10^{-110}$	
Females only	Low risk	31,531	0.021	0.001	$2x10^{-87}$	$2x10^{-6}$
	High risk	31,513	0.025	0.001	$1x10^{-118}$	

References

1. Ahmad S, Rukh G, Varga TV, et al. Gene x physical activity interactions in obesity: combined analysis of 111,421 individuals of European ancestry. *PLoS Genet.* 2013;9(7):e1003607.
2. Kilpelainen TO, Qi L, Brage S, et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med.* Nov 2011;8(11):e1001116.
3. Li S, Zhao JH, Luan J, et al. Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. *PLoS Med.* 2010;7(8).
4. Qi Q, Chu AY, Kang JH, et al. Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies. *BMJ.* 2014;348:g1610.
5. Qi Q, Chu AY, Kang JH, et al. Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med.* Oct 11 2012;367(15):1387-1396.
6. Qi Q, Downer MK, Kilpelainen TO, et al. Dietary Intake, FTO Genetic Variants, and Adiposity: A Combined Analysis of Over 16,000 Children and Adolescents. *Diabetes.* Jul 2015;64(7):2467-2476.
7. Qi Q, Li Y, Chomistek AK, et al. Television watching, leisure time physical activity, and the genetic predisposition in relation to body mass index in women and men. *Circulation.* Oct 9 2012;126(15):1821-1827.
8. Agency FS. *McCance and Widdowson's The Composition of Foods, Sixth summary edition. .* Cambridge: Royal Society of Chemistry; 2002.
9. IPAQ. *Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ).* 2005.
10. Tyrrell J, Jones SE, Beaumont R, et al. Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *BMJ.* Mar 8 2016;352.