Associations between dietary index for gut microbiota and stroke, and the mediating role of inflammation: a prospective cohort study

Contents

Supplementary Table 1. Characteristics of participants in dietary assessments in the UK Biobank over time.

Time	Characteristics	Respondent	Non-respondent
Baseline	N	70,676	
Conducted in assessment center	Age (mean (SD))	56.21 (8.16)	
April 2009 to September 2010	Sex, Male (%)	31,616 (44.73)	
	Ethnic,White (%)	65,954 (93.32)	
	College degree or professional qualification (%)	30,768 (43.83)	
	TDI (mean (SD))	-1.27 (2.84)	
	Previous or current smoker (%)	30,623 (43.46)	
	Previous or current drinker (%)	67,894 (96.17)	
	BMI, overweight (%)	45,129 (64.95)	
	Hypertension (%)	48,993 (70.18)	
First Dietary Follow-up	N	100,553	205,758
February 2011 to April 2011	Age (mean (SD))	56.10 (7.83)	55.41 (8.11)
Response rate: 32.83%	Sex, Male (%)	44,954 (44.7)	97,591 (47.4)
	Ethnic,White (%)	97,158 (96.6)	192,190 (93.4)
	College degree or professional qualification (%)	51,297 (51.2)	85,076 (41.7)
	TDI (mean (SD))	-1.67 (2.84)	-1.52 (2.94)
	Previous or current smoker (%)	43,232 (43.1)	91,720 (44.8)
	Previous or current drinker (%)	97,535 (97.1)	197,817 (96.4)
	BMI, overweight (%)	60,933 (61.5)	137,645 (68.1)
	Hypertension (%)	67,822 (71.1)	139,028 (70.8)
Second Dietary Follow-up	N	83,222	234,994
June 2011 to September 2011	Age (mean (SD))	56.24 (7.77)	55.51 (8.11)
Response rate: 26.15%	Sex, Male (%)	36,882 (44.3)	110,683 (47.1)
	Ethnic,White (%)	80,498 (96.7)	220,416 (93.8)
	College degree or professional qualification (%)	42,565 (51.3)	98,682 (42.4)
	TDI (mean (SD))	-1.65 (2.86)	-1.55 (2.93)
	Previous or current smoker (%)	35,799 (43.1)	104,291 (44.6)
	Previous or current drinker (%)	80,732 (97.1)	226,115 (96.4)
	BMI, overweight (%)	50,705 (61.8)	155,300 (67.3)
	Hypertension (%)	56,028 (71.1)	158,345 (70.9)
Third Dietary Follow-up	N	103,736	214,665
October 2011 to December 2011	Age (mean (SD))	56.17 (7.81)	55.48 (8.13)
Response rate: 32.58%	Sex, Male (%)	46,265 (44.6)	101,365 (47.2)
	Ethnic,White (%)	100,172 (96.6)	200,924 (93.6)
	College degree or professional qualification (%)	53,685 (51.9)	87,647 (41.2)
	TDI (mean (SD))	-1.66 (2.85)	-1.53 (2.94)
	Previous or current smoker (%)	44,095 (42.6)	96,077 (44.9)

	Previous or current drinker (%)	100,661 (97.1)	206,366 (96.3)
	BMI, overweight (%)	62,005 (60.7)	144,092 (68.3)
	Hypertension (%)	69,331 (70.5)	145,159 (71.2)
Fourth Dietary Follow-up	N	100,197	228,896
April 2012 to June 2012	Age (mean (SD))	56.26 (7.77)	55.53 (8.12)
Response rate: 30.45%	Sex, Male (%)	44,100 (44.0)	108,417 (47.4)
	Ethnic,White (%)	96,792 (96.6)	214,713 (93.8)
	College degree or professional qualification (%)	52,230 (52.3)	93,247 (41.1)
	TDI (mean (SD))	-1.65 (2.86)	-1.55 (2.93)
	Previous or current smoker (%)	42,660 (42.7)	102,255 (44.9)
	Previous or current drinker (%)	97,096 (97.0)	220,232 (96.4)
	BMI, overweight (%)	59,635 (60.4)	153,433 (68.2)
	Hypertension (%)	66,792 (70.5)	154,616 (71.3)

Abbreviations: N, number of participants; TDI, Townsend deprivation index; BMI, body mass index; SD, standard deviation.

Supplementary Table 2. The components and field IDs of DI-GM in the UK Biobank.

Components of DI-GM	Food items included in UK Biobank		Field ID
Beneficial to gut microbiota	Avocados[#]	—	104100
	Broccoli[#]	—	104140
	Green tea[#]	—	100420
	Soybean[#]	Meat substitutes - soy	26137
		Soy desserts and yogurt	26086
		Soy milk	26136
	Fiber[†]	—	26017
	Coffee[†]	Coffee, caffeinated	26081
		Coffee, decaffeinated	26082
	Fermented dairy[†]	Full fat yogurt	26096
		Low fat yogurt	26102
		High fat cheese	26099
		Medium and low fat cheese	26103
	Whole grains[†]	Biscuit cereal	26075
		Bran cereal	26076
		Muesli	26105
		Oat cereal (non sugar)	26077
		Oat cereal (sugar)	26078
		Wholemeal bread	26074
		Wholemeal pasta, brown rice and other wholegrains	26114
Unfavorable to gut microbiota	Processed meat*	—	26122
	Red meat*	Beef	26066
		Lamb	26100
		Pork	26117
	Refined grains*	Biscuits	26068
		Pizza	26116
		Samosa, pakora	26128
		Savoury crackers	26083
		White bread	26073
		White pasta and rice	26113
		Grain dishes - added fat	26097
		Other bread	26072
		Other cereal (sugar)	26079
		Savoury snacks	26134
		Other desserts and cakes and pastries	26085
	High-fat diet (% energy)[&]		
		Total energy	26002
		Total fat	26008

[#]Score 1: Consumption > sex-specific median, score 0: Otherwise;

[†]Score 1: Consumption ≥ sex-specific median, score 0: Otherwise;

*Score 0: Consumption ≥ sex-specific median, score 1: Otherwise;

[&]Score 0: Consumption ≥ 40%, score 1: Otherwise.

Abbreviation: DI-GM, dietary index for gut microbiota.

Since the sex-specific median for avocados, broccoli, green tea and soybean in the UKB was 0, we considered it illogical to assign a score of 1 to all participants for the consumption of these foods. Therefore, for these items, participants with consumption exceeding the sex-specific median were assigned a score of 1; otherwise, the score was 0.

Supplementary Table 3. Definitions of inflammatory biomarkers and INFLA-score.

Item	Abbreviation	Unit	Field ID
Inflammatory biomarkers			
C-reactive protein	CRP	mg/L	30710
Neutrophil count	NEUT	10^9 /L	30140
Monocyte count	MONO	10^9 /L	30130
White blood cell (leukocyte) count	WBC	10^9 /L	30000
Platelet count	PLT	10^9 /L	30080
Lymphocyte count	LYMPH	10^9 /L	30120
Neutrophil-to-lymphocyte ratio	NLR	/	30140 and 30120
INFLA-score			
C-reactive protein	CRP	mg/L	30710
White blood cell (leukocyte) count	WBC	10^9 /L	30000
Platelet count	PLT	10^9 /L	30080
Neutrophil-to-lymphocyte ratio	NLR	/	30140 and 30120

Abbreviation: INFLA, low-grade chronic inflammation score.

Supplementary Table 4. The numbers (percentages) of participants with missing covariates in the analysis.

Covariates	N	%
Education	387	0.31
Townsend deprivation index (TDI)	147	0.12
Smoking status	266	0.21
Alcohol status	89	0.07
Body mass index (BMI)	268	0.21
Systolic Blood Pressure	5,656	4.53
Diastolic Blood Pressure	5,652	4.52
Family history of stroke	1,993	1.60
Physical activity	25,996	20.81

Abbreviations: N, number of participants.

Supplementary Table 5. The association of DI-GM with stroke risk after including participants with only one diet questionnaire (N = 207,666).

Predictor	Stroke		Ischemic stroke		Hemorrhage stroke	
	HR [95% CI]	p-value	HR [95% CI]	p-value	HR [95% CI]	p-value
DI-GM	0.97 [0.95-0.98]	8.32E-06	0.96 [0.94-0.98]	6.75E-05	0.98 [0.95-1.02]	0.409
DI-GM group						
0-3	Reference		Reference		Reference	
4	0.92 [0.86-1.00]	0.045	0.98 [0.88-1.10]	0.780	0.76 [0.61-0.95]	0.014
5	0.89 [0.83-0.96]	0.003	0.88 [0.78-0.98]	0.026	0.89 [0.73-1.09]	0.259
⩾6	0.86 [0.80-0.92]	1.12E-05	0.84 [0.76-0.93]	8.85E-04	0.94 [0.78-1.12]	0.465
Trend test		1.22E-05		1.22E-04		0.888
BGMS	0.93 [0.91-0.95]	8.11E-15	0.91 [0.89-0.94]	5.00E-10	0.96 [0.91-1.00]	0.072
UGMS	1.01 [0.97-1.05]	0.626	1.03 [0.99-1.08]	0.112	1.04 [0.97-1.12]	0.279

Multivariable Cox proportional hazard model was used to examine the associations of DI-GM with stroke risk, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval; N, number of participants.

Supplementary Table 6. The association of DI-GM with stroke risk after excluding participants with unreasonable energy intake and atypical diets (N = 102,005).

Predictor	Stroke		Ischemic stroke		Hemorrhage stroke	
	HR [95% CI]	p-value	HR [95% CI]	p-value	HR [95% CI]	p-value
DI-GM	0.97 [0.95-0.99]	0.006	0.96 [0.93-0.99]	0.016	0.97 [0.92-1.02]	0.204
DI-GM group						
0-3	Reference		Reference		Reference	
4	0.89 [0.80-1.00]	0.056	0.94 [0.79-1.12]	0.501	0.83 [0.61-1.13]	0.241
5	0.88 [0.79-0.99]	0.026	0.88 [0.75-1.05]	0.153	0.88 [0.65-1.18]	0.377
⩾6	0.85 [0.77-0.94]	1.02E-03	0.84 [0.73-0.98]	0.026	0.87 [0.67-1.13]	0.302
Trend test		0.002		0.018		0.454
BGMS	0.94 [0.91-0.97]	6.27E-06	0.93 [0.89-0.97]	3.30E-04	0.93 [0.86-1.00]	0.035
UGMS	1.03 [0.99-1.07]	0.162	1.02 [0.96-1.09]	0.481	1.04 [0.93-1.16]	0.494

Multivariable Cox proportional hazard model was used to examine the associations of DI-GM with stroke risk, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval; N, number of participants.

Supplementary Table 7. The association of DI-GM with stroke risk after excluding participants who experienced a stroke within the first two years of follow-up (N = 124,515).

Predictor	Stroke		Ischemic stroke		Hemorrhage stroke	
	HR [95% CI]	p-value	HR [95% CI]	p-value	HR [95% CI]	p-value
DI-GM	0.98 [0.96-1.00]	0.014	0.96 [0.93-0.99]	0.018	0.98 [0.93-1.04]	0.524
DI-GM group						
0-3	Reference		Reference		Reference	
4	0.88 [0.79-0.99]	0.029	1.02 [0.86-1.21]	0.800	0.87 [0.64-1.18]	0.370
5	0.86 [0.77-0.96]	0.008	0.91 [0.77-1.08]	0.263	0.92 [0.69-1.24]	0.584
⩾6	0.88 [0.80-0.97]	0.009	0.90 [0.77-1.04]	0.148	0.95 [0.73-1.23]	0.694
Trend test		0.022		0.065		0.906
BGMS	0.95 [0.92-0.97]	5.22E-05	0.93 [0.90-0.97]	8.64E-04	0.95 [0.88-1.02]	0.122
UGMS	1.03 [0.99-1.07]	0.177	1.01 [0.95-1.08]	0.701	1.06 [0.96-1.18]	0.261

Multivariable Cox proportional hazard model was used to examine the associations of DI-GM with stroke risk, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval; N, number of participants.

Supplementary Table 8. The association of DI-GM with stroke risk after excluding participants with AF, atherosclerosis, dementia, and depression at baseline (N = 112,442).

Predictor	Stroke		Ischemic stroke		Hemorrhage stroke	
	HR [95% CI]	p-value	HR [95% CI]	p-value	HR [95% CI]	p-value
DI-GM	0.97 [0.95-0.99]	7.33E-03	0.97 [0.94-1.00]	0.027	0.96 [0.91-1.01]	0.142
DI-GM group						
0-3	Reference		Reference		Reference	
4	0.85 [0.76-0.95]	0.005	0.97 [0.82-1.15]	0.700	0.76 [0.56-1.03]	0.076
5	0.85 [0.76-0.95]	4.66E-03	0.88 [0.74-1.04]	0.143	0.80 [0.60-1.07]	0.139
⩾6	0.85 [0.77-0.93]	8.90E-04	0.87 [0.75-1.01]	0.075	0.82 [0.64-1.06]	0.124
Trend test		4.76E-03		0.048		0.254
BGMS	0.95 [0.93-0.98]	1.62E-04	0.94 [0.91-0.98]	4.79E-03	0.93 [0.87-1.00]	0.043
UGMS	1.01 [0.97-1.05]	0.656	1.00 [0.94-1.06]	0.930	1.01 [0.91-1.13]	0.855

Multivariable Cox proportional hazard model was used to examine the associations of DI-GM with stroke risk, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score; AF, atrial fibrillation; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval; N, number of participants.

Supplementary Table 9. The association of DI-GM with stroke risk after performing multiple imputation for missing covariates (N = 124,943).

Predictor	Stroke		Ischemic stroke		Hemorrhage stroke	
	HR [95% CI]	p-value	HR [95% CI]	p-value	HR [95% CI]	p-value
DI-GM	0.97 [0.95-0.99]	4.01E-04	0.96 [0.94-0.99]	0.010	0.96 [0.91-1.00]	0.076
DI-GM group						
0-3	Reference		Reference		Reference	
4	0.86 [0.78-0.95]	0.004	0.97 [0.83-1.13]	0.680	0.75 [0.57-0.99]	0.042
5	0.85 [0.77-0.94]	1.04E-03	0.89 [0.76-1.03]	0.120	0.84 [0.65-1.08]	0.176
⩾6	0.84 [0.77-0.92]	9.75E-05	0.87 [0.76-0.99]	0.038	0.82 [0.66-1.03]	0.094
Trend test		4.69E-04		0.021		0.239
BGMS	0.94 [0.92-0.96]	9.18E-07	0.94 [0.90-0.97]	4.89E-04	0.92 [0.87-0.98]	0.011
UGMS	1.01 [0.97-1.05]	0.626	1.01 [0.95-1.06]	0.833	1.02 [0.92-1.12]	0.730

Multivariable Cox proportional hazard model was used to examine the associations of DI-GM with stroke risk, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes, family history of stroke, atrial fibrillation, atherosclerosis, dementia, depression and physical activity.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval; N, number of participants.

Supplementary Table 10. The association between DI-GM and inflammatory biomarkers.

Name	Beta	p-value	FDR
C-reactive protein	-0.040	4.33E-141	1.73E-140
Neutrophil count	-0.030	7.28E-67	1.46E-66
Monocyte count	-0.019	1.65E-30	2.43E-30
White blood cell (leukocyte) count	-0.032	1.45E-75	3.87E-75
Platelet count	-0.020	1.82E-30	2.43E-30
Lymphocyte count	-0.015	3.70E-18	4.23E-18
Neutrophil-to-lymphocyte ratio	-0.013	1.45E-13	1.45E-13
INFLA-score	-0.258	2.04E-145	1.63E-144

Multiple linear regression was used to examine the associations of baseline DI-GM with baseline inflammatory biomarkers, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke.

Abbreviations: DI-GM, dietary index for gut microbiota; INFLA, low-grade chronic inflammation score; TDI, Townsend Deprivation Index; BMI, body mass index; FDR, false discovery rate.

Supplementary Table 11. Effect of baseline DI-GM on longitudinal changes in inflammatory biomarkers.

Name	Beta	p-value	FDR
C-reactive protein	-0.025	1.32E-05	2.89E-05
Neutrophil count	-0.027	3.40E-06	1.36E-05
Monocyte count	-0.016	0.005	0.006
White blood cell (leukocyte) count	-0.031	9.33E-08	7.46E-07
Platelet count	-0.011	0.057	0.065
Lymphocyte count	-0.019	1.18E-03	0.002
Neutrophil-to-lymphocyte ratio	-0.006	0.265	0.265
INFLA-score	-0.025	1.44E-05	2.89E-05

Multiple linear regression was used to examine the associations of baseline DI-GM with the rate of longitudinal changes in inflammatory biomarkers, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes, and family history of stroke. The rate of longitudinal changes in inflammatory biomarkers was calculated using linear mixed-effects models.

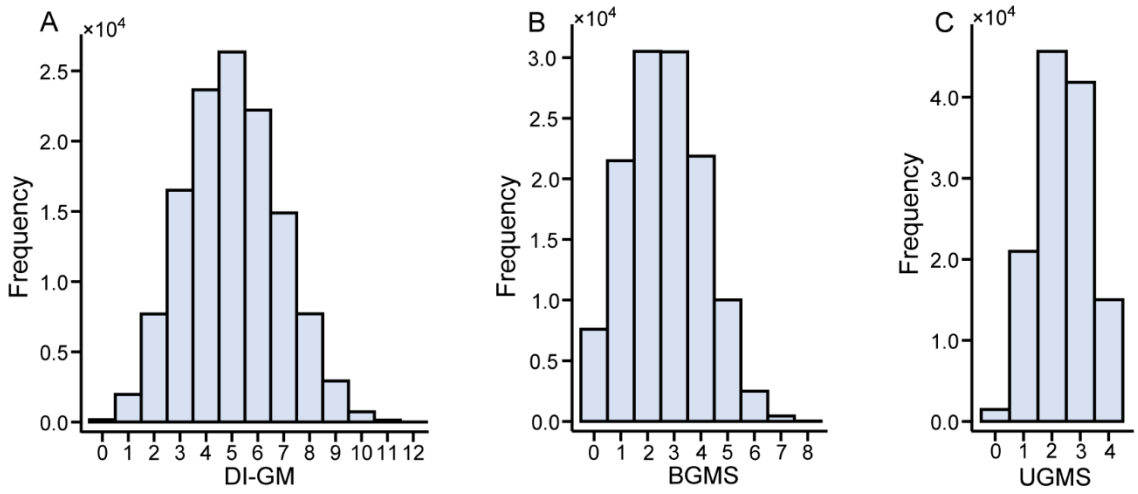
Abbreviations: DI-GM, dietary index for gut microbiota; INFLA, low-grade chronic inflammation score; TDI, Townsend Deprivation Index; BMI, body mass index; FDR, false discovery rate.

Supplementary Table 12. The association of inflammatory biomarkers with stroke risk.

Name	HR	95% CI	p-value	FDR
C-reactive protein	1.08	1.04-1.12	0.0001	0.0010
Neutrophil count	1.06	1.02-1.10	0.0052	0.0084
Monocyte count	1.07	1.03-1.11	0.0004	0.0017
White blood cell (leukocyte) count	1.05	1.01-1.09	0.0173	0.0231
Platelet count	0.97	0.94-1.01	0.1499	0.1713
Lymphocyte count	0.98	0.95-1.02	0.3264	0.3264
Neutrophil-to-lymphocyte ratio	1.06	1.03-1.10	0.0008	0.0017
INFLA-score	1.01	1.01-1.02	0.0007	0.0017

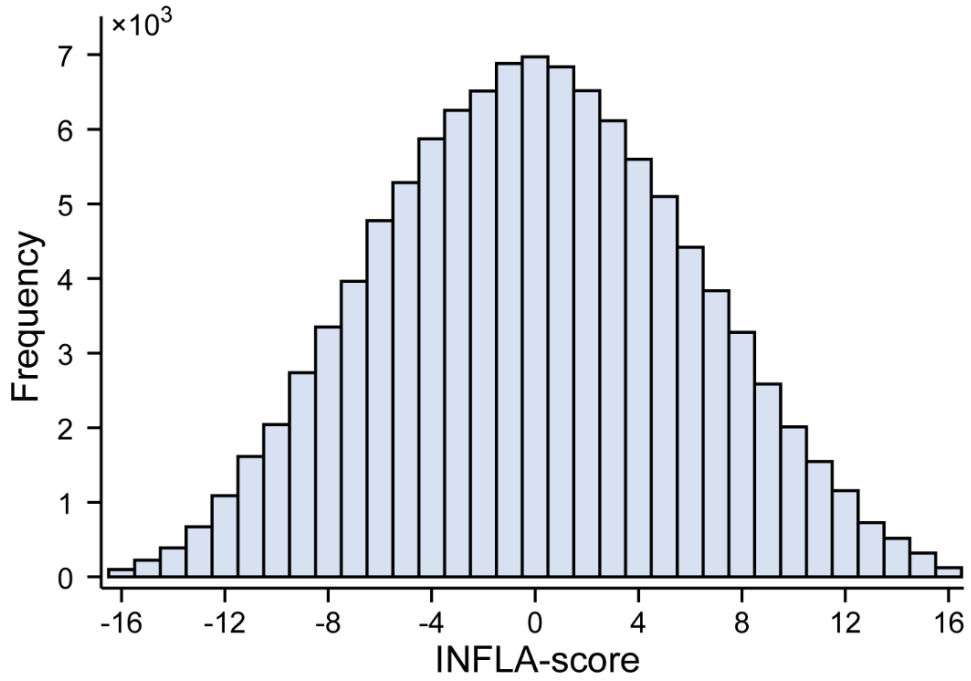
Multivariable Cox proportional hazard model was used to examine the associations of inflammatory biomarkers with stroke risk, adjusting for potential confounders including age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke.

Abbreviations: INFLA, low-grade chronic inflammation score; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval; FDR, false discovery rate.



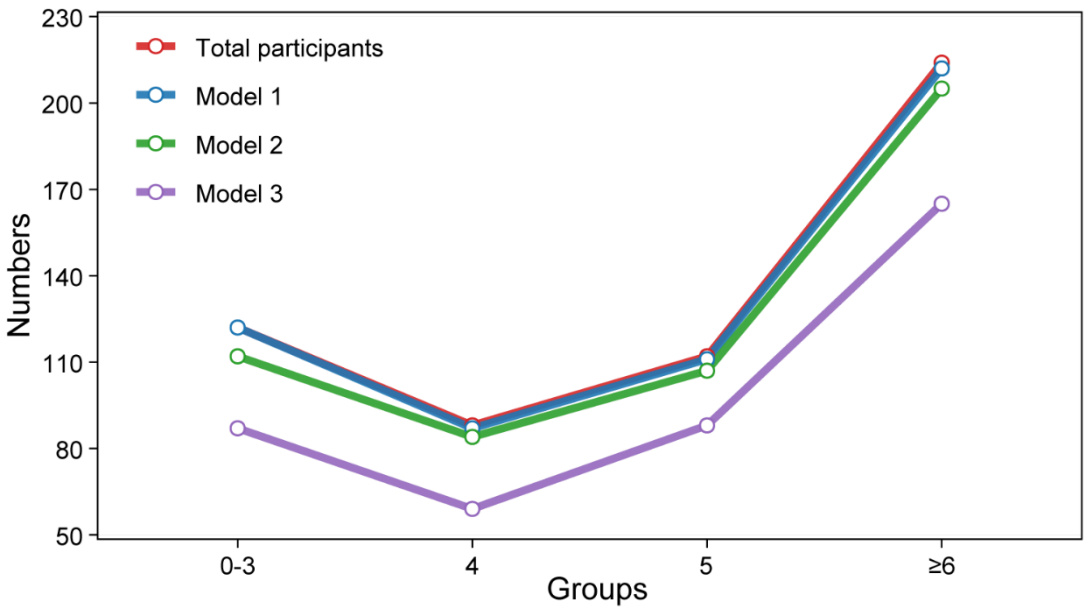
Supplementary Figure 1. Data distributions of DI-GM, BGMS and UGMS from 124,943 participants.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score.



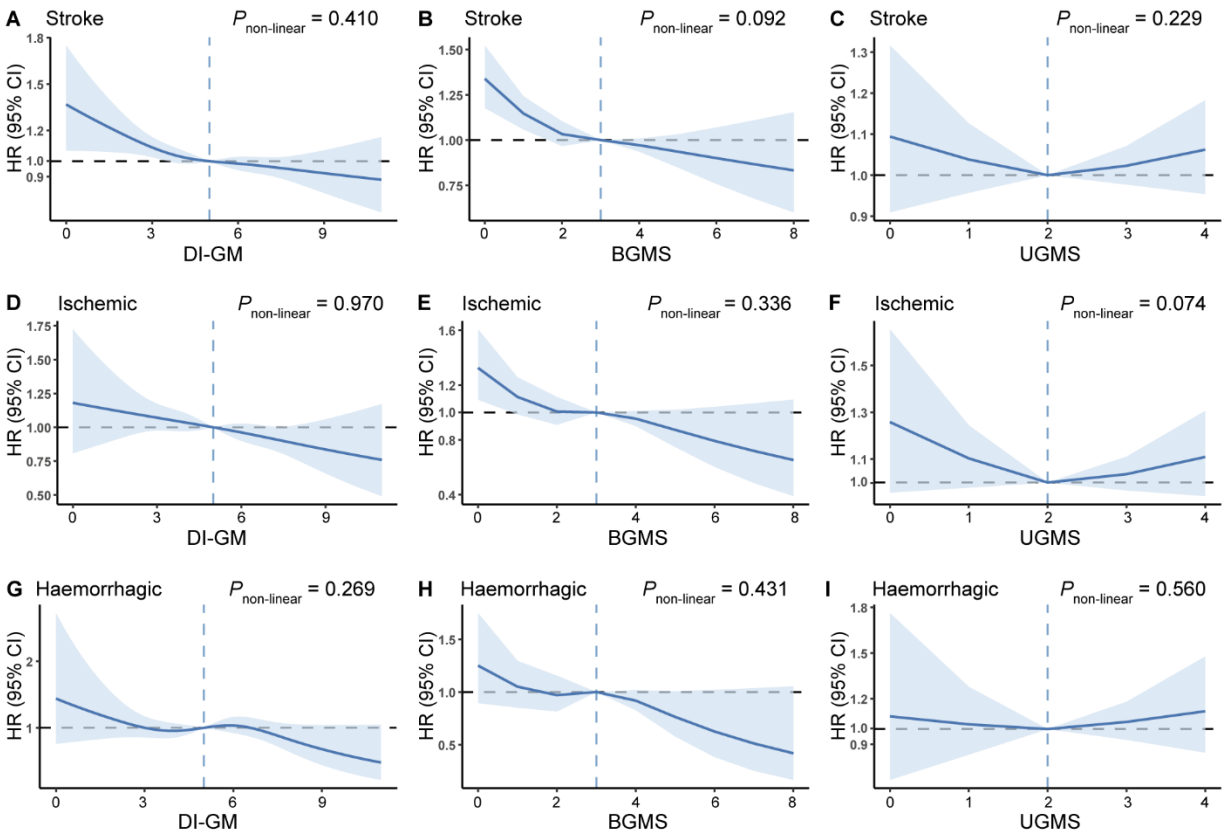
Supplementary Figure 2. Data distributions of the INFLA-score from 109,422 participants.

Abbreviations: INFLA, low-grade chronic inflammation score.



Supplementary Figure 3. Number of hemorrhagic stroke events by DI-GM score in the total population and after exclusion of missing covariates in three models.

Abbreviations: DI-GM, dietary index for gut microbiota.



Supplementary Figure 4. Non-linear associations between DI-GM and the risk of stroke and its subtypes.

Two-sided Cox regression models were adjusted for age, sex, ethnicity, TDI, education level, smoking status, alcohol consumption, energy intake, BMI, blood pressure, diabetes and family history of stroke. Solid lines (center of the error bands) indicate HRs, and shading indicate 95% CIs. Blue dashed line indicates the reference point.

Abbreviations: DI-GM, dietary index for gut microbiota; BGMS, beneficial to gut microbiota score; UGMS, unfavorable to gut microbiota score; TDI, Townsend Deprivation Index; BMI, body mass index; HR, hazard ratio; CI, confidence interval.