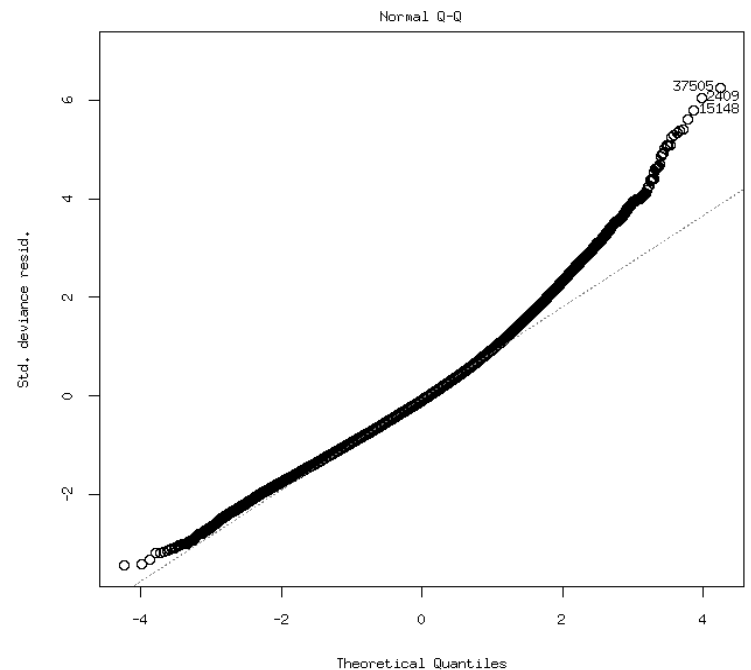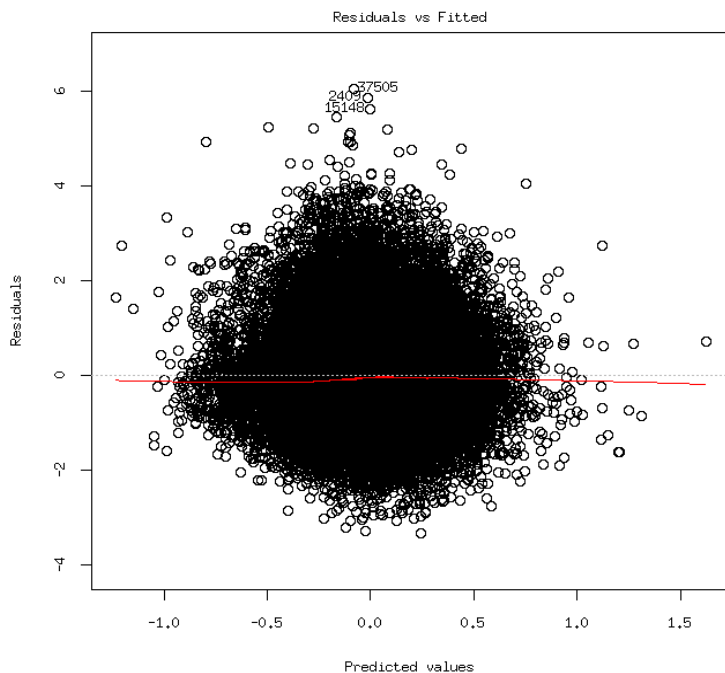Regression coefficients and variation explained for the diet exposures and bmi:

the dependent variable bmi was log transformed and standardized by (x-mean(x)) / sd(x)
the independent variables were standardized only

The independent dataset of Swedish only (47107 subjects):

|	Regression coefficient	Variance explained	p-value
POLYsum1	0.066335	0.000121	0.01884
MONOsum1	0.086867	0.000764	3.51e-09
mfetsum1	-0.144135	0.001219	8.58e-14
fettsum1	0.299495	0.001869	< 2e-16
acids	-0.111648	0.000429	9.62e-06
kolhsum1	0.225986	0.003014	< 2e-16
sacksum1	-0.025627	0.000419	1.22e-05
TRANSsum1	-0.065616	0.001406	1.12e-15
protsum1_anim	0.167995	0.009045	< 2e-16
kolesum1	-0.030810	0.000166	0.00592
FULLKsum1	-0.020467	0.00031	0.00017
NATRsum1	0.044330	0.001527	< 2e-16

The 2 visits subset of Swedish only (66228 subjects),
visit==1 (33114 subjects):

	Regression coefficient	Variance explained	p-value
POLYsum1	0.106650	0.000511	5.48e-05
MONOsum1	0.027709	0.000139	0.035241
mfetsum1	-0.205119	0.003082	< 2e-16
fettsum1	0.256505	0.001609	8.18e-13
acids	-0.136136	0.001106	2.98e-09
kolhsum1	0.136386	0.001336	6.97e-11
sacksum1	-0.035134	0.000863	1.60e-07
TRANSsum1	-0.031824	0.000415	0.000277
protsum1_anim	0.134038	0.007262	< 2e-16
kolesum1	-0.035067	0.000345	0.000929
FULLKsum1	-0.032645	0.000849	2.03e-07
NATRsum1	0.048958	0.001766	6.41e-14
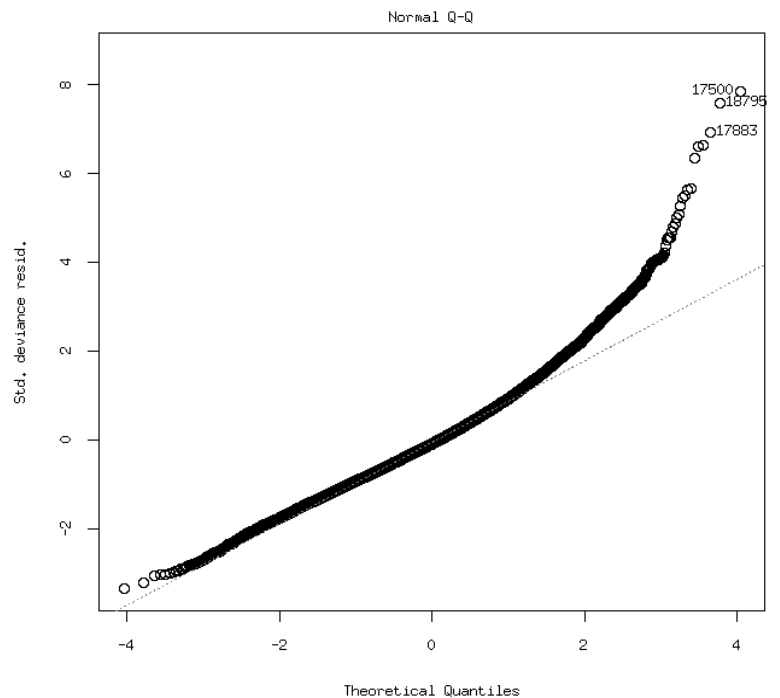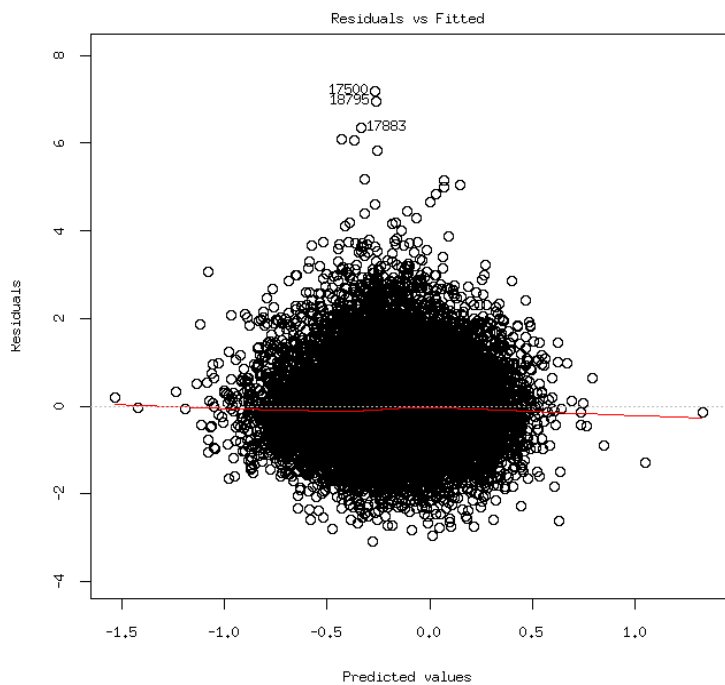
The 2 visits subset (66228 subjects),
visit==2 (33114 subjects):

	Regression coefficient	Variance explained	p-value
POLYsum1	0.127463	0.000451	0.000134
MONOsum1	0.048996	0.000184	0.014670
mfetsum1	-0.209340	0.002011	7.08e-16
fettsum1	0.399718	0.002753	< 2e-16
acids	-0.186931	0.001155	9.75e-10
kolhsum1	0.276879	0.005074	< 2e-16
sacksum1	-0.053294	0.001900	4.40e-15
TRANSsum1	-0.022010	0.000159	0.023362
protsum1_anim	0.176034	0.010027	< 2e-16
kolesum1	-0.031344	0.000168	0.019769
FULLKsum1	-0.044796	0.001462	6.01e-12
NATRsum1	0.059482	0.002741	< 2e-16
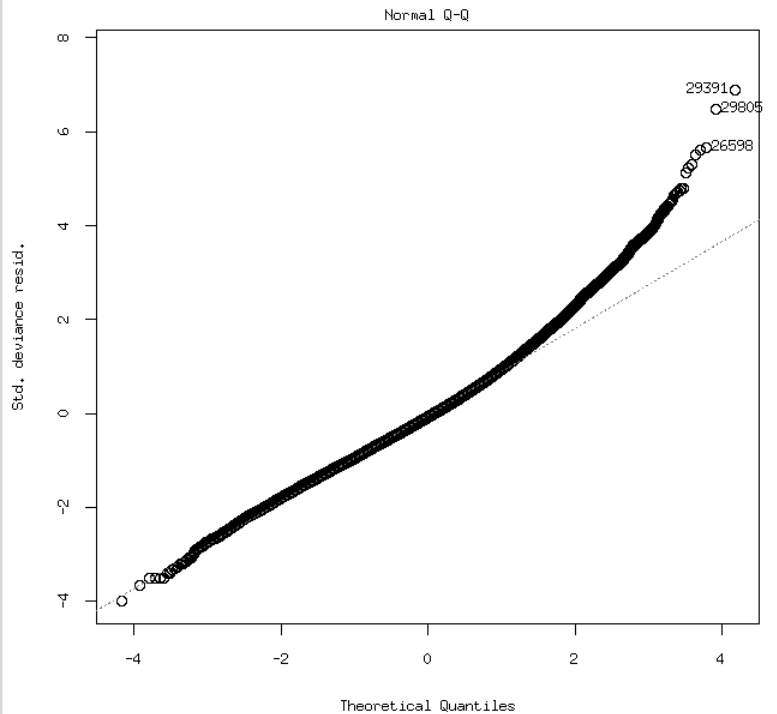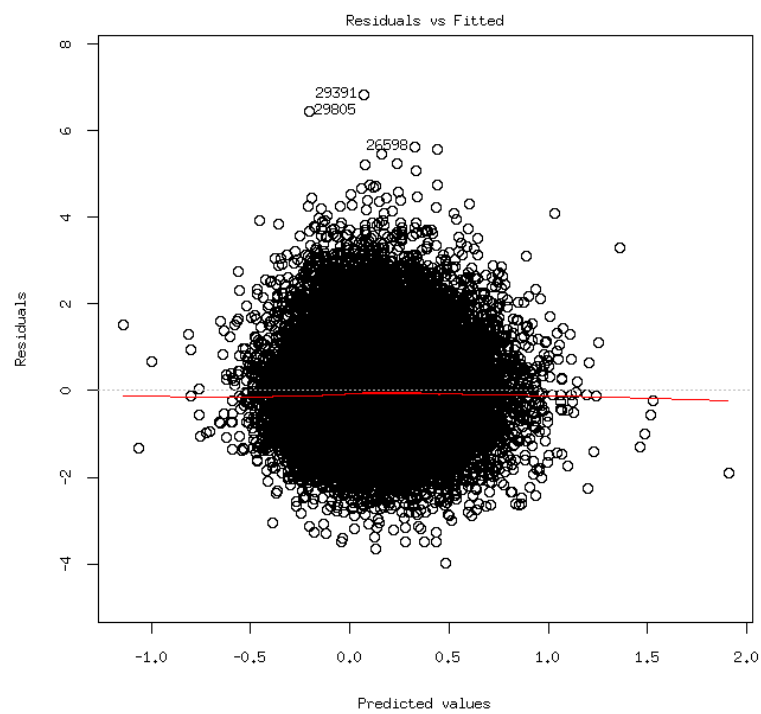
Categorizing each variable based on the guidelines:

Since the focus is on the obesogenic type of unhealthy environment, 0 score is given to subjects below the recommended intake of the nutrient, 1 to those that are in the limits of the recommended intake and 2 to those that are above the upper limit of the recommended nutrient intake.
There are variables where the guidelines will only recommend one limit. In those cases, there can only be two categories and we will have decide how to code for this, for now I assign only 1 and 2, without the 0, that is why some variables will have no subjects in the category 0.
There are variables where the recommendations have not been specified, like for example trans fat is recommended to be as low as possible and cholesterol is said to be in the recommended limits as a consequence of other recommendations. For those variables, where recommendations are not specified, tertiles were calculated. This might not be the best solution, as other variables are very disproportionate, including bmi.

Number of subjects in each category for each variable:

	Number of subject with 0	Number of subject with 1	Number of subject with 2
POLYsum1	29632	33028	1697
MONOsum1	14121	50070	166
mfetsum1	0	5364	58993
fettsum1	3735	48772	11850
acids	0	61995	2362
kolhsum1	19385	42937	2035
sacksum1	0	58641	5716
TRANSsum1	21238	21237	21882
protsum1_anim	21238	21237	21882
kolesum1	21238	21237	21882
FULLKsum1	21882	21237	21238
NATRsum1	0	64325	32
ensum1	21238	21237	21882

	bmi<18.5	18.5=<bmi<25	25=<bmi<30	30=<bmi
bmi	424	31957	25884	7805

Regression coefficients and variation explained for the categorized diet exposures and bmi:

visit==1:

	Regression coefficient, unstandardized	Regression coefficient, standardized	Variance explained	p-value
POLYsum1_score	0.028938	0.015458	0.000193	0.013327
MONOsum1_score	0.028404	0.011943	9e-05	0.090641
mfetsum1_score	-0.153107	-0.037843	0.001071	5.24e-09
fettsum1_score	-0.008232	-0.003716	6e-06	0.657579
acids_score	-0.069953	-0.013039	0.000164	0.022218
kolhsum1_score	-0.032045	-0.015523	0.000108	0.063449
sacksum1_score	-0.036238	-0.011847	0.000133	0.039644
TRANSsum1_score	-0.057104	-0.046739	0.001153	1.37e-09
protsum1_anim_score	0.109930	0.089976	0.006267	< 2e-16
kolesum1_score	-0.050186	-0.041077	0.000953	3.66e-08
FULLKsum1_score	0.033025	0.027030	0.000661	4.54e-06
NATRsum1_score	0.554910	0.013165	0.000184	0.015543
ensum1_score	0.003916	0.003205	9e-06	0.598938

visit==2:

	Regression coefficient, unstandardized	Regression coefficient, standardized	Variance explained	p-value
POLYsum1_score	0.0223012	0.0122590	0.000108	0.061551
MONOsum1_score	0.0623012	0.0259191	0.000384	0.000424
mfetsum1_score	-0.1145644	-0.0345362	0.000803	3.45e-07
fettsum1_score	0.0979603	0.0488105	0.001031	7.66e-09
acids_score	0.0426673	0.0080906	6.2e-05	0.158186
kolhsum1_score	0.0516267	0.0273561	0.000306	0.001638
sacksum1_score	-0.0546422	-0.0126053	0.000153	0.026092
TRANSsum1_score	-0.0544908	-0.0446007	0.001261	1.68e-10
protsum1_anim_score	0.1353006	0.1107435	0.008416	< 2e-16
kolesum1_score	-0.0052186	-0.0042714	8e-06	0.609593
FULLKsum1_score	0.0415119	0.0339775	0.000975	1.94e-08
NATRsum1_score	0.4906168	0.0101970	0.000108	0.061490
ensum1_score	0.0251363	0.0205740	0.000343	0.000861

Three different diet scores:
-diet score 1: combining all individual scores as they are
-diet score 2: combining all individual scores, but multiplying them with the standardized beta coefficient of the continuous variable from an independent data set
-diet score 3: combining only scores from the variables which are complying with the guidelines used to create the scores

visit=1:

	Regression coefficient, unstandardized	Regression coefficient, standardized	Variance explained	p-value
diet score 1	0.005563	0.014578	0.000171	0.01954
diet score 2	0.448963	0.072653	0.00521	< 2e-16
diet score 3	0.038703	0.057360	0.003297	< 2e-16

visit=2:

	Regression coefficient, unstandardized	Regression coefficient, standardized	Variance explained	p-value
diet score 1	0.0267616	0.0726858	0.00461	< 2e-16
diet score 2	0.612431	0.105275	0.010795	< 2e-16
diet score 3	0.067903	0.106172	0.010786	< 2e-16

For PA only the following variables had sufficient data: g1_a, g1_b, g1_c, g1_d, g3_a, g3_b, g6. The variables g1_N are related since they are seasonal variables and we were discussing to combine them, so I had checked several different combinations:

1. Separate variables.
2. Combining all g1_N
3. Combining only the significant g1_N

Visit 1:

1.Separate variables:

	Regression coefficient, unstandardized	Regression coefficient, standardized	p-value
g1_a	-0.026707	-0.035912	0.01397
g1_b	-0.001490	-0.002117	0.85631
g1_c	-0.006834	-0.009016	0.52182
g1_d	-0.028987	-0.031211	0.00242
g3_a	-0.005845	-0.006559	0.31450
g3_b	-0.004055	-0.005348	0.46532
g6	-0.054040	-0.063570	< 2e-16

2.Combining all g1_N:

	Regression coefficient, unstandardized	Regression coefficient, standardized	p-value
g1_abcd	-0.014986	-0.070790	< 2e-16
g3_a	-0.005632	-0.006320	0.32869
g3_b	-0.003804	-0.005016	0.48647
g6	-0.053058	-0.062415	< 2e-16

3.Combining only the significant g1_N:

	Regression coefficient, unstandardized	Regression coefficient, standardized	p-value
g1_ad	-0.031370	-0.071011	< 2e-16
g3_a	-0.005596	-0.006279	0.33173
g3_b	-0.005429	-0.007159	0.31257
g6	-0.053258	-0.062651	< 2e-16

Visit 2:

1.Separate variables:

	Regression coefficient, unstandardized	Regression coefficient, standardized	p-value
g1_a	-0.0253117	-0.0332984	0.02985
g1_b	0.0006195	0.0008662	0.94453
g1_c	-0.0009063	-0.0011778	0.93659
g1_d	-0.0666347	-0.0716823	3.49e-11
g3_a	-0.0402458	-0.0420984	3.26e-10
g3_b	-0.0169811	-0.0228948	0.00373
g6	-0.0680579	-0.0912764	< 2e-16

2.Combining all g1_N:

	Regression coefficient, unstandardized	Regression coefficient, standardized	p-value
g1_abcd	-0.020930	-0.097848	< 2e-16
g3_a	-0.042628	-0.044591	2.12e-11
g3_b	-0.013237	-0.017847	0.021643
g6	-0.068313	-0.091619	< 2e-16

3.Combining only the significant g1_N:

	Regression coefficient, unstandardized	Regression coefficient, standardized	p-value
g1_ad	-0.044993	-0.100966	< 2e-16
g3_a	-0.042335	-0.044284	2.83e-11
g3_b	-0.014702	-0.019822	0.00904
g6	-0.068315	-0.091621	< 2e-16

Variance explained calculated separately and adjusting only for the basic variables, since PA variables have missing values in mixed places.

Visit 1:

	Variance explained	p-value
g1_a	0.006326	< 2e-16
g1_b	0.005051	< 2e-16
g1_c	0.005886	< 2e-16
g1_d	0.005701	< 2e-16
g3_a	0.000649	5.212801e-06
g3_b	0.002416	< 2e-16
g6	0.005676	< 2e-16
g1_abcd	0.006849	< 2e-16
g1_ad	0.006977	< 2e-16

Visit 2:

	Variance explained	p-value
g1_a	0.012533	< 2e-16
g1_b	0.009751	< 2e-16
g1_c	0.01193	< 2e-16
g1_d	0.013767	< 2e-16
g3_a	0.006293	< 2e-16
g3_b	0.010456	< 2e-16
g6	0.012445	< 2e-16
g1_abcd	0.014299	< 2e-16
g1_ad	0.01496	< 2e-16

Probably best to keep just g1_a, g1_d and g6 for the final PA score, out of all combinations, the beta coefficient and r² were the biggest:

Visit 1:

	Regression coefficient, unstandardized	Regression coefficient, standardized	Variance explained	p-value
PA score	-0.038565	-0.102699	0.010857	< 2e-16

Visit 2:

	Regression coefficient, unstandardized	Regression coefficient, standardized	Variance explained	p-value
PA score	-0.0559818	-0.1538940	0.023603	< 2e-16

Combining diet score and PA(PA reverse scoring):

Separate:

diet score 1
Visit 1:

	Regression coefficient, standardized	p-value
diet score 1	0.002307	0.71732
PA score	0.103639	< 2e-16

Visit 2:

	Regression coefficient, standardized	p-value
diet score 1	0.057348	< 2e-16
PA score	0.147512	< 2e-16

diet score 2
Visit 1:

	Regression coefficient, standardized	p-value
diet score 2	0.072328	< 2e-16
PA score	0.104734	< 2e-16

Visit 2:

	Regression coefficient, standardized	p-value
diet score 2	0.101192	< 2e-16
PA score	0.151720	< 2e-16

diet score 3
Visit 1:

	Regression coefficient, standardized	p-value
diet score 3	0.048991	< 2e-16
PA score	0.100540	< 2e-16

Visit 2:

	Regression coefficient, standardized	p-value
diet score 3	0.092354	< 2e-16
PA score	0.145231	< 2e-16

Summed:

diet score 1
Visit 1:

	Regression coefficient, standardized	Variance explained	p-value
diet score 1 + PA score	0.067369	0.004034	< 2e-16

Visit 2:

	Regression coefficient, standardized	Variance explained	p-value
diet score 1 + PA score	0.1213241	0.012926	< 2e-16

diet score 2
Visit 1:

	Regression coefficient, standardized	Variance explained	p-value
diet score 2 + PA score	0.076094	0.005849	< 2e-16

Visit 2:

	Regression coefficient, standardized	Variance explained	p-value
diet score 2 + PA score	0.102899	0.010333	< 2e-16

diet score 3

Visit 1:

	Regression coefficient, standardized	Variance explained	p-value
diet score 3 + PA score	0.088926	0.007851	< 2e-16

Visit 2:

	Regression coefficient, standardized	Variance explained	p-value
diet score 3 + PA score	0.1319158	0.016556	< 2e-16