

Supplementary Table 1a. Description of the phenotypes used in the study.

	Trait	Abv.	Description	Units	Log Transformed?	References
anthropometric	Height	Height	Body Height	cm	No	
	Weight	Weight	Body Weight	kg	Yes	
	Bio-impedance Analysis Fat	BIA Fat	% of Fat (Tanita scales)	%	No	
	Waist	Waist	Waist Circumference	cm	Yes	
	Hips	Hips	Hip Circumference	cm	Yes	
	Waist to Hip Ratio	WHR	Waist/Hips	(None)	Yes	
	Body Mass Index	BMI	Weight/Height ²	kg*cm ⁻²	Yes	
	A Body Shape Index	ABSI	Waist/(BMI ^{2/3} *Height ^{1/2})	m ^{11/6} *kg ^{-2/3}	Yes	Krakauer & Krakauer (2012). Plos One 7, e39504.
metabolic	Creatinine	Creat	Creatinine level in serum	μmol*l ⁻¹	Yes	
	Total Cholesterol	TC	Total Cholesterol level in serum	μmol*l ⁻¹	Yes	
	HDL	HDL	HDL Cholesterol level in serum	μmol*l ⁻¹	Yes	

Supplementary Table 1b. Correlation between all the phenotypes used in the study.

[illegible]

Supplementary Table 2a. Description of the covariates used in the study.

Covariate group	Covariate	Abv.	Type	Description
base	Sex	Sex	Discrete	Sex of the individual (M: male, F: female)
	Age	Age	Continuous	Age of the individual at clinic appointment
	Clinic	Clinic	Discrete	Clinic where the measurements were taken (9 values)
socioeconomic	SIMD	SIMD	Continuous	Scottish index of multiple deprivation: A deprivation ranking based on living area and calculated using seven domains: income, employment, crime, education, health, housing and geographical access
	Years of education	YE	Continuous	Number of years attending to school/study fulltime
	Household size	HHS	Continuous	Number of people living in the current household
	VehicleRatio	VR	Continuous	Number of vehicles in the household / Household size
	Job status	Job	Discrete	Jobless status (1: currently unemployed, 0: currently working)
lifestyle	Alcohol units	Alcohol	Continuous	Units of alcohol taken in a week
	Smoking status	Smok	Discrete	Smoking history (1: currently smoking, 0: currently non-smoker)
	Activity level	Activ	Discrete	Level of activity (1: active, 0: non-active)
	FruitDayUnits	FruitU	Continuous	Number of pieces of fruit eat per day
	Fruit Consumption	FruitC	Discrete	Consumption of fruit (1: usual, 0: non-usual)
	Vegetables Consumption	VegC	Discrete	Consumption of vegetables (1: usual, 0: non-usual)
	Fish Consumption	FishC	Discrete	Consumption of fish (1: usual, 0: non-usual)
	Poultry Consumption	PoultC	Discrete	Consumption of poultry (1: usual, 0: non-usual)
	Meat Consumption	MeatC	Discrete	Consumption of meat (1: usual, 0: non-usual)
	Eggs Consumption	EggC	Discrete	Consumption of eggs (1: usual, 0: non-usual)
	Dairy Consumption	DairyC	Discrete	Consumption of dairy (1: usual, 0: non-usual)
genetic	Geographical Principal Components	gPC	Continuous	Variables capturing the genetic stratification in the sample

Supplementary Table 2b. Correlations between socioeconomic and lifestyle covariates used in the study.

[illegible]

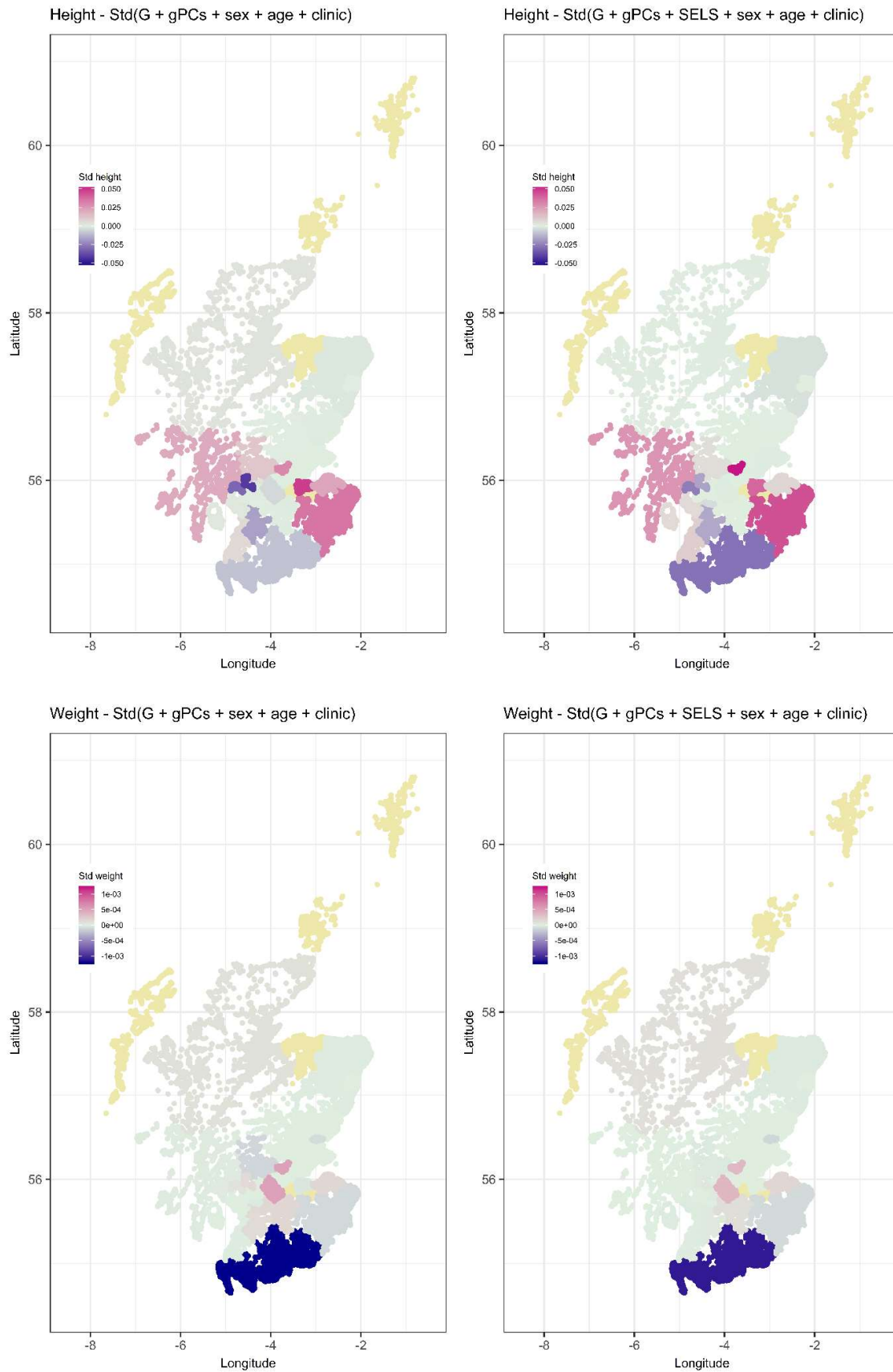
Supplementary Table 4. Significance of region on phenotypes in the full framework. The values show the significance (p-values) of region in four models: 1. Differences in the traits (first column, Basal Model); 2. Correcting for kinship (second column, Family Model); 3. Correcting for kinship and genetic structure (third column, Structure Model); 4. Correcting for kinship and environmental covariates (fourth column, Environment Model); 5. Correcting for kinship and genetic and environmental covariates (fifth column, Structure and Environment Model). The asterisk marks the estimates that are significantly different from zero. All models corrected for sex, age and clinic.

Trait	Model				
	Basal	Family	Structure	Environment	Structure and Environment
Height	3.50E-06*	0.299	0.492	0.928	0.956
Weight	0.129	0.519	0.522	0.840	0.836
BIA Fat	0.002*	0.165	0.233	0.527	0.527
Waist	1.10E-05*	0.022*	0.042*	0.232	0.251
Hips	0.070	0.328	0.342	0.600	0.595
WHR	4.52E-07*	0.002*	0.009*	0.187	0.247
BMI	9.46E-06*	0.023*	0.040*	0.478	0.501
ABSI	2.70E-04*	0.006*	0.011*	0.033*	0.034*
Creatinine	1.81E-11*	1.15E-04*	1.51E-04*	1.62E-04*	2.03E-04*
TC	0.250	0.428	0.422	0.508	0.502
HDL	1.98E-05*	0.013*	0.016*	0.201	0.207

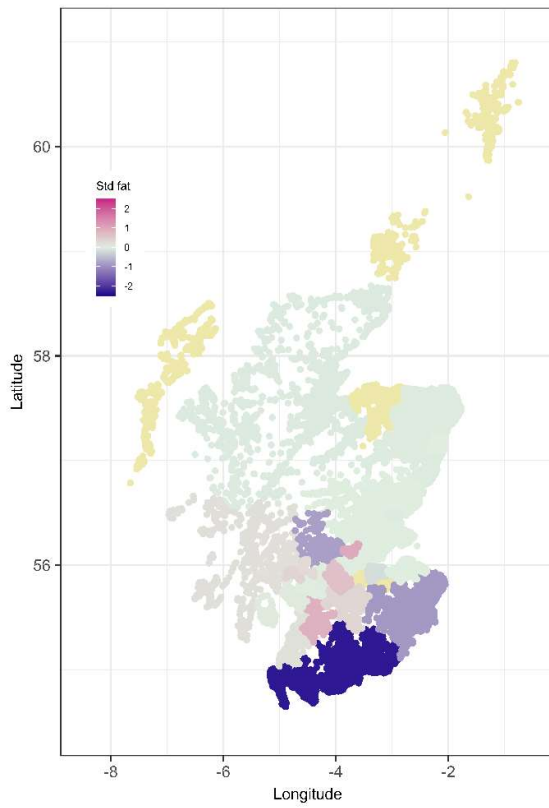
Supplementary Table 5. Variance explained by the fitted covariates in the model.

	Height	Weight	BIA Fat	Waist	Hips	WHR	BMI	ABSI	Creatinine	TC	HDL
SIMD	0.00120	0.00550	0.00410	0.00980	0.00670	0.00500	0.01280	0.00130		0.00080	0.00350
YE	0.00190	0.00110	0.00280	0.00260	0.00260	0.00130	0.00460				0.00140
Household Size	0.00100	0.00140		0.00080		0.00060	0.00050				
Vehicle Ratio	0.00040	0.00100					0.00060	0.00170		0.00320	
Jobless				0.00010	0.00010				0.00010		
Alcohol Units Std	0.00060	0.00040				0.00070			0.00230	0.00740	0.04760
Smoking Status									0.00130		
Active Status		0.11010	0.05770				0.11550	0.09370	0.16930		0.18880
Fruit Day Units	0.00030		0.00030					0.00120	0.00110	0.00200	0.00070
FruitC					0.02060			0.01270			
VeggiC	0.00090		0.00280			0.00120	0.00400	0.02620	0.01360		0.00640
FishC		0.00003	0.00010		0.00010	0.00010	0.00005	0.00010			0.00004
PoultC			0.00450	0.00590		0.01470		0.04640			0.01400
MeatC		0.00310	0.01240			0.00140	0.00710	0.03790	0.00280		0.00420
EggsC		0.01080	0.02400				0.01340	0.13210		0.01840	0.02070
DairyC			0.00020	0.00020	0.00040			0.00220	0.00020		0.00060
Total Environment	0.00640	0.13340	0.10890	0.01950	0.03040	0.02490	0.15850	0.35570	0.19060	0.03180	0.28800
gPC1	0.00320		0.00180	0.00150		0.00330	0.00150	0.00180			
gPC2	0.00030										
gPC3	0.00120								0.00040		
gPC4											
gPC5				0.00070		0.00040		0.00060			0.00120
gPC6											
gPC7											
gPC8											
gPC9											
gPC10											
Total gPCs	0.00480	0.00000	0.00180	0.00220	0.00000	0.00370	0.00150	0.00230	0.00040	0.00000	0.00120

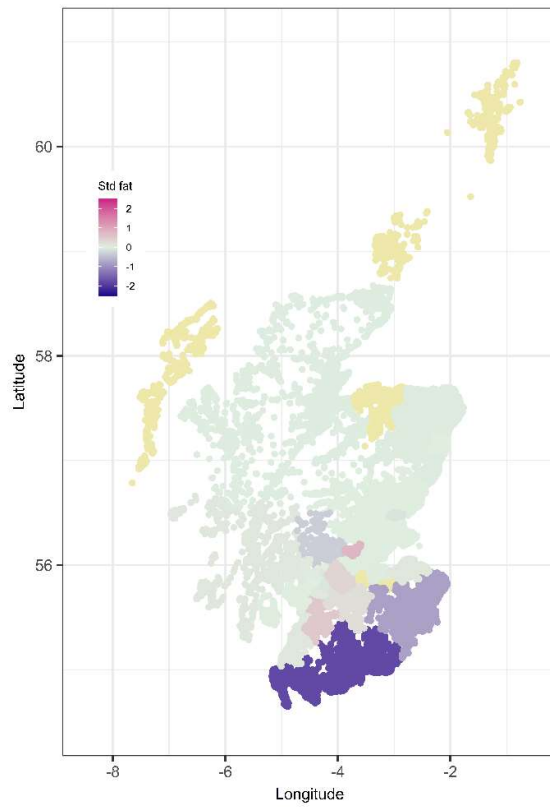
Supplementary Figure 1. Changes in the standardised means for each trait per area before and after adjusting for the lifestyle and socioeconomic covariates. Yellow: regions with less than 20 individuals, not considered.



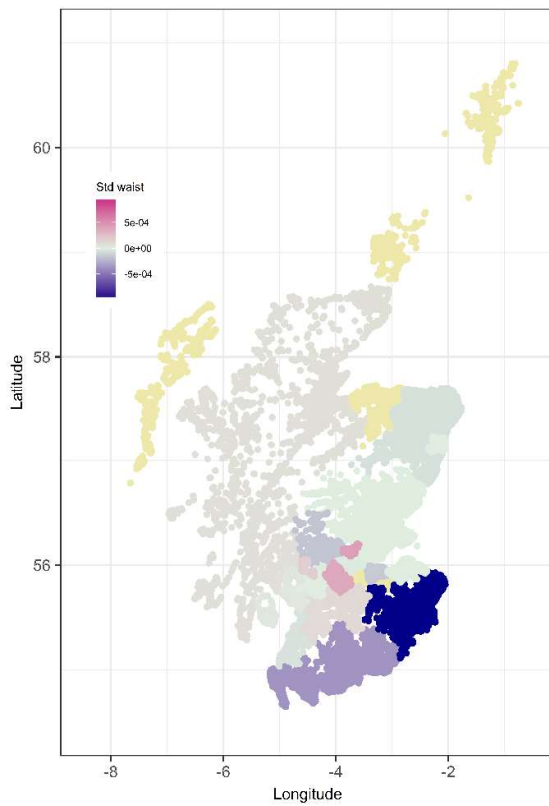
BIA Fat - Std(G + gPCs + sex + age + clinic)



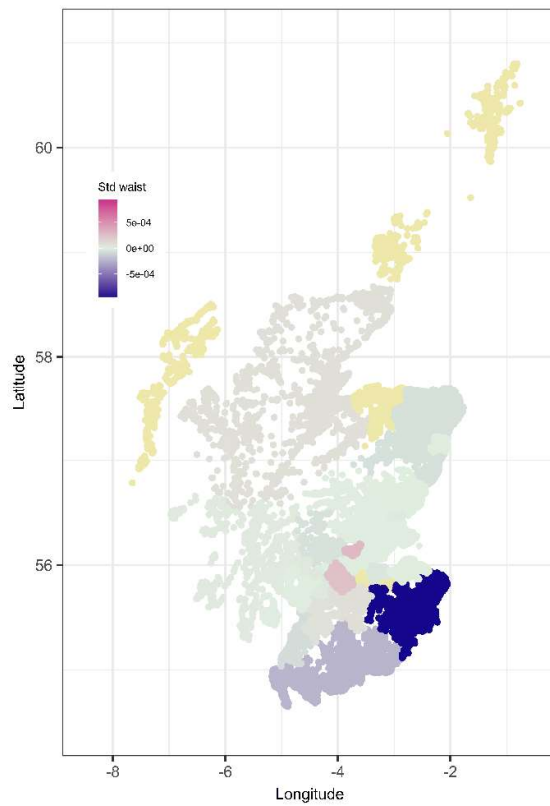
BIA Fat - Std(G + gPCs + SELS + sex + age + clinic)



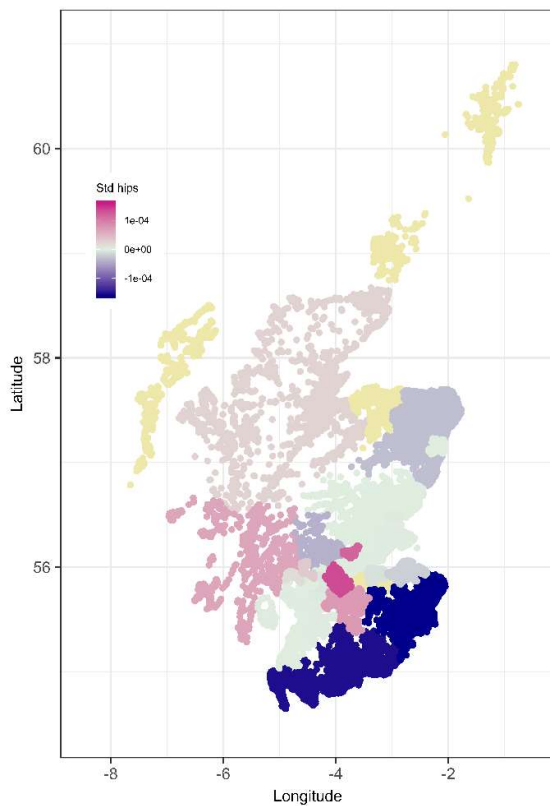
Waist - Std(G + gPCs + sex + age + clinic)



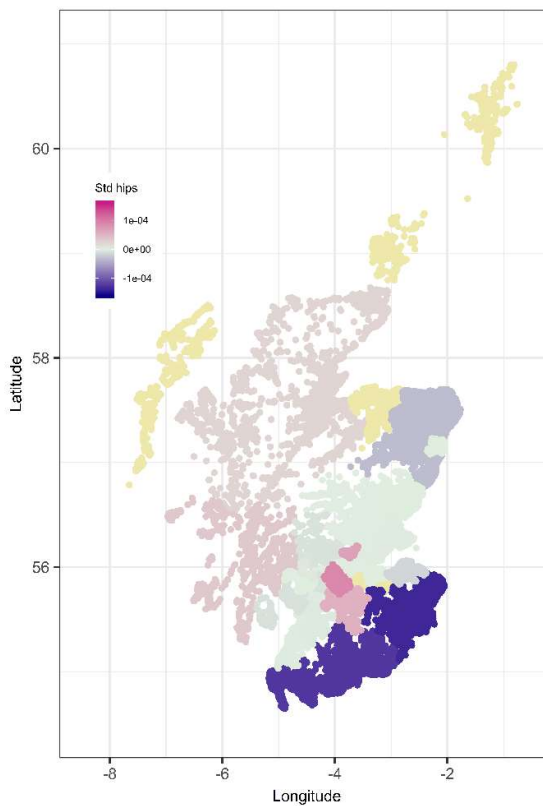
Waist - Std(G + gPCs + SELS + sex + age + clinic)



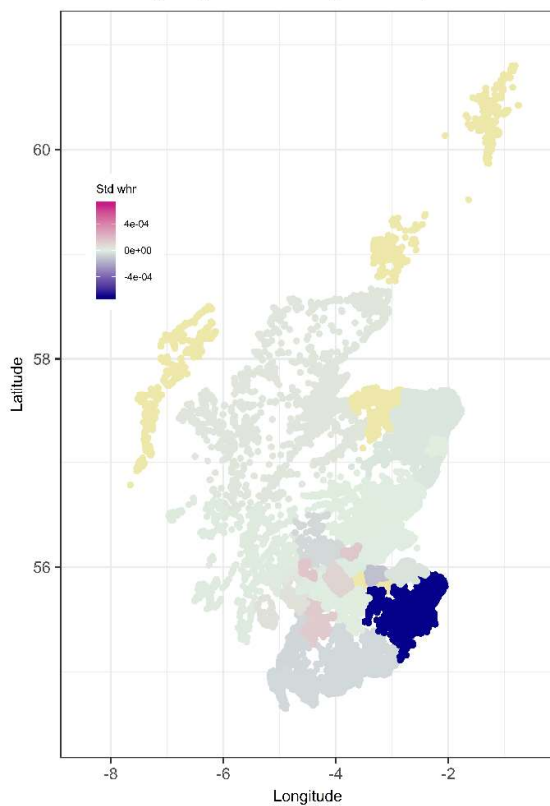
Hips - Std(G + gPCs + sex + age + clinic)



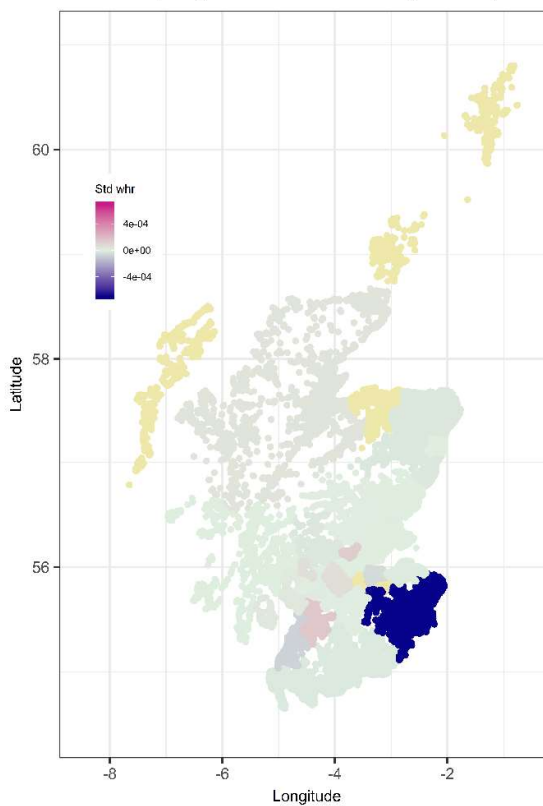
Hips - Std(G + gPCs + SELS + sex + age + clinic)



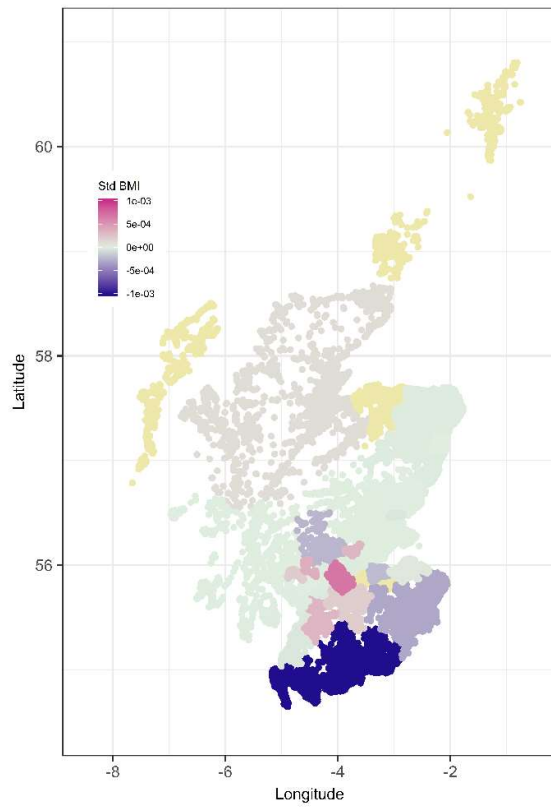
WHR - Std(G + gPCs + sex + age + clinic)



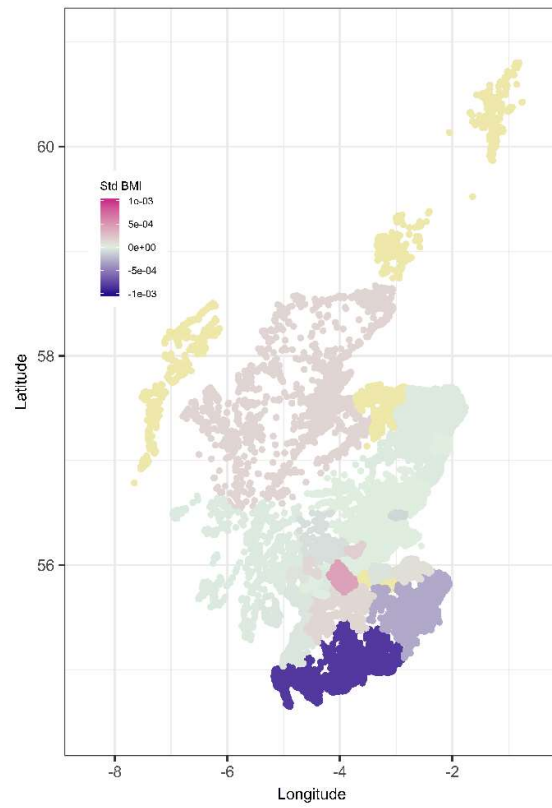
WHR - Std(G + gPCs + SELS + sex + age + clinic)



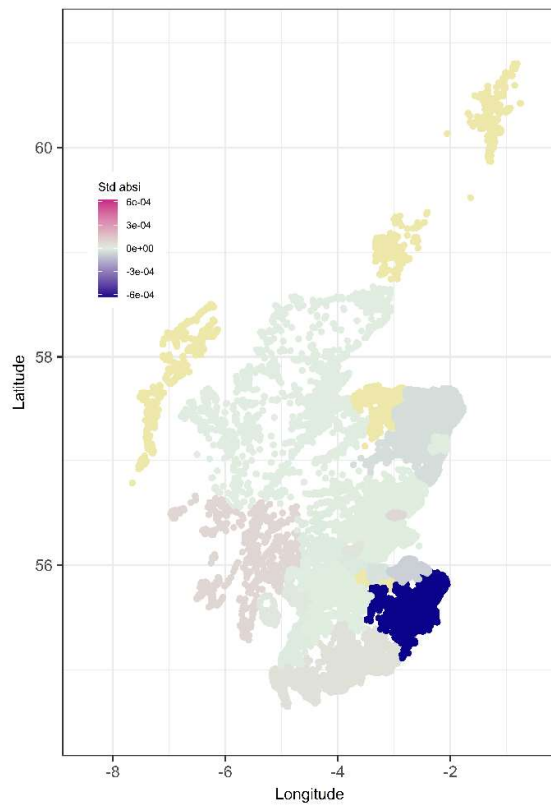
BMI - Std(G + gPCs + sex + age + clinic)



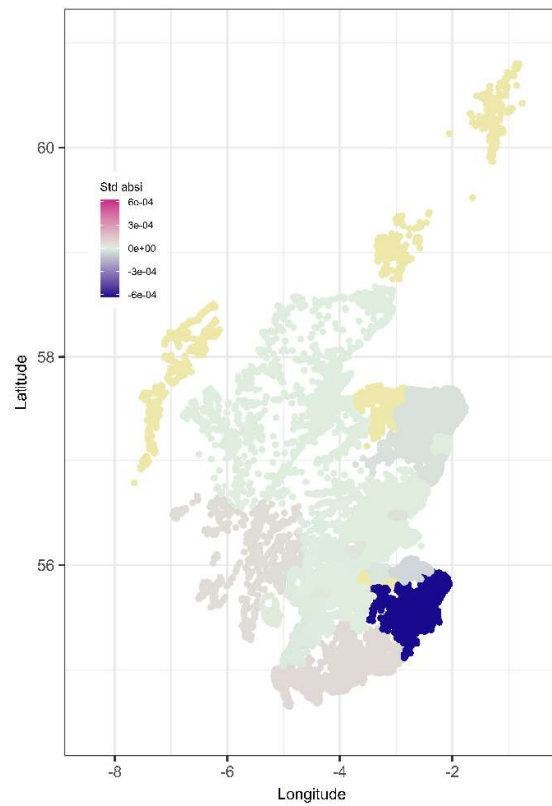
BMI - Std(G + gPCs + SELS + sex + age + clinic)



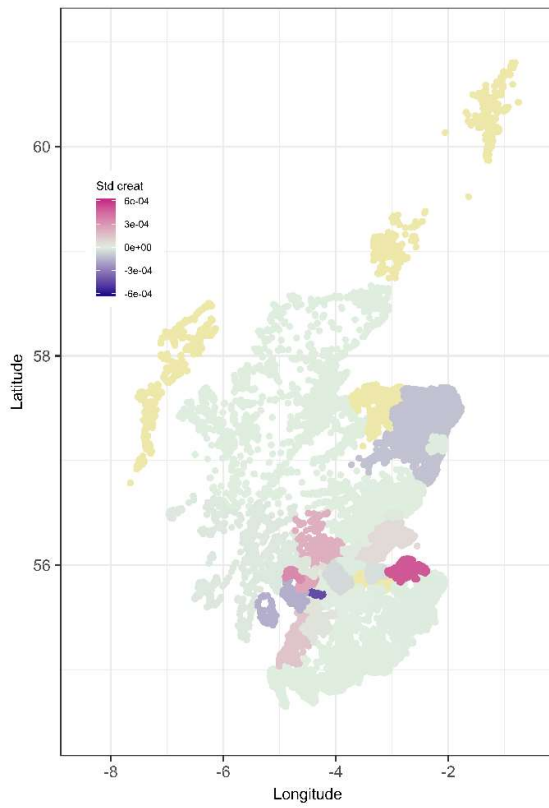
ABSI - Std(G + gPCs + sex + age + clinic)



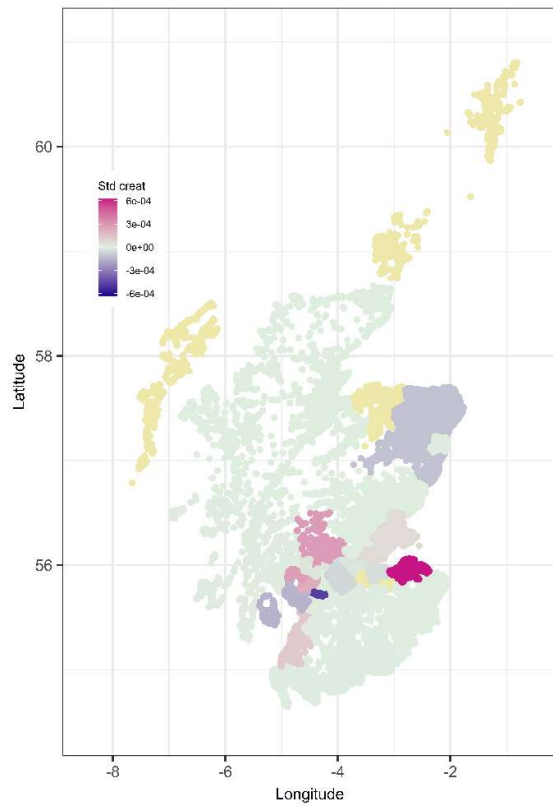
ABSI - Std(G + gPCs + SELS + sex + age + clinic)



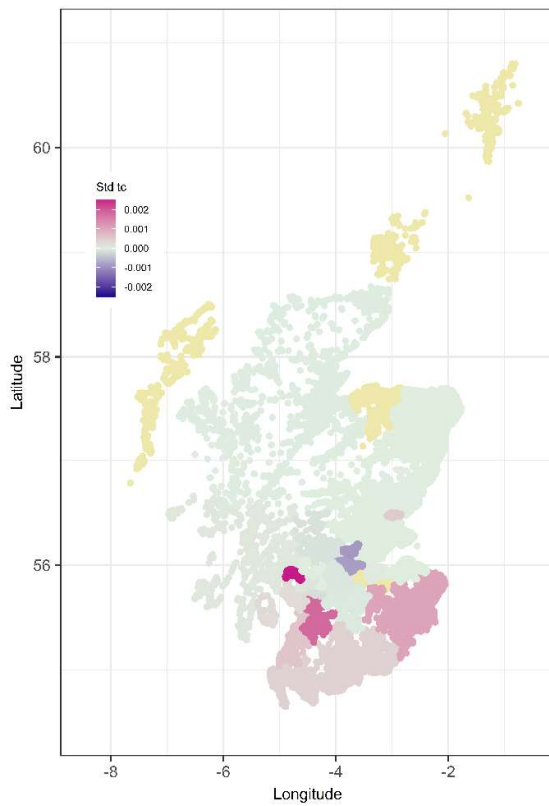
Creatinine - Std(G + gPCs + sex + age + clinic)



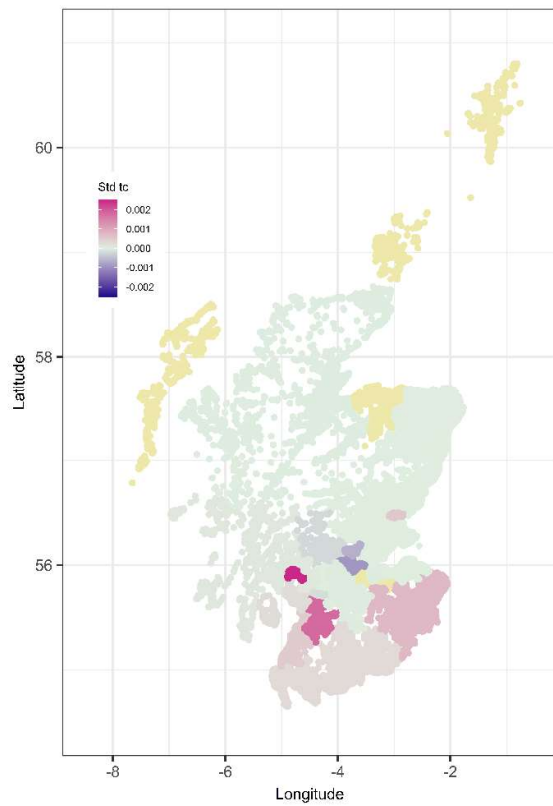
Creatinine - Std(G + gPCs + SELS + sex + age + clinic)



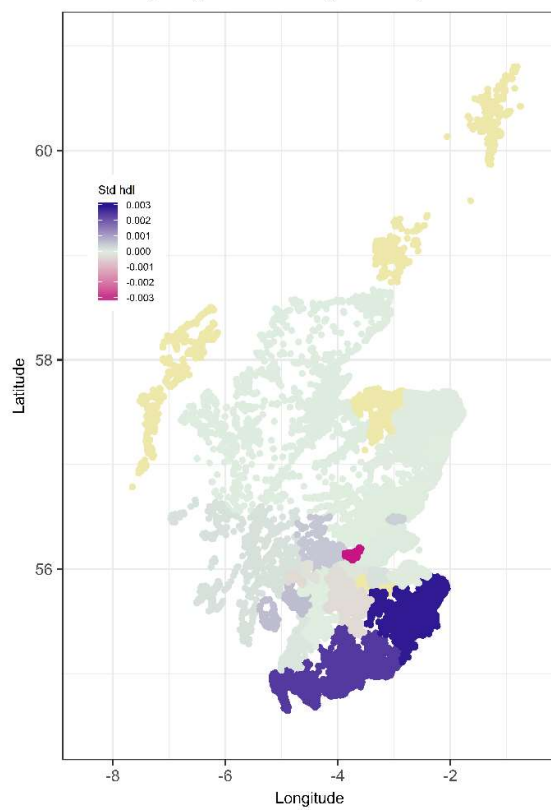
TC - Std(G + gPCs + sex + age + clinic)



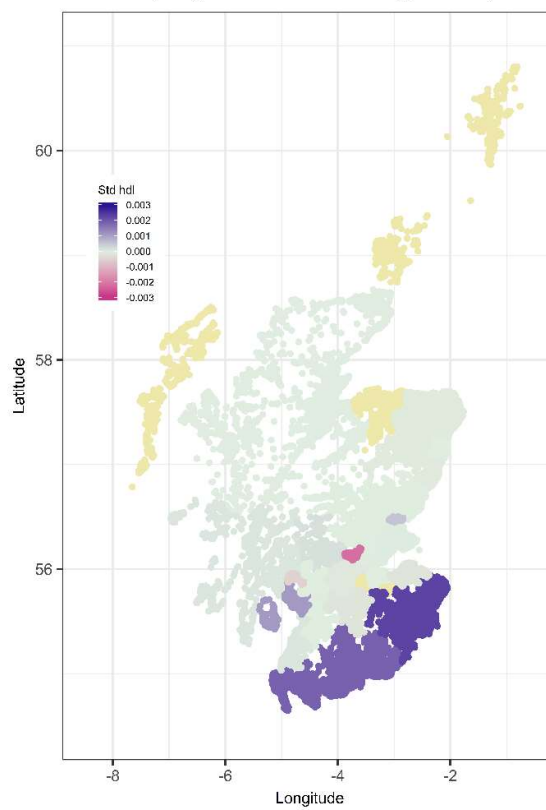
TC - Std(G + gPCs + SELS + sex + age + clinic)



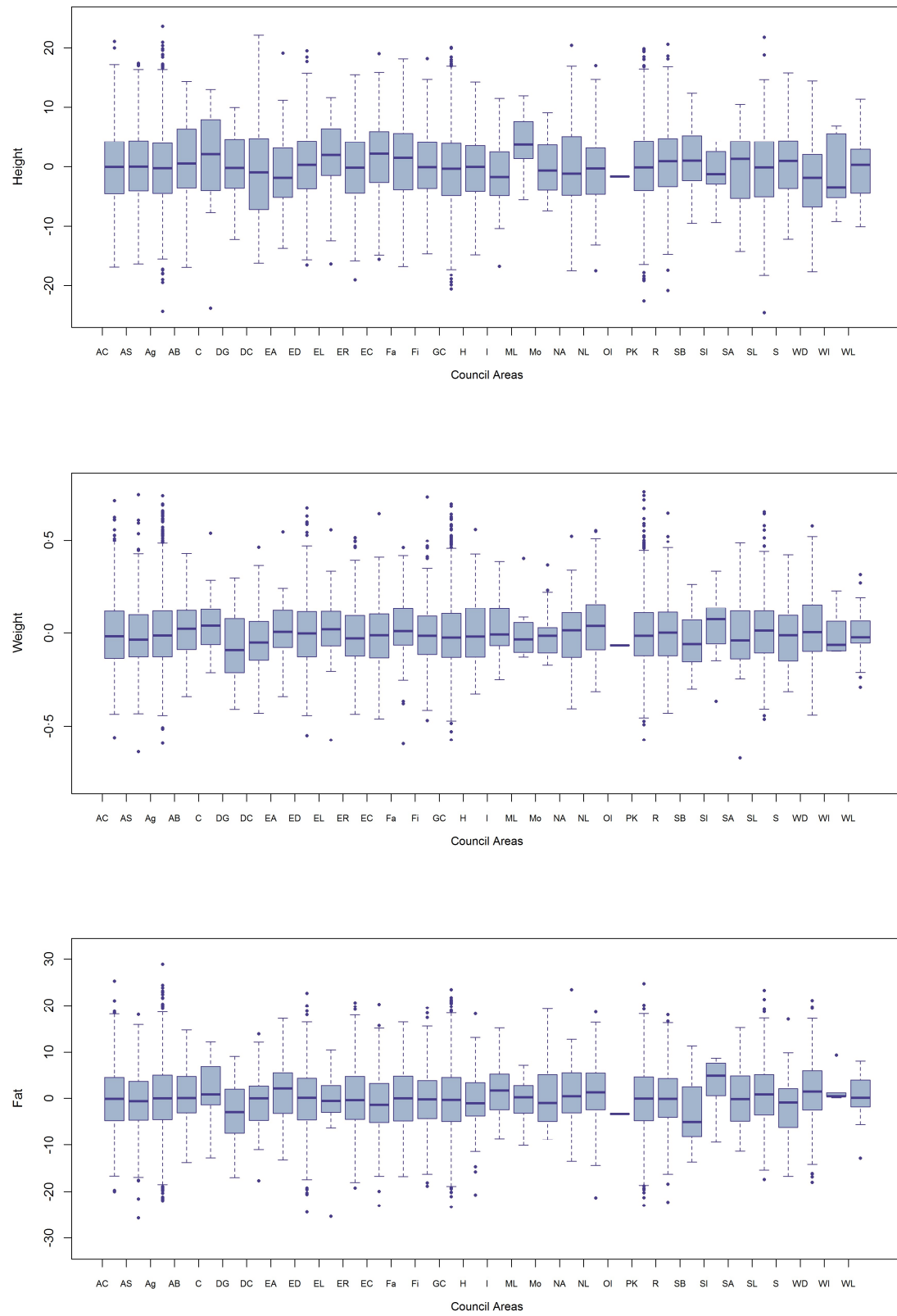
HDL - Std(G + gPCs + sex + age + clinic)

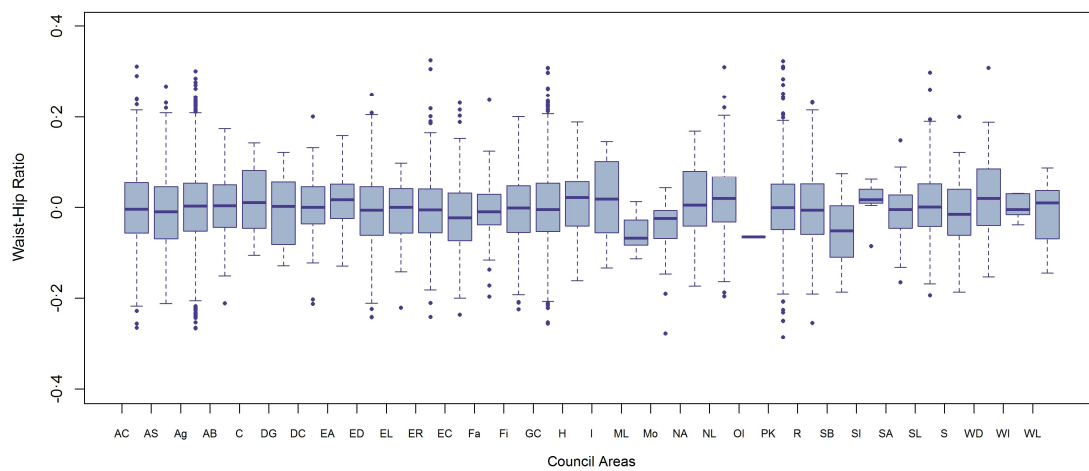
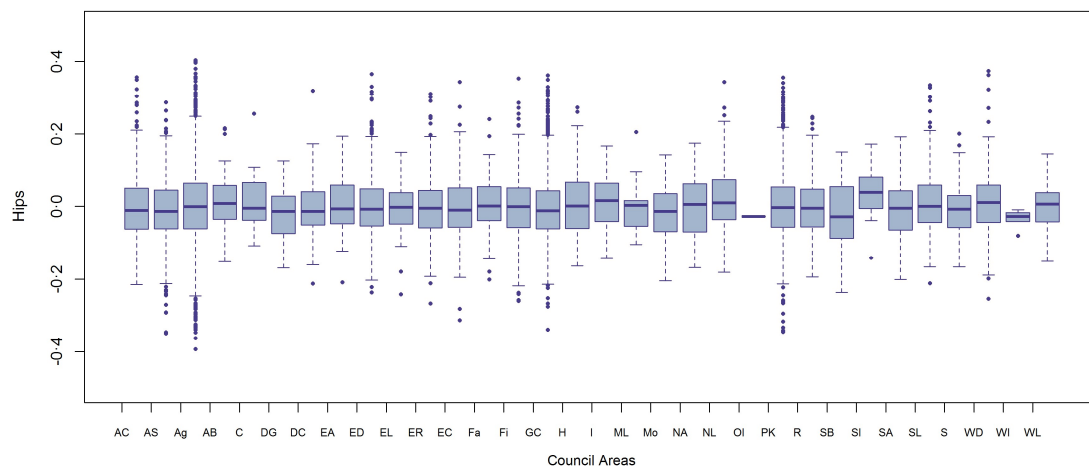
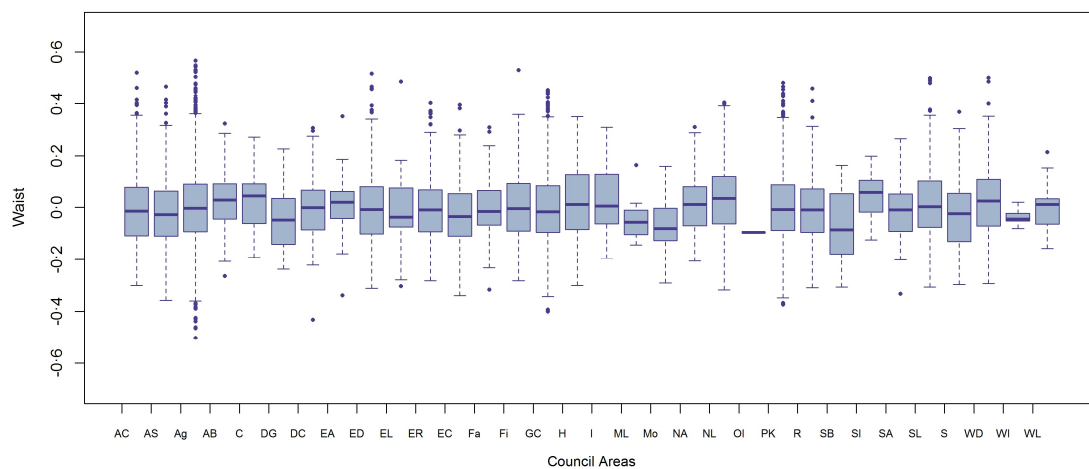


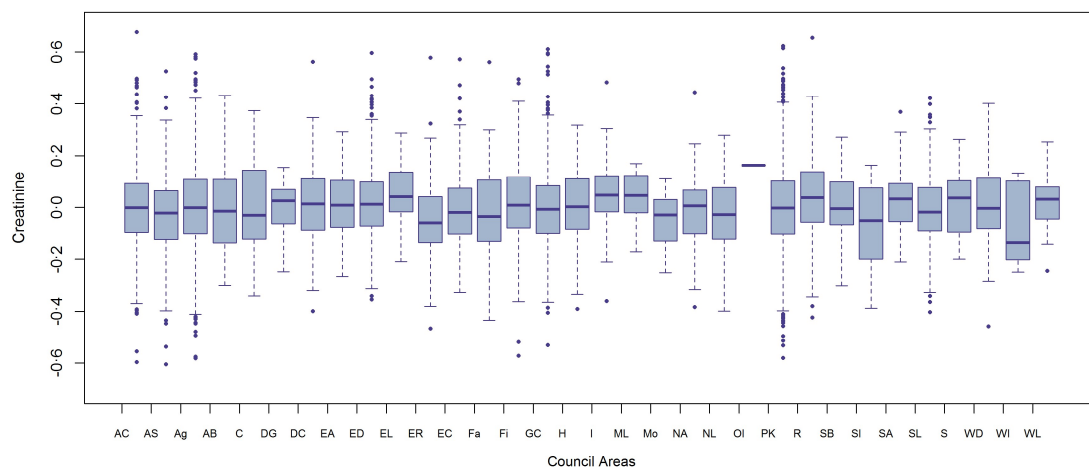
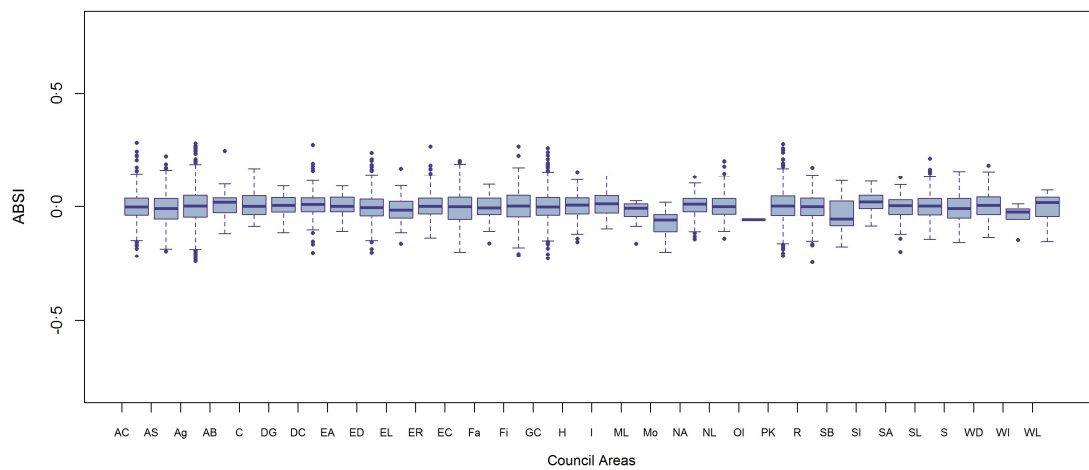
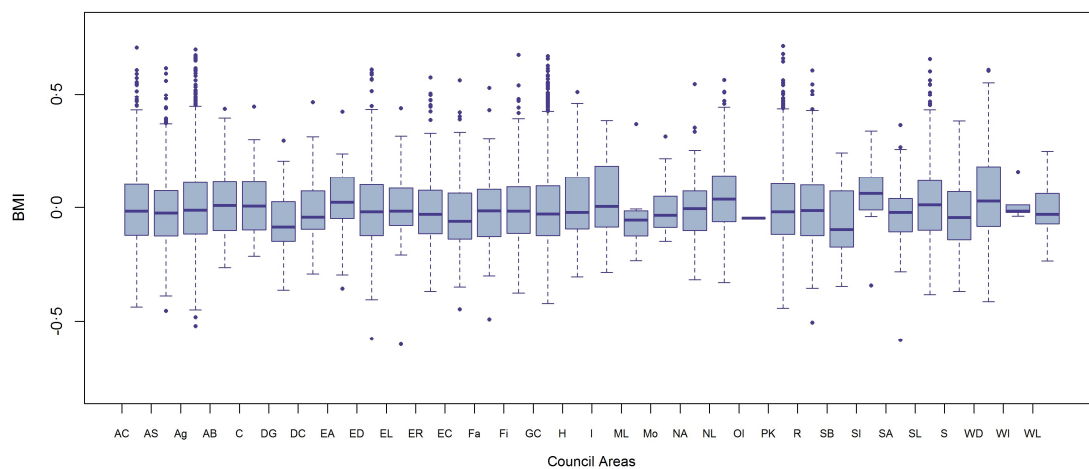
HDL - Std(G + gPCs + SELS + sex + age + clinic)

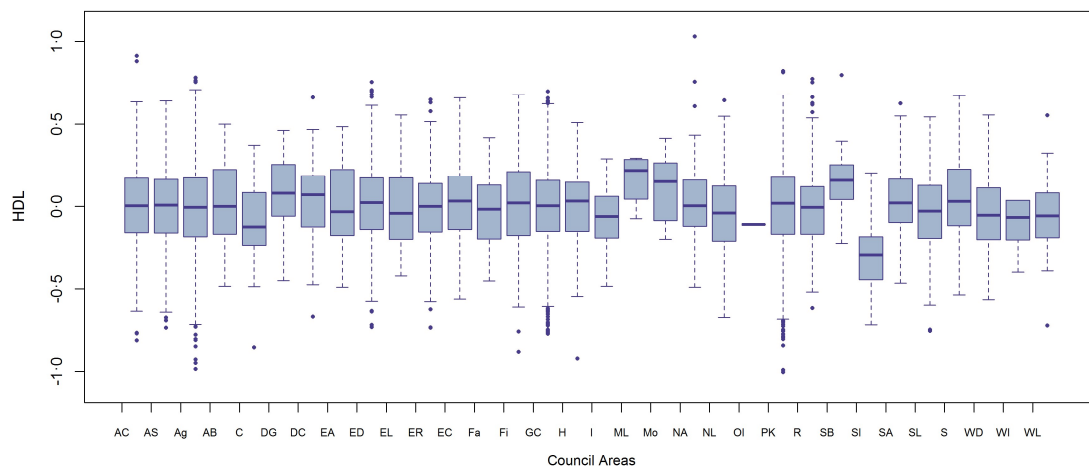
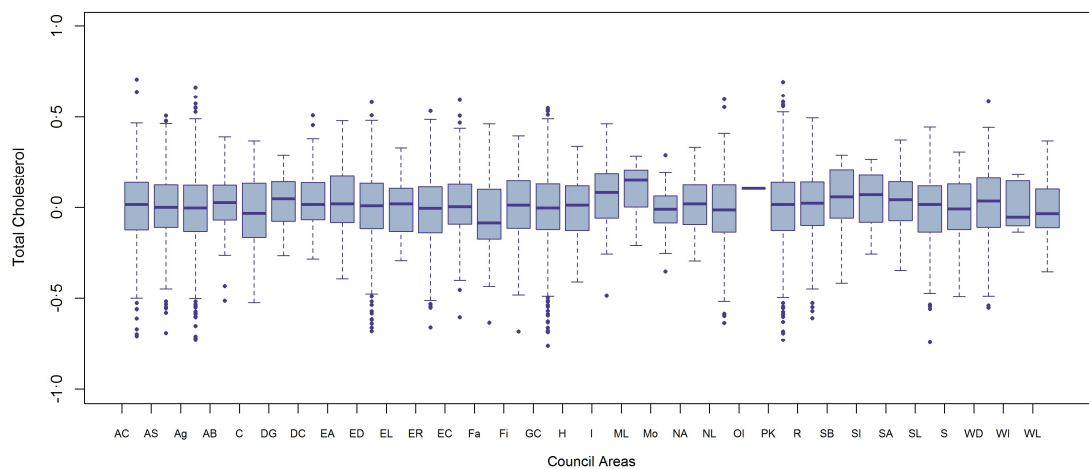


Supplementary Figure 2. Boxplot of the traits in the different council areas.

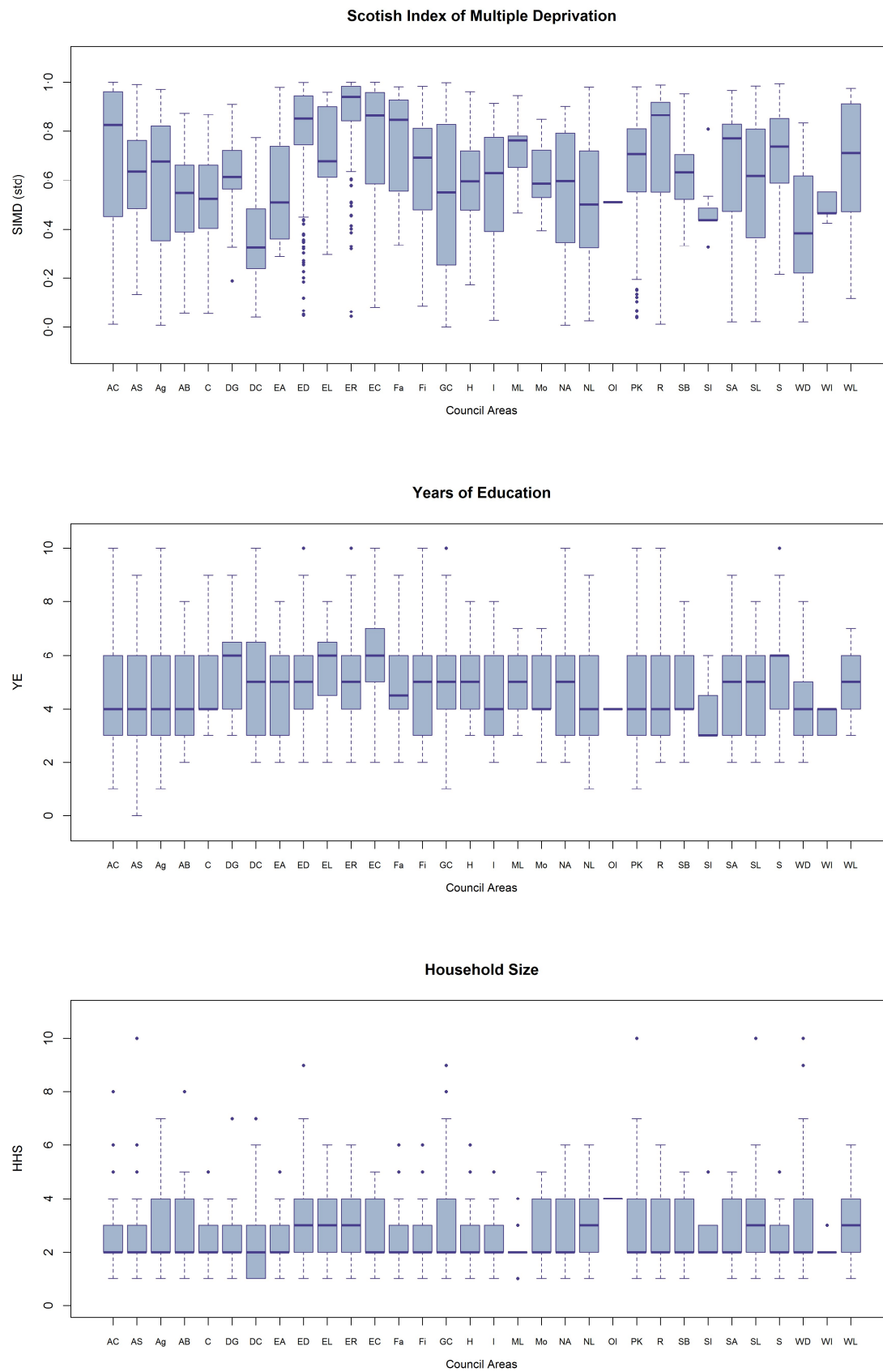


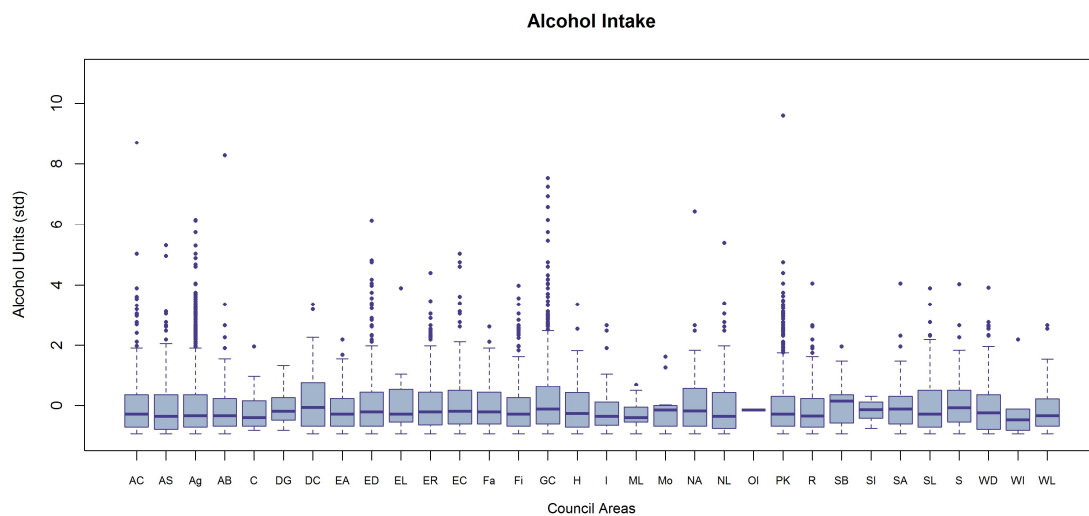
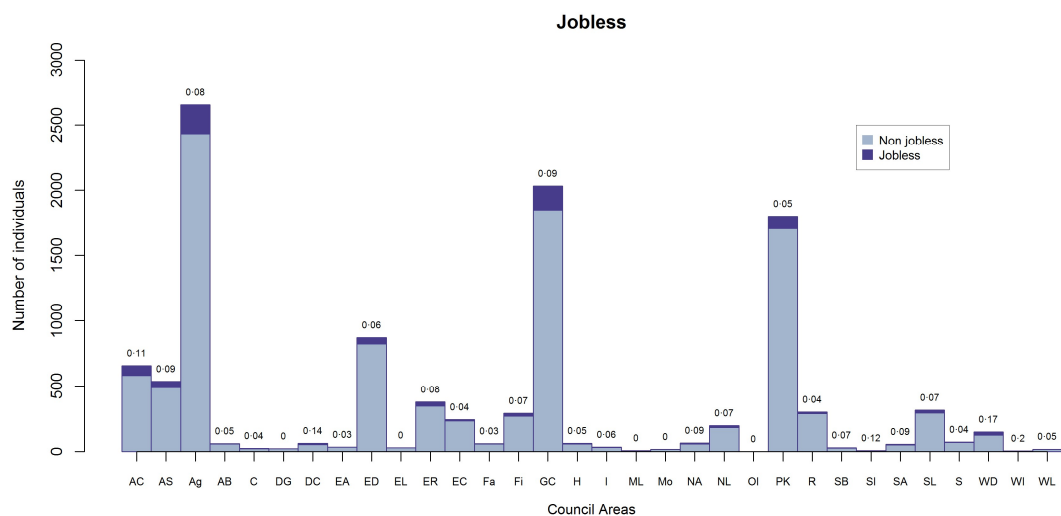
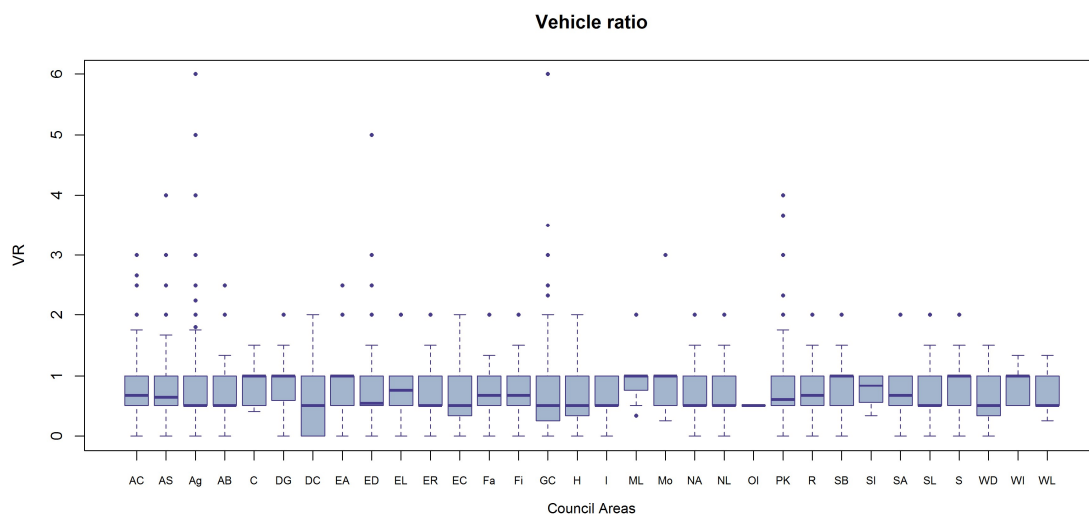


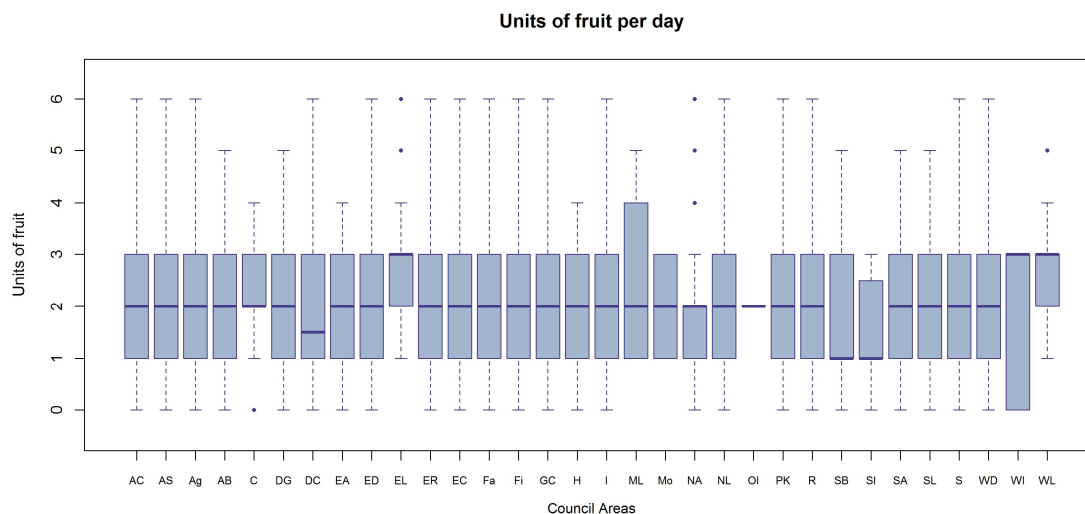
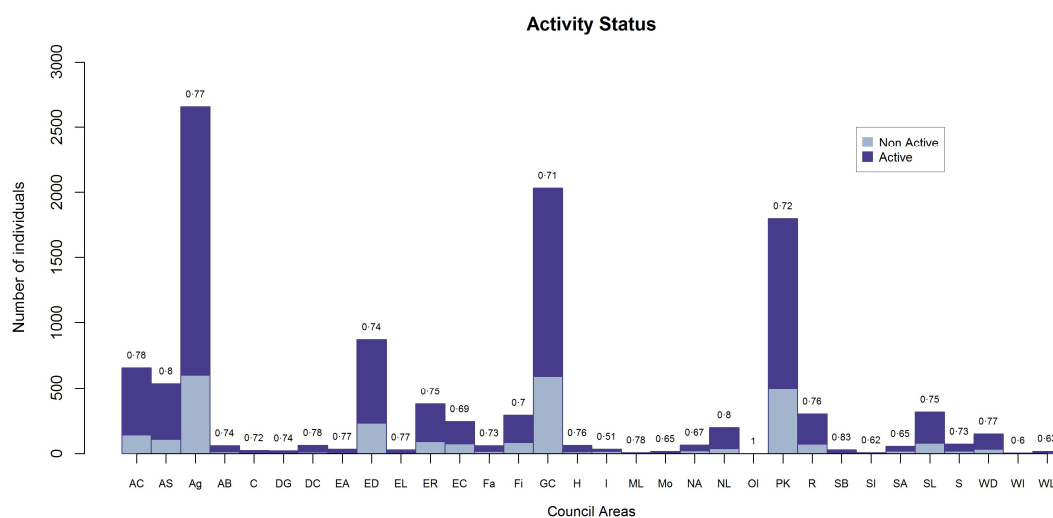
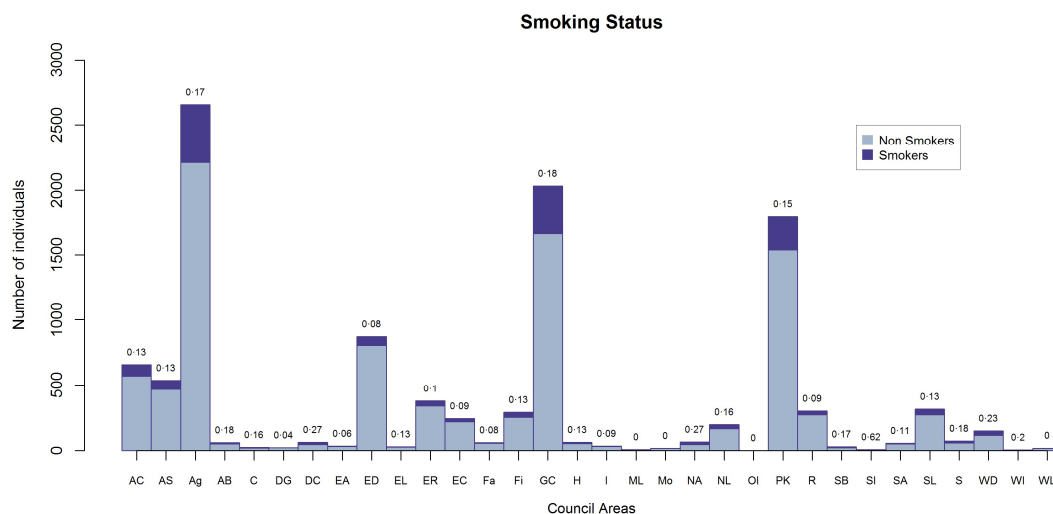


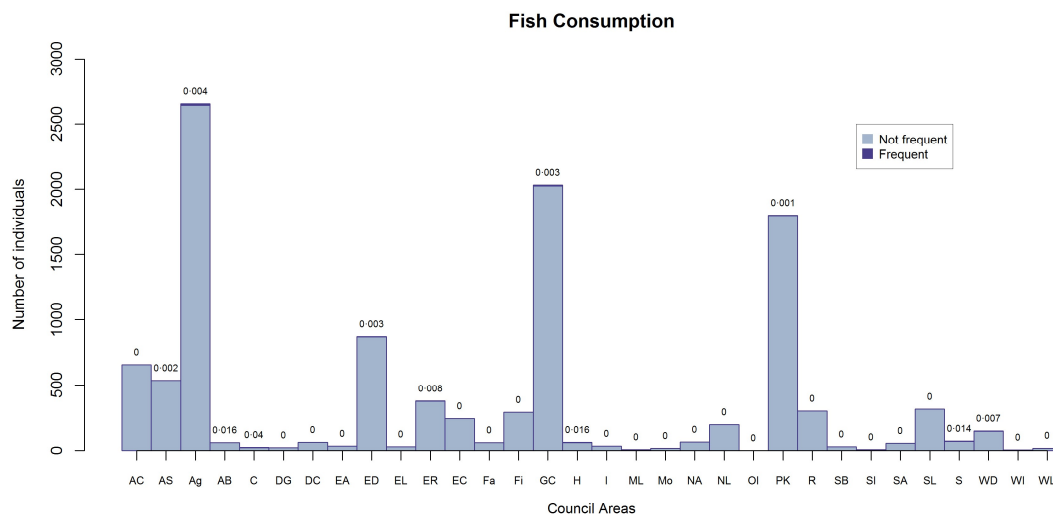
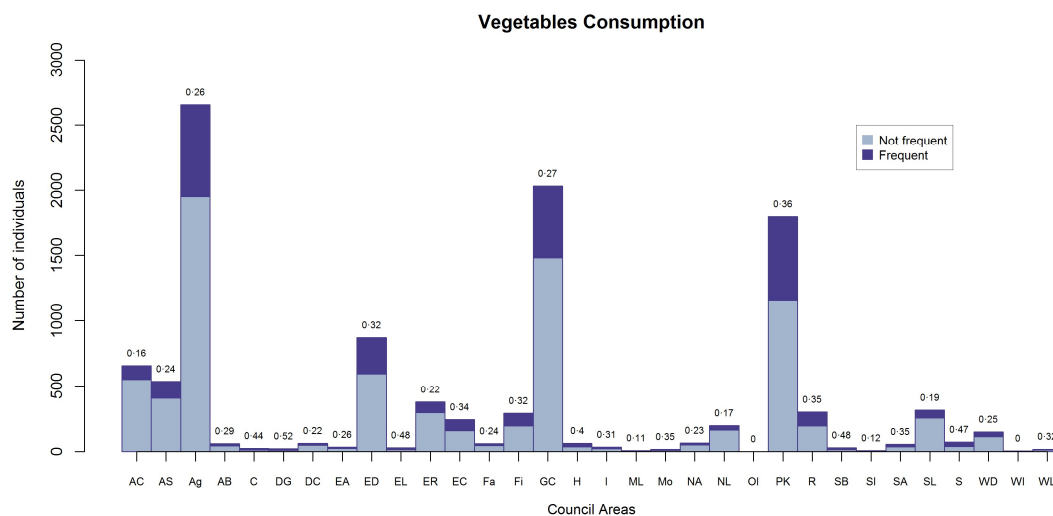
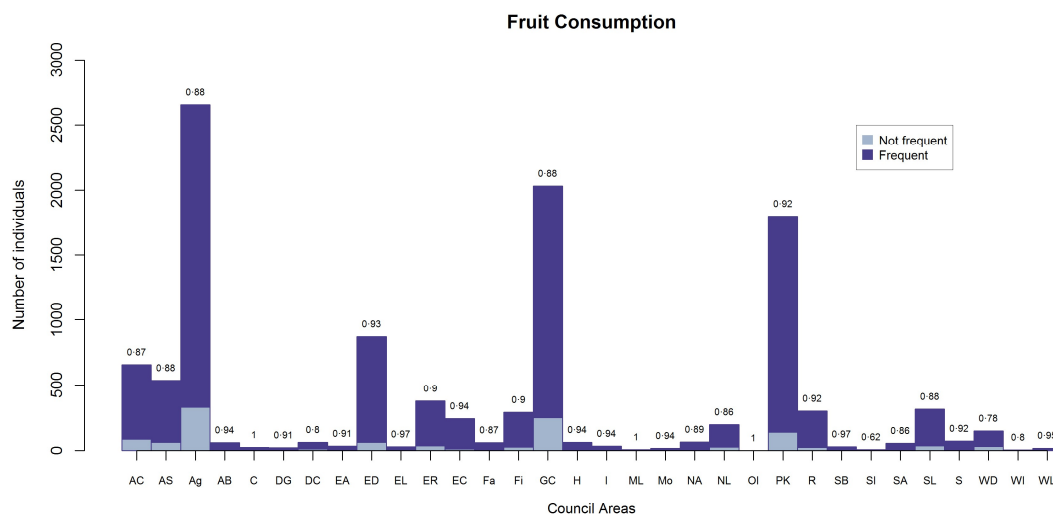


Supplementary Figure 3. Boxplot of the covariates in the different council areas.

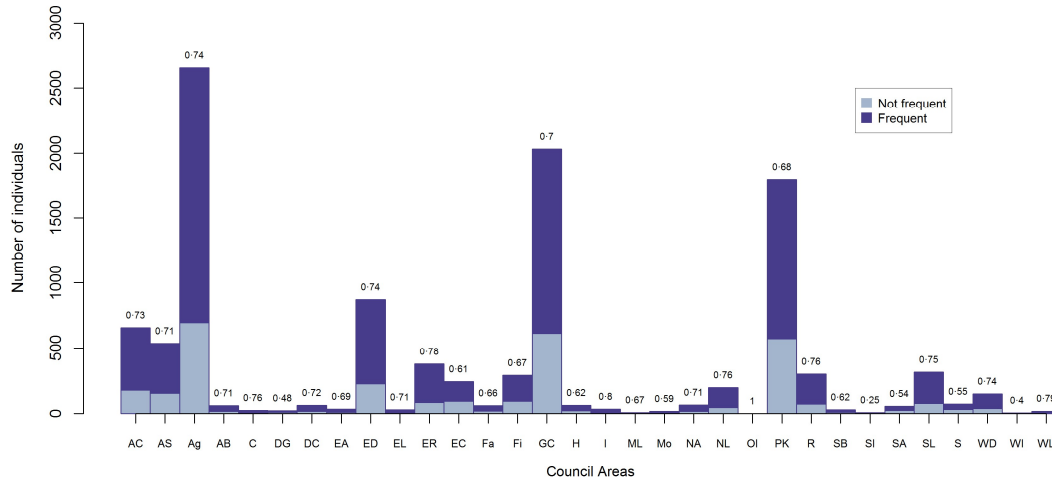




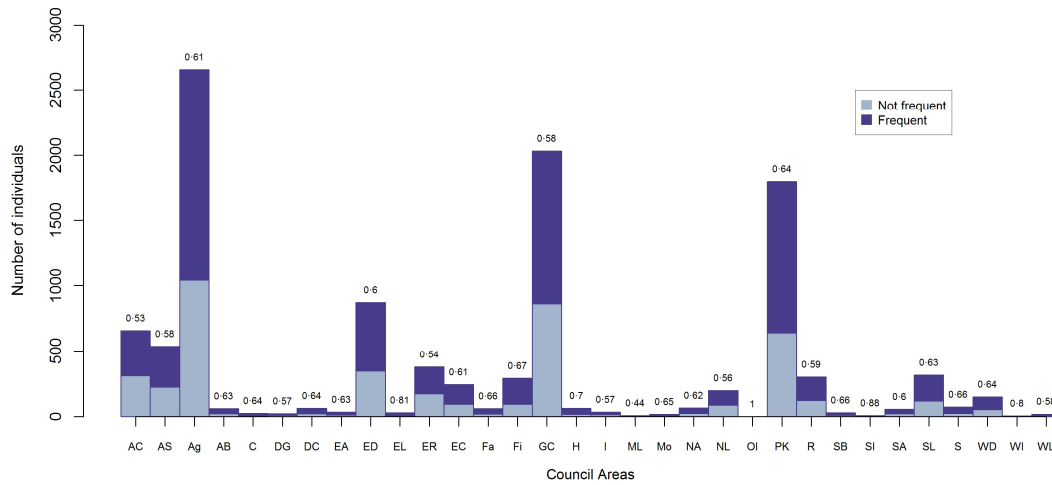




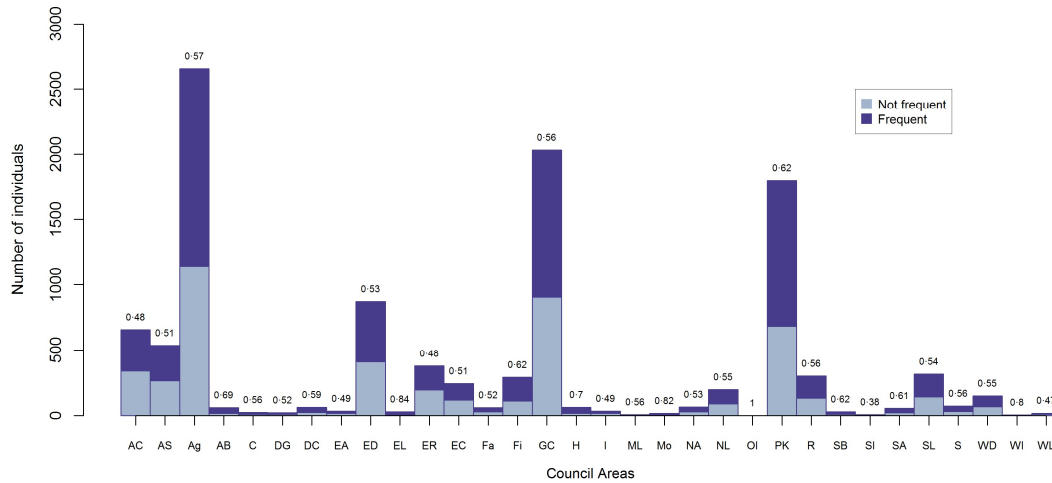
Poultry Consumption



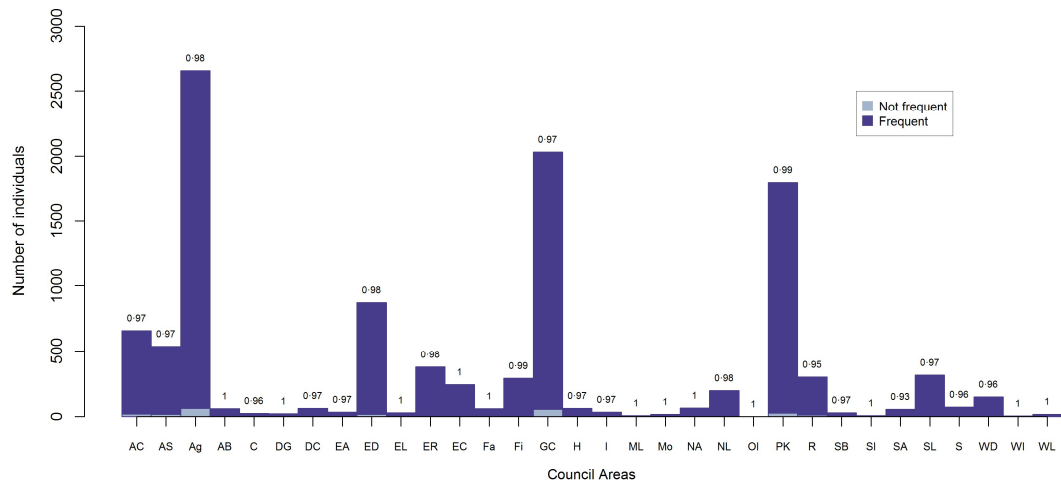
Meat Consumption



Eggs Consumption



Dairy Consumption



Supplementary Note 1. Quality control of lifestyle and socioeconomic variables.

Two different questionnaires (named V2 and V5) were taken by different subsets of the ~24,000 individuals in GS:SFHS. Each contained different questions to gather similar information. In order to obtain comparable values for variables collected from both questionnaires, so that information on each variable was available for all individuals, the quality control and standardisation described below was performed for each variable named.

Alcohol units intake information

In V2 people were asked number of units drank the previous week whereas in V5 they were asked about amount drank of specific drinks (also the previous week). V5 drinks were converted into units based on the values suggested in V2 for self-calculation of units, namely:

- 1 Glass of wine (red or white) = 1.5 units
- 1 Pint of beer/cider = 2 units
- 1 Bottle of beer/cider = 2 units
- 1 Measurement of spirits = 1 unit

After standardisation, data from both questionnaires were merged and an outlier removed (ID 128029 in V2, with more than 300 Units consumed during the week previous to answering the questionnaire).

Smoking status information

Both questionnaires shared the question *Have you ever smoked tobacco?* With possible answers:

1 - Yes, currently smoke; 2 - Yes, but stopped within past 12 months; 3 - Yes, but stopped more than 12 months ago; 4 - No, never smoked

The data was merged for that variable and the variable Smoking Status was created: 1: Smoker (4014 individuals), 0: NonSmoker (19265 individuals) (269 missing).

Activity level information

V2 separated work and leisure activities whereas V5 asked general questions and separated walking in a different category. To harmonise the information gathered by each questionnaire, 2 new measures summarising the hours per week spent doing vigorous and moderate activities were created (vigorous (VA) and moderate (MA) activity).

From V2, we obtained VA as the sum of time “very active” per week during work and during non-working time for vigorous activities; MA was the sum of time moderately active per week during work and during non-working time for moderate activity.

From V5, we obtained VA as the from the Time very active last week for vigorous activities and MA as the sum of time moderately active last week and the time spent walking last week for moderate activity.

The variables obtained were standardised before merging the data from the different questionnaires. For those individuals with missing values for VA but responses for MA, VA was set to zero instead to missing (assuming therefore that if they replied to moderate activity section and did not to vigorous activity one that they did not perform any vigorous activity).

The values for vigorous (VA) and moderate (MA) activity were converted into a new variable: Activity level, with 0 (corresponding to non-active) for those with $VA_{std} \leq 0.5$ and $MA_{std} \leq 1$, and 1 (active) for the rest.

17517 individuals were classified as active and 6031 as non-active.

Diet information

Individuals completing the V2 questionnaire were asked to quantify the amount of different foods eaten while those completing the V5 questionnaire were asked about the frequency of intake of the different types of food (possible answers were 1. Daily; 2. 5-6 times per week; 3. 2-4 times per week; 4. Once per week; 5. Less than once per week; 6. Less than once per month; 7. Never).

For responses to the V2 questionnaire, the amount consumed for each food group was converted into yearly consumption and that into the V5 intake frequency categories.

A new set of dichotomous variables was created from the categories. Intake for each food group was classified as “not usual consumption” (0, less than 1 day per week) or “usual consumption” (1, more than 1 day per week) creating the variables: Fruit Consumption, Vegetables Consumption (green vegetables + other

vegetables), Fish Consumption (oily and other types of fish), Poultry Consumption, Meat Consumption, Eggs Consumption and Dairy Consumption.

Since the calculation in V2 implied that the response could be a fractional number, when these occurred they were rounded to the closest integer. Individuals (always from V2, probably because of how the question was asked) with values >10 were considered outliers and set to missing.

A common variable for both questionnaires was Fruit Units (per day), so it was also merged.

Years of education

Both questionnaires shared the question *How many years altogether did you attend school/study full-time?*

The values are categorical but were fitted as a continuous variable in all the models. The meaning of the categories was:

0: 0; 1: 1-4; 2: 5-9; 3: 10-11; 4: 12-13; 5: 14-15; 6: 16-17; 7: 18-19; 8: 20-21; 9: 22-23; 10: 24+