Supplementary materials

The dietary index for gut microbiota, genetically predicted gut microbiome and risk of chronic kidney disease: a cohort study

Zheng Ma [1, 2]*, Yixiao Zhang [3]*, Honghao Yang [1, 2], Gang Zheng [1, 2], Liuxin Li [1, 2], Ruiyao Zhang[1, 2], Xiang Ji [1, 2], Chao Ji [1, 2], Yang Xia [1, 2, 4#], Yuhong Zhao [1, 2#]

[#] Correspondence to: Yang Xia or Yuhong Zhao.

Supplemental results

Supplementary Table 9. Associations between the DI-GM and the risk of CKD when participants with at least 2 times 24-hour dietary recall questionnaire with typical diet and credible energy (n=89,896)

Supplementary Table 10. Associations between the DI-GM and the risk of CKD further adjusted for eGFR (n=166,364)

Supplementary Table 11. Associations between the DI-GM and the risk of CKD using competing risk regression (n=166,364)

Supplementary Figure 3. The dose-response association between the DI-GM and risk of CKD diabetes using restricted cubic spline with three knots

Supplementary Figure 4. The dose-response association between the DI-GM and risk of CKD diabetes using restricted cubic spline with four knots

Supplementary Table 12. Associations between the DI-GM and the risk of CKD including participants lost to follow-up (n=166,821)

Supplementary Table 13. Associations between the DI-GM and the risk of CKD further adjusted for history of antibiotic use, history of bowel resection surgery, and IBD (n=166,364)

Supplementary Table 14. Associations between the DI-GM and the risk of CKD excluding participants who history of antibiotic use, history of bowel resection surgery, and IBD (n=151,160).

Supplementary Figure 5. Associations of the DI-GM with the risk of CKD according to genetically predicted gut microbial abundance, additionally adjusted for antibiotic use, bowel resection surgery, and IBD history

Supplement Table 1. The assessment for CKD in the UK Biobank cohort [a].

Diagnosed diseases	UK Biobank algorithm [b]	ICD-10 codes	Self-reported codes [c]	OPCS-4 codes	Others
At baseline					
CKD	131290 (Date I12 first reported) 131292 (Date I13 first reported) 132032 (Date N18 first reported)	E10.2, E11.2, T86.1, Z94.0	1192, 1193, 1194, 1196	L74, M01, M02.3, M08.4, M17, X40, X41	eGFR<60 mL/min/1.73 m² or UACR >30 mg/g
During the follow-up					
CKD	131290 (Date I12 first reported) 131292 (Date I13 first reported) 132032 (Date N18 first reported)	E10.2, E11.2, T86.1, Z94.0	-	L74, M01, M02.3, M08.4, M17, X40, X41	-

[a] Abbreviations: CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; ICD-10, International Classification of Disease, 10th version; OPCS-4, Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4; UACR, urinary albumin-to-creatinine ratio.

[b] "Date of disease (shown as ICD-10 codes) first reported" captures data from multiple sources, including primary care records, hospital inpatient records, self-reported medical conditions, and death registers. This algorithm is implemented by the UK Biobank team. The numbers in the table represent UKB field IDs (e.g., 131290).

[c] Self-reported codes for non-cancer illness (field ID: 20002).

Supplementary Table 2. Components and scoring criteria of DI-GM in UK Biobank.

Components of DI-GM	Food items included in UK Biobank	Scoring criteria
Beneficial to gut microbiota	Avocados Broccoli Pulses (including kidney beans/Chickpeas/butter beans etc.) Coffee Dried fruit (including e.g., raisins, apple rings, Cranberries (not with breakfast cereal) Fermented dairy Fiber green tea Soybean Whole grains	Score 1 - Consumption > sex-specific median Score 0 - Otherwise
Unfavorable to gut microbiota	Refined grains Processed meat Red meat High-fat diet (% energy)	Score 0 - Consumption ≤ sex-specific median Score 1 - Otherwise Score 0 - Consumption≥40%

Supplementary Table 3. Characteristics of the 19 SNPs associated with gut microbiome abundance [a].

SNP	Effect Allele	Chr	Beta	Phylum
rs10769159	G	11	-0.06403	Firmicutes
rs10805326	G	4	0.077515	Firmicutes
rs11098863	T	4	-0.0966	Actinobacteria
rs11110281	T	12	-0.13404	Firmicutes
rs12320842	C	12	0.094833	Firmicutes
rs12781711	C	10	-0.06561	Firmicutes
rs17159861	C	7	0.096223	Firmicutes
rs182549	C	2	0.116333	Actinobacteria
rs4428215	G	3	0.12561	Proteobacteria
rs602075	A	9	0.168974	Firmicutes
rs61841503	G	10	0.092427	Firmicutes
rs67476743	T	19	0.132164	Firmicutes
rs7221249	A	17	0.083986	Firmicutes
rs7322849	T	13	0.11126	Actinobacteria
rs736744	C	9	0.117882	Proteobacteria
rs75754569	C	3	0.181434	Firmicutes
rs8009993	G	14	-0.13594	Firmicutes
rs830151	G	19	0.194561	Firmicutes
rs9864379	A	3	-0.16052	Cyanobacteria

Abbreviations: SNP, single nucleotide polymorphism; Chr, chromosome.

Supplementary Table 4. SNPs used to construct the eGFR-GRS in the UK Biobank study [a].

SNP	Effect Allele	Chr	Effect	Locus
rs61830291	A	1	-0.0036	LINC01352
rs2490391	A	1	-0.0024	SDCCAG8
rs12061708	A	1	-0.0026	KLHDC7A
rs2749153	A	1	-0.0033	ZNF436-AS1
rs688540	A	1	-0.003	FOXD2
rs3845534	A	1	-0.0019	LOC100422212
rs1011731	A	1	-0.0019	DNM3
rs78444298	A	1	-0.0105	EDEM3
rs78329830	A	1	-0.0054	PLA2G4A
rs1887252	C	1	-0.0019	LINC01362
rs7543734	C	1	0.0031	BCAR3
rs74748843	T	1	-0.0048	CASZ1
rs659437	T	1	-0.0027	AKR1A1
rs10159261	T	1	-0.0034	AGMAT
rs267738	T	1	-0.0048	CERS2
rs17413465	A	1	0.0025	MIR4422HG
rs11211257	A	1	0.0027	PIK3R3
rs1757915	A	1	0.0021	LINC01755
rs11166440	A	1	0.002	CDC14A
rs10857788	A	1	0.003	SYPL2
rs4971100	A	1	0.002	TRIM46
rs3850625	A	1	0.0046	CACNA1S
rs2808454	A	1	0.0019	PFKFB2
rs12736457	C	1	0.0054	PPM1J
rs75625374	C	1	0.0045	CD34
rs7536433	T	1	0.0021	AK5
rs679843	T	1	0.0021	MGC27382
rs3118119	T	1	0.003	LOC105371433
rs4656220	T	1	0.002	PRRX1
rs3795503	T	1	0.002	KIAA1614
rs7535253	T	1	0.0021	PTPN14
rs2577134	T	1	0.002	RNU5F-1
rs417237	T	1	0.0018	OBSCN
rs1047891	A	2	-0.0065	CPS1
rs17050272	A	2	-0.0022	LINC01101
rs2971880	A	2	-0.0024	SPTBN1
rs60980181	A	2	-0.0027	CALCRL

rs4664475	T	2	-0.002	NEB
rs35472707	T	2	-0.0073	LRP2
rs2301343	T	2	-0.0023	SLC8A1
rs35284526	A	2	0.0029	NFE2L2
rs4491726	A	2	0.0032	RDH14
rs6546869	A	2	0.0059	ALMS1P1
rs11694902	A	2	0.0041	TFCP2L1
rs7425436	A	2	0.0024	ORC4
rs35669853	A	2	0.0024	MIR5702
rs10197255	A	2	0.0018	LINC01812
rs10865189	C	2	0.0024	ZFP36L2
rs187355703	C	2	0.01	HOXD8
rs780093	T	2	0.0044	GCKR
rs11123169	T	2	0.0025	PSD4
rs1548945	T	2	0.0036	TNP1
rs1050816	T	2	0.0026	SPEG
rs13003198	T	2	0.0018	SAG
rs807624	T	2	0.0032	DDX1
rs4666821	T	2	0.002	PDE1A
rs3791221	A	2	0.0022	SH3YL1
rs6779998	A	3	-0.0017	TGFBR2
rs4625	A	3	-0.0023	DAG1
rs9828976	C	3	-0.0024	SLC35G2
rs56065557	C	3	-0.0029	SENP2
rs6778731	T	3	-0.0017	WNT7A
rs3774726	T	3	-0.0021	ATXN7
rs2289746	T	3	-0.0019	CBLB
rs35320690	T	3	-0.0025	MSL2
rs11919484	T	3	-0.0026	KNG1
rs2581820	A	3	0.0021	SFMBT1
rs9868185	A	3	0.0026	SLC15A2
rs1397764	A	3	0.0043	TFDP2
rs9823161	A	3	0.0022	LINC02028
rs11914389	T	3	0.003	ACVR2B
rs7651407	T	3	0.0025	PLXNB1
rs10934754	T	3	0.002	ALDH1L1-AS2
rs7624084	T	3	0.0017	ZBTB38
rs76272256	T	3	0.0024	MECOM
rs795009	T	3	0.002	SYN2
rs3775932	A	4	-0.0018	WDR1

rs16874073	T	4	-0.0045	PPARGC1A
rs4864890	T	4	-0.0023	DCUN1D4
rs12509595	T	4	-0.0035	FGF5
rs75501914	A	4	0.0039	HGFAC
rs71606723	A	4	0.0025	UGT8
rs223471	C	4	0.0028	LOC102723704
rs55929207	C	4	0.0019	ETNPPL
rs28817415	T	4	-0.0073	SHROOM3
rs12163971	A	5	-0.0029	AFF4
rs13159523	A	5	-0.0024	TPPP
rs13157326	A	5	-0.0027	RAI14
rs2010352	A	5	-0.0018	AK6
rs1362800	T	5	-0.0049	DAB2
rs3812036	T	5	-0.0065	SLC34A1
rs72759880	T	5	-0.0056	PIK3R1
rs3797537	A	5	0.0019	DMGDH
rs12520984	C	5	0.0019	FST
rs12777	C	5	0.005	SLC22A4
rs11746506	T	5	0.0017	MRPS30
rs11743174	T	5	0.0019	ABLIM3
rs495237	T	5	0.0027	LINC00603
rs79760705	T	5	0.0056	ARL15
rs881858	A	6	-0.0054	LINC01512
rs72912510	A	6	-0.0024	RRAGD
rs9375818	A	6	-0.0031	ARG1
rs3822939	A	6	-0.0025	EYA4
rs12207180	A	6	-0.0085	SLC22A2
rs12212034	T	6	-0.0018	PKHD1
rs6458868	T	6	-0.002	GSTA2
rs3925003	T	6	-0.0018	HMGCLL1
rs13200335	A	6	0.0024	TFEB
rs11755724	A	6	0.0027	RREB1
rs1857859	A	6	0.0019	SIM1
rs1268168	A	6	0.0024	FOXO3
rs9397738	A	6	0.0027	SCAF8
rs77915916	A	6	0.0046	CRIP3
rs7740107	A	6	0.0027	L3MBTL3
rs3765502	T	6	0.0024	DCDC2
rs144100226	T	6	0.0059	HMGA1
rs720989	T	6	0.0021	SUPT3H

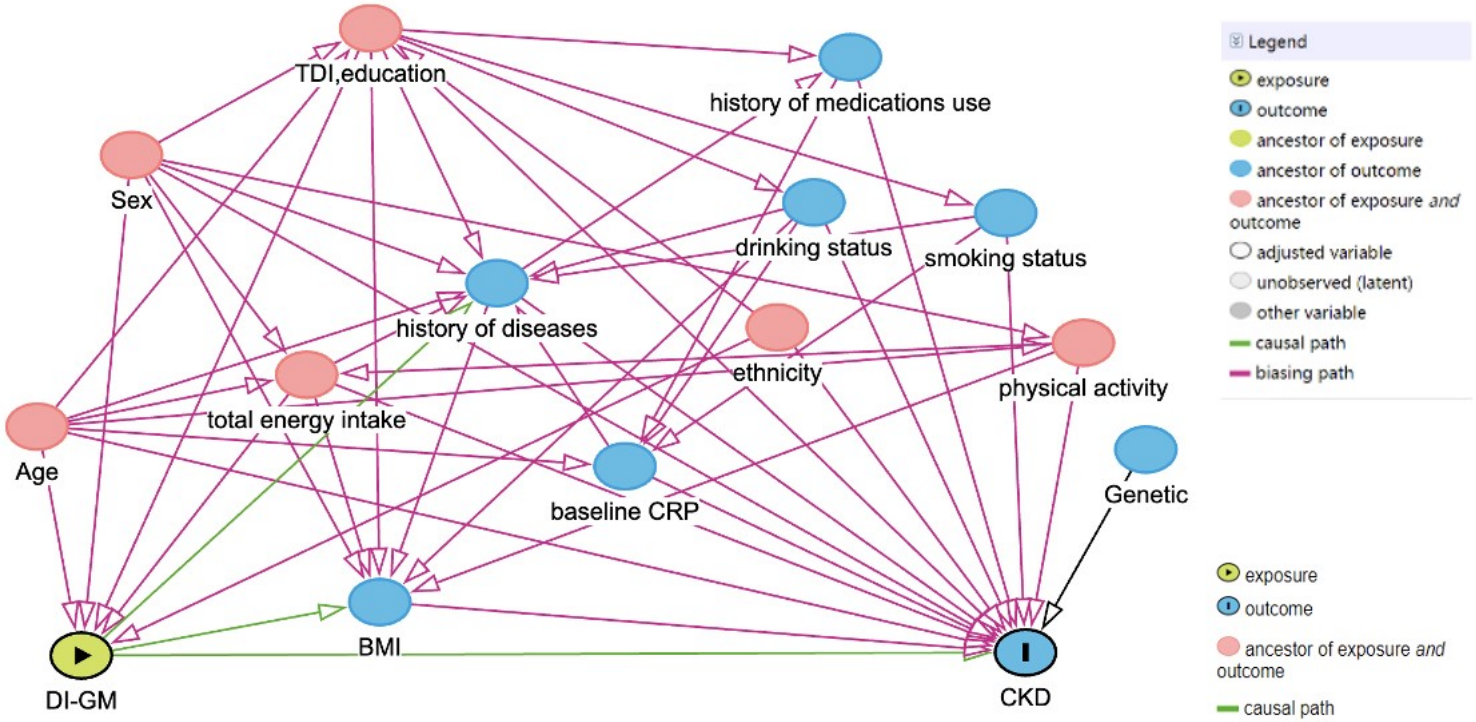
rs35154268	A	7	-0.0022	SND1
rs6968554	A	7	-0.0019	AHR
rs62491533	T	7	-0.0027	UBE2H
rs10254101	T	7	-0.0068	PRKAG2
rs62435145	T	7	-0.006	UNCX
rs3750081	T	7	-0.0022	KBTBD2
rs801193	T	7	-0.002	GS1-124K5.11
rs6973656	A	7	0.0035	TMEM60
rs700753	C	7	0.0031	LOC730338
rs55773927	T	7	0.0019	VKORC1L1
rs41301394	T	7	0.0023	POR
rs3757387	T	7	0.003	IRF5
rs12671694	T	7	0.0025	SHH
rs868822	T	7	0.0029	LINC01006
rs11783418	A	8	-0.002	XKR6
rs10102889	C	8	-0.0036	NRG1
rs2976178	C	8	-0.0025	WWP1
rs2980423	T	8	-0.0023	PRAG1
rs35353426	T	8	-0.0026	LOC157273
rs10098664	T	8	-0.0021	BLK
rs34861762	T	8	-0.0043	STC1
rs1533059	A	8	0.0025	MFHAS1
rs7832708	T	8	0.0022	MSRA
rs2954017	T	8	0.0024	TRIB1
rs12377027	A	9	-0.0026	MLLT3
rs13287724	A	9	-0.003	B4GALT1-AS1
rs1321917	C	9	-0.0023	ASTN2
rs544169	A	9	0.0022	UBAP2
rs2039424	A	9	0.0044	PIP5K1B
rs7024579	T	9	0.0023	QSOX2
rs2068888	A	10	-0.0024	CYP26A1
rs4918943	A	10	-0.0022	SORBS1
rs12240572	A	10	-0.0032	DNAJC9-AS1
rs7095954	A	10	-0.0018	TSPAN14
rs816850	C	10	-0.002	KCNMA1
rs1536225	T	10	-0.0021	PDCD11
rs7072591	A	10	0.0019	PARD3-AS1
rs10821905	A	10	0.0037	A1CF
rs1055256	A	10	0.0025	EEF1AKMT2
rs80282103	A	10	0.0078	LARP4B

rs6481598	C	10	0.0024	SVIL
rs8474	C	10	0.002	PARG
rs7475348	T	10	0.0031	MYPN
rs9420446	T	10	0.0023	FAM35A
rs10821944	T	10	0.002	ARID5B
rs284859	T	10	0.0026	WBP1L
rs1541937	A	11	-0.0029	OR52H1
rs1783827	A	11	-0.002	MIR130A
rs3892895	A	11	-0.0023	TPCN2
rs963837	T	11	-0.0057	DCDC1
rs6484504	T	11	-0.0026	DNAJC24
rs2727040	T	11	-0.0026	TRIM49B
rs948493	T	11	-0.0033	MIR1234
rs11237450	A	11	0.0032	GAB2
rs63934	A	11	0.0041	KCNQ1
rs6589750	A	11	0.002	USP2-AS1
rs11564722	T	11	0.0033	INS-IGF2
rs61897431	T	11	0.0029	SLC39A13
rs7127946	T	11	0.0023	OR4B1
rs1813937	T	11	0.0022	LOC646813
rs10790452	T	11	0.002	SORL1
rs10846157	A	12	-0.0034	RERG
rs11062167	A	12	-0.0039	SLC6A13
rs632887	A	12	0.0032	TSPAN9
rs117113238	A	12	0.0039	BCL2L14
rs2634675	A	12	0.0025	ZNF641
rs1275609	A	12	0.0024	PHLDA1
rs4238020	T	12	0.0029	C12orf4
rs12313306	T	12	0.0029	R3HDM2
rs41284816	T	13	-0.0078	DLEU2
rs500830	T	13	0.0029	DACH1
rs61993680	A	14	-0.0019	SLC25A29
rs72683923	T	14	-0.0074	L2HGDH
rs6574652	T	14	-0.0017	STON2
rs17184313	T	14	-0.0029	RIN3
rs1028455	A	14	0.002	SPATA7
rs690428	A	15	-0.0039	WDR72
rs1994887	A	15	-0.002	CGNL1
rs351237	A	15	-0.0018	STRA6
rs4886696	A	15	-0.0032	SIN3A

rs6492982	T	15	-0.0033	INO80
rs11071738	T	15	-0.0025	APH1B
rs11071939	T	15	-0.0039	SMAD3
rs1145077	T	15	-0.0085	GATM
rs59646751	T	15	-0.0023	IGF1R
rs4886755	A	15	0.0041	NRG4
rs17507300	A	15	0.0024	BTBD1
rs7169629	C	15	0.0018	WDR73
rs12913015	T	15	0.0027	C15orf54
rs956006	T	15	0.0019	MGC15885
rs2472297	T	15	0.0039	CYP1A1
rs166906	T	15	0.0033	SCAPER
rs9932625	A	16	-0.003	LINC01571
rs28581385	A	16	-0.0028	LINC01229
rs154656	A	16	-0.003	CHMP1A
rs1635404	T	16	-0.0025	TRAP1
rs193538	T	16	-0.002	ABCC1
rs7185391	T	16	-0.0027	SLC7A6
rs62053077	T	16	-0.0021	MARVELD3
rs7203398	A	16	0.0025	CHD9
rs77924615	A	16	0.0098	PDILT
rs62050038	A	16	0.0028	WWP2
rs438339	T	16	0.0035	RPL3L
rs1858800	T	16	0.002	ZFHX3
rs883541	A	17	-0.0022	PRKAR1A
rs2411192	A	17	-0.0024	MYO19
rs8866	C	17	-0.0018	PITPNC1
rs2349648	T	17	-0.0017	MPRIP
rs4794813	A	17	0.0055	CDK12
rs35662455	C	17	0.003	TEX14
rs9903801	C	17	0.0047	BCAS3
rs9891340	T	17	0.0024	SMCR2
rs2440165	T	17	0.004	SLC47A1
rs9895661	T	17	0.0069	BCAS3
rs28735420	T	17	0.0039	MAP2K4
rs227731	T	17	0.0018	C17orf67
rs16942751	A	18	-0.0029	AQP4
rs1719934	A	18	0.0026	EPB41L3
rs8096658	C	18	0.005	NFATC1
rs4940525	T	18	0.0025	LINC01544

rs34647824	A	19	-0.0021	RRAS
rs78241494	T	19	-0.003	ZNF585A
rs281380	T	19	-0.0021	MAMSTR
rs2974751	A	19	0.0018	CALR
rs8101667	T	19	0.0044	CEP89
rs7251730	T	19	0.0024	ZNF260
rs113445505	T	19	0.0037	ZNF781
rs6087579	A	20	-0.0028	ITCH
rs4408777	A	20	-0.0021	RGS19
rs2235826	A	20	-0.003	PCK1
rs1041606	T	20	-0.0021	MACROD2
rs17216707	T	20	-0.0051	CYP24A1
rs1509117	A	20	0.0024	PLCB1
rs72629024	C	20	0.0035	PPDPF
rs62187537	T	20	0.0039	FKBP1ASDCBP2
rs1407040	T	20	0.0018	GNAS
rs35636653	T	20	0.0022	OSBPL2
rs2273684	T	20	0.0032	GSS
rs2823139	A	21	-0.0026	NRIP1
rs2834317	A	21	-0.0035	LOC101928126
rs2244237	T	21	0.0027	CLDN14
rs80576	A	22	-0.0028	APOL3
rs4820324	C	22	-0.0023	MAFF
rs131263	T	22	0.0024	ZMAT5
rs112880707	T	22	0.0052	MKL1
rs738527	T	22	0.0032	A4GALT

Abbreviations: SNP, single nucleotide polymorphism; eGFR, estimated glomerular filtration rate; GRS, genetic risk score; Chr, chromosome.



Supplementary Figure 1. Directed acyclic graph for the associations of the DI-GM with the risk of CKD.

Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; TDI, Townsend deprivation index; CRP, C-reactive protein.

Supplementary Table 5. Baseline characteristics of participants according to categories of the DI-GM (n=166,364) [a].

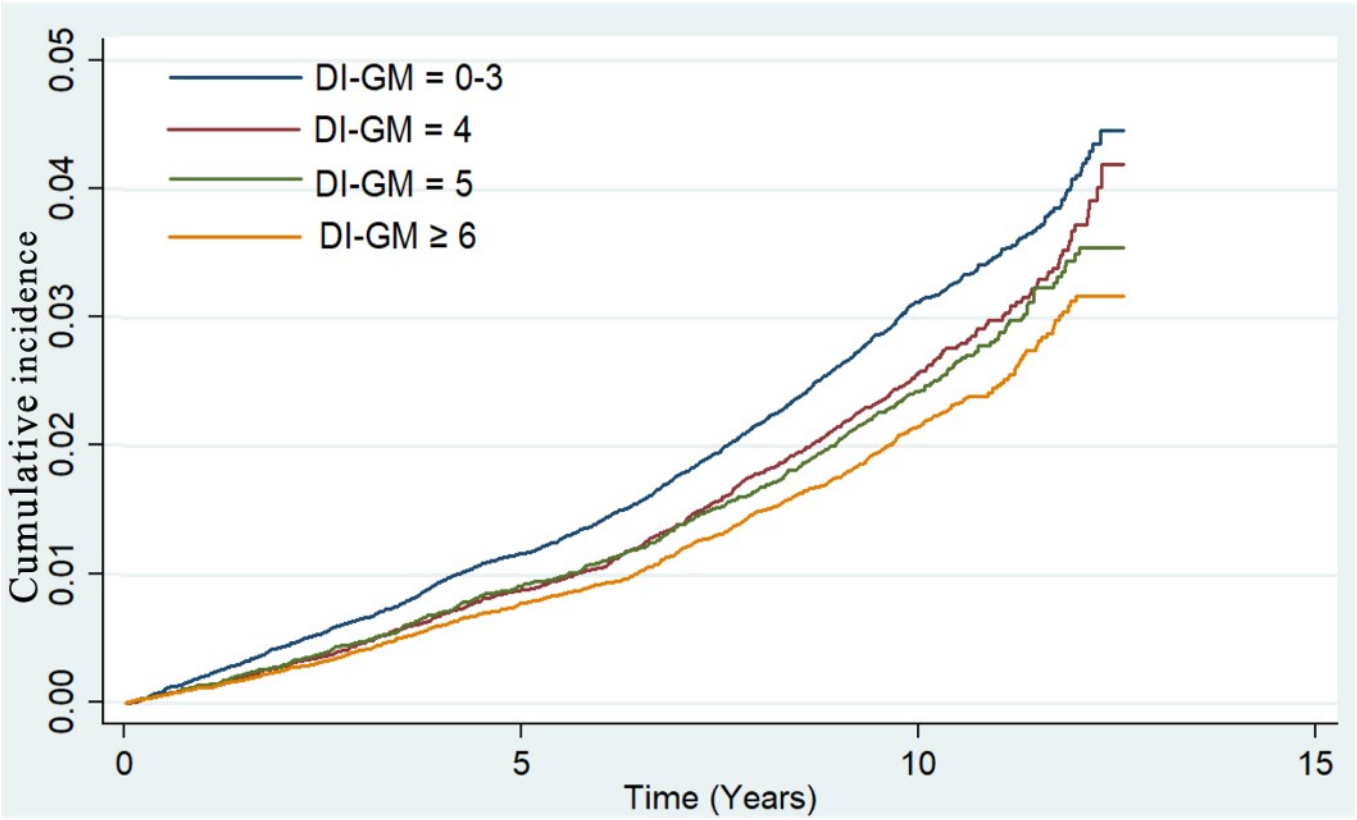
Characteristics	DI-GM				*P*-value
	0–3 (n = 54,747)	4 (n = 34,904)	5 (n = 31,705)	≥ 6 (n = 45,008)	
Demographic information					
Age at recruitment (years)	55.1 (7.97)	56.0 (7.95)	56.4 (7.84)	56.8 (7.74)	<0.0001
Sex (female, %)	50.0	53.6	55.8	60.9	<0.0001
Race (white, %)	95.0	95.8	96.1	96.3	<0.0001
Education level (≥ College graduate, %)	37.4	41.1	43.8	50.3	<0.0001
Townsend Deprivation Index	-1.49 (2.92)	-1.62 (2.84)	-1.73 (2.79)	-1.66 (2.83)	<0.0001
Lifestyles					
Smoking status (%)					<0.0001
Current smoker	10.2	8.08	6.64	5.15	
Ex-smoker	34.4	35.8	35.6	35.9	
Non-smoker	55.5	56.1	57.8	59.0	
Alcohol drinking status (%)					0.004
Current drinker	93.9	94.1	93.7	93.5	
Ex-drinker	2.76	2.87	3.09	3.47	
Non-drinker	3.31	3.02	3.26	3.03	
Physical activity (%)					<0.0001
Low	21.5	18.8	16.6	13.5	
Moderate	42.7	42.6	42.8	41.2	
High	35.8	38.7	40.6	45.2	
Total energy intake (kJ/day)	8220.0 (2324.3)	8406.2 (2466.8)	8668.0 (2443.3)	8889.7 (2314.1)	<0.0001
Anthropometric indexes					
BMI (kg/m^2)	27.4 (4.71)	27.0 (4.55)	26.7 (4.38)	26.0 (4.27)	<0.0001

Blood pressure and biochemistry					
SBP (mmHg)	136.8 (17.9)	136.7 (17.9)	136.3 (18.1)	135.5 (18.1)	<0.0001
DBP (mmHg)	82.4 (10.0)	81.9 (9.89)	81.4 (9.84)	80.7 (9.80)	<0.0001
eGFR (mL/min/1.73 m^2)	91.9 (11.9)	92.0 (11.7)	92.1 (11.5)	92.9 (11.3)	<0.0001
CRP (mg/L)	2.52 (4.19)	2.34 (3.97)	2.16 (3.78)	1.92 (3.45)	<0.0001
Individual history of disease (%)					
Diabetes	5.85	5.48	5.48	4.17	<0.0001
Hypertension	52.6	51.3	50.0	47.2	<0.0001
CVD	4.14	4.02	3.86	3.38	<0.0001
Cancer	9.17	9.63	10.00	11.1	<0.0001
Antihypertensive medication use (%)	18.2	17.7	16.0	14.4	<0.0001
Cholesterol-lowering medication use (%)	15.1	14.7	14.4	12.6	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; CRP, C-reactive protein; CVD, cardiovascular disease.

[b] Continuous variables are expressed as mean (standard deviation) and categorical variables are expressed as percentages.

[c] Analysis of variance or logistic regression.



Supplementary Figure 2. The Cumulative Incidence of CKD according to categories of the DI-GM

Abbreviations: CKD, chronic kidney disease; DI-GM, dietary index for gut microbiota.

Supplementary Table 6. Associations between the DI-GM and the risk of CKD when excluded participants who developed CKD within first 2 years of follow-up (n=165,805) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	54,504	34,802	31,610	44,889	-
Number of cases	1,304	731	622	761	-
Person-years	501,843	322,519	291,986	411,857	-
Model 1 [d]	1.00 (reference)	0.87 (0.79, 0.95) [c]	0.82 (0.74, 0.90)	0.72 (0.66, 0.78)	<0.0001
Model 2 [e]	1.00 (reference)	0.82 (0.75, 0.89)	0.77 (0.70, 0.85)	0.69 (0.63, 0.76)	<0.0001
Model 3 [f]	1.00 (reference)	0.84 (0.76, 0.92)	0.81 (0.74, 0.89)	0.76 (0.69, 0.83)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.

Supplementary Table 7. Associations between the DI-GM and the risk of CKD when excluded participants who developed CKD within first 5 years of follow-up (n= 164,786) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	54,115	34,599	31,413	44,659	-
Number of cases	915	528	425	531	-
Person-years	500,495	321,812	291,293	411,046	-
Model 1 [d]	1.00 (reference)	0.89 (0.80, 0.99) [c]	0.80 (0.71, 0.89)	0.72 (0.64, 0.80)	<0.0001
Model 2 [e]	1.00 (reference)	0.84 (0.75, 0.93)	0.75 (0.67, 0.84)	0.69 (0.62, 0.77)	<0.0001
Model 3 [f]	1.00 (reference)	0.85 (0.76, 0.95)	0.78 (0.70, 0.88)	0.75 (0.67, 0.83)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.

Supplementary Table 8. Associations between the DI-GM and the risk of CKD when excluding participants who had a history of cancer or CVD (n=144,191) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	47,730	30,335	27,465	38,661	-
Number of cases	1,150	624	522	626	-
Person-years	438,829	280,731	253,287	353,756	-
Model 1 [d]	1.00 (reference)	0.85 (0.77, 0.93)	0.79 (0.71, 0.87)	0.68 (0.62, 0.75)	<0.0001
Model 2 [e]	1.00 (reference)	0.80 (0.72, 0.88)	0.73 (0.66, 0.81)	0.65 (0.59, 0.72)	<0.0001
Model 3 [f]	1.00 (reference)	0.82 (0.75, 0.91)	0.78 (0.70, 0.86)	0.70 (0.64, 0.78)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension and diabetes), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.

Supplementary Table 9. Associations between the DI-GM and the risk of CKD when participants with at least 2 times 24-hour dietary recall questionnaire with typical diet and credible energy (n=89,896) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	29,429	17,397	16,399	26,671	-
Number of cases	788	421	360	480	-
Person-years	256,568	152,048	143,592	233,660	-
Model 1 [d]	1.00 (reference)	0.90 (0.80, 1.02)	0.82 (0.72, 0.92)	0.67 (0.60, 0.75)	<0.0001
Model 2 [e]	1.00 (reference)	0.87 (0.77, 0.97)	0.78 (0.69, 0.89)	0.68 (0.61, 0.76)	<0.0001
Model 3 [f]	1.00 (reference)	0.90 (0.80, 1.01)	0.83 (0.73, 0.94)	0.74 (0.66, 0.83)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.

Supplementary Table 10. Associations between the DI-GM and the risk of CKD further adjusted for eGFR (n=166,364) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	54,747	34,904	31,705	45,008	-
Number of cases	1,547	833	717	880	-
Person-years	502,093	322,627	292,082	411,981	-
Model 1 [d]	1.00 (reference)	0.84 (0.77, 0.91) [c]	0.80 (0.73, 0.87)	0.70 (0.64, 0.76)	<0.0001
Model 2 [e]	1.00 (reference)	0.78 (0.72, 0.85)	0.75 (0.68, 0.82)	0.67 (0.62, 0.73)	<0.0001
Model 3 [f]	1.00 (reference)	0.84 (0.77, 0.91)	0.85 (0.78, 0.93)	0.85 (0.78, 0.92)	0.0003

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, the genotype measurement batch, and baseline eGFR.

Supplementary Table 11. Associations between the DI-GM and the risk of CKD using competing risk regression (n=166,364) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	54,747	34,904	31,705	45,008	-
Number of cases	1,547	833	717	880	-
Person-years	502,093	322,627	292,082	411,981	-
Model 1 [d]	1.00 (reference)	0.84 (0.77, 0.91) [c]	0.80 (0.73, 0.87)	0.70 (0.64, 0.76)	<0.0001
Model 2 [e]	1.00 (reference)	0.79 (0.72, 0.86)	0.75 (0.69, 0.82)	0.68 (0.62, 0.74)	<0.0001
Model 3 [f]	1.00 (reference)	0.81 (0.74, 0.88)	0.79 (0.72, 0.87)	0.74 (0.68, 0.80)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

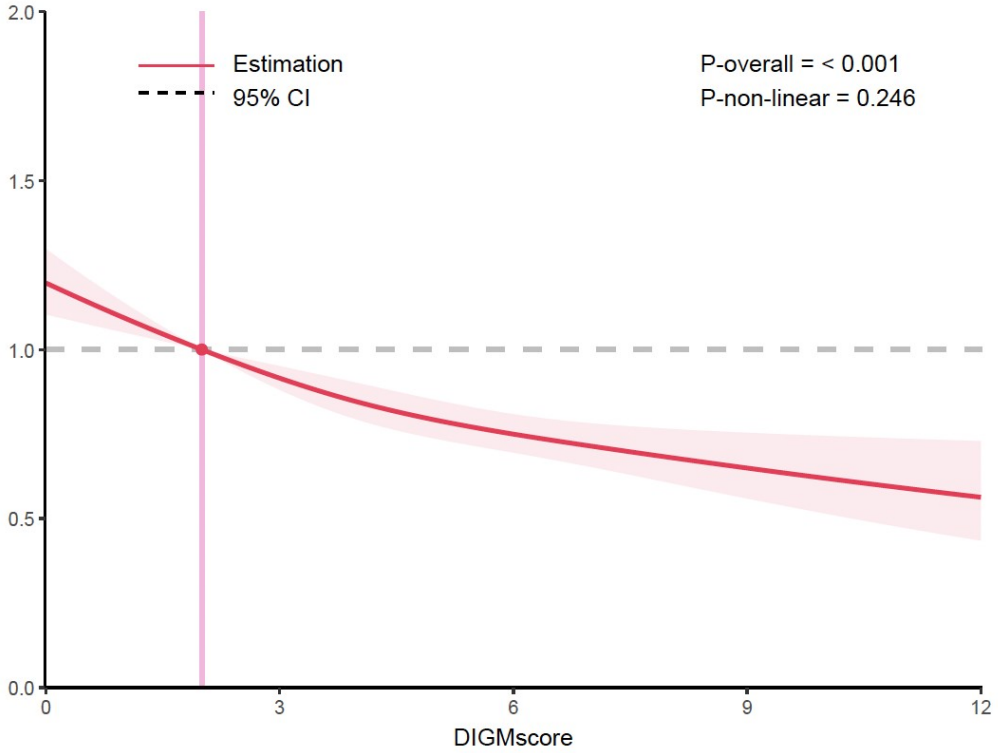
[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

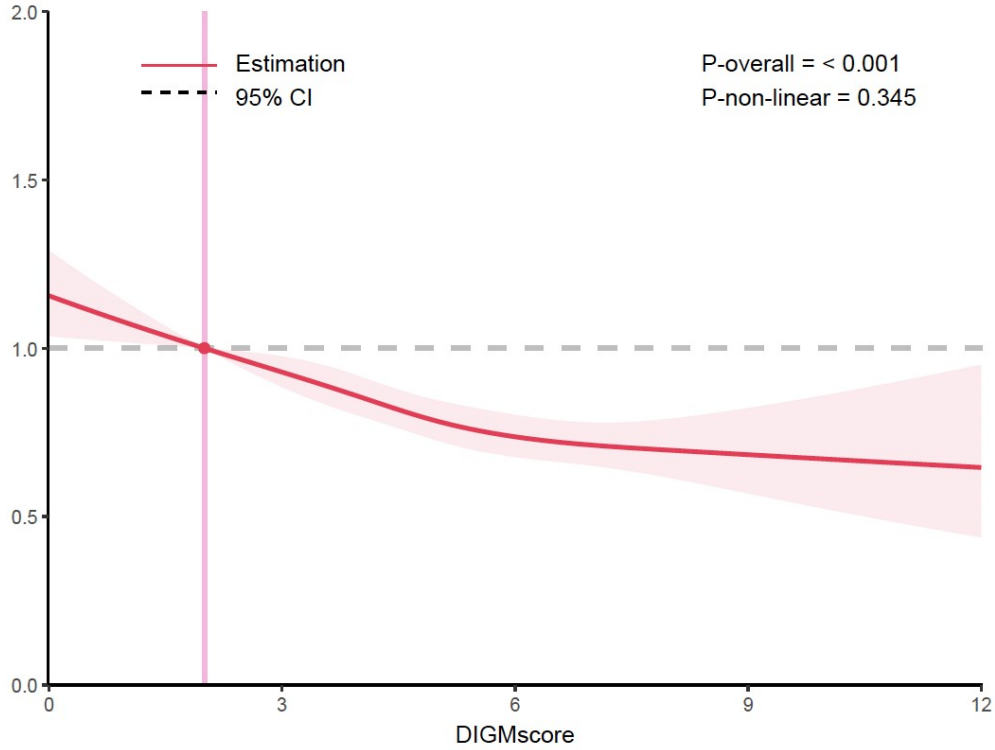
[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.



Supplementary Figure 3. The dose-response association between the DI-GM and risk of CKD diabetes using restricted cubic spline with three knots.

Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; HR, hazard ratio; CI, confidence interval; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

The model was adjusted for age, sex, BMI, ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.



Supplementary Figure 4. The dose-response association between the DI-GM and risk of CKD diabetes using restricted cubic spline with four knots.

Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; HR, hazard ratio; CI, confidence interval; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

The model was adjusted for age, sex, BMI, ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.

Supplementary Table 12. Associations between the DI-GM and the risk of CKD including participants lost to follow-up (n=166,821) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	54,893	35,004	31,799	45,125	-
Number of cases	1,547	833	717	880	-
Person-years	503,054	323,284	292,704	412,734	-
Model 1 [d]	1.00 (reference)	0.84 (0.77, 0.91) [c]	0.80 (0.73, 0.87)	0.70 (0.64, 0.76)	<0.0001
Model 2 [e]	1.00 (reference)	0.78 (0.72, 0.85)	0.75 (0.68, 0.82)	0.67 (0.62, 0.73)	<0.0001
Model 3 [f]	1.00 (reference)	0.81 (0.74, 0.88)	0.79 (0.72, 0.86)	0.73 (0.67, 0.80)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, and the genotype measurement batch.

Supplementary Table 13. Associations between the DI-GM and the risk of CKD further adjusted for history of antibiotic use, history of bowel resection surgery, and IBD (n=166,364) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	54,747	34,904	31,705	45,008	-
Number of cases	1,547	833	717	880	-
Person-years	502,093	322,627	292,082	411,981	-
Model 1 [d]	1.00 (reference)	0.84 (0.77, 0.91) [c]	0.80 (0.73, 0.87)	0.70 (0.64, 0.76)	<0.0001
Model 2 [e]	1.00 (reference)	0.78 (0.72, 0.85)	0.75 (0.68, 0.82)	0.67 (0.62, 0.73)	<0.0001
Model 3 [f]	1.00 (reference)	0.81 (0.74, 0.88)	0.79 (0.72, 0.86)	0.74 (0.68, 0.80)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein, IBD, inflammatory bowel disease.

[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, the genotype measurement batch, history of antibiotic use, history of bowel resection surgery, and IBD.

Supplementary Table 14. Associations between the DI-GM and the risk of CKD excluding participants who history of antibiotic use, history of bowel resection surgery, and IBD (n=151,160) [a].

□	DI-GM				*P* for trend [b]
	0-3	4	5	≥ 6	
Number of participants	49,764	31,796	28,876	40,724	-
Number of cases	1,395	749	657	806	-
Person-years	456,275	294,104	265,994	372,705	-
Model 1 [d]	1.00 (reference)	0.83 (0.76, 0.91) [c]	0.81 (0.74, 0.89)	0.71 (0.65, 0.78)	<0.0001
Model 2 [e]	1.00 (reference)	0.78 (0.71, 0.85)	0.75 (0.69, 0.83)	0.68 (0.63, 0.75)	<0.0001
Model 3 [f]	1.00 (reference)	0.80 (0.73, 0.87)	0.79 (0.72, 0.87)	0.74 (0.68, 0.81)	<0.0001

[a] Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein, IBD, inflammatory bowel disease.

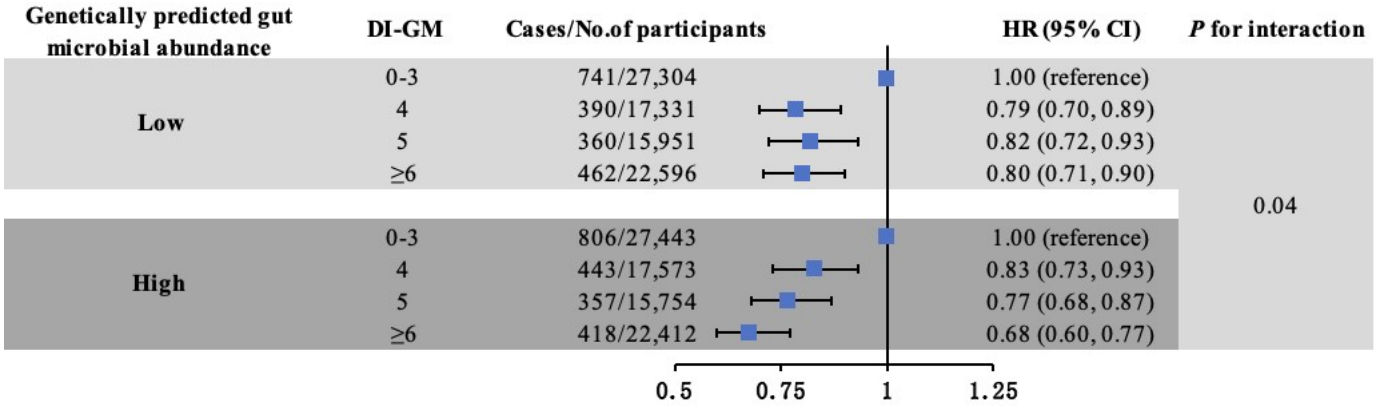
[b] Calculated by using DI-GM as a continuous variable.

[c] Hazard ratios (95% confidence interval) (all such values).

[d] Model 1 was a crude model.

[e] Model 2 was adjusted for age, sex, and BMI.

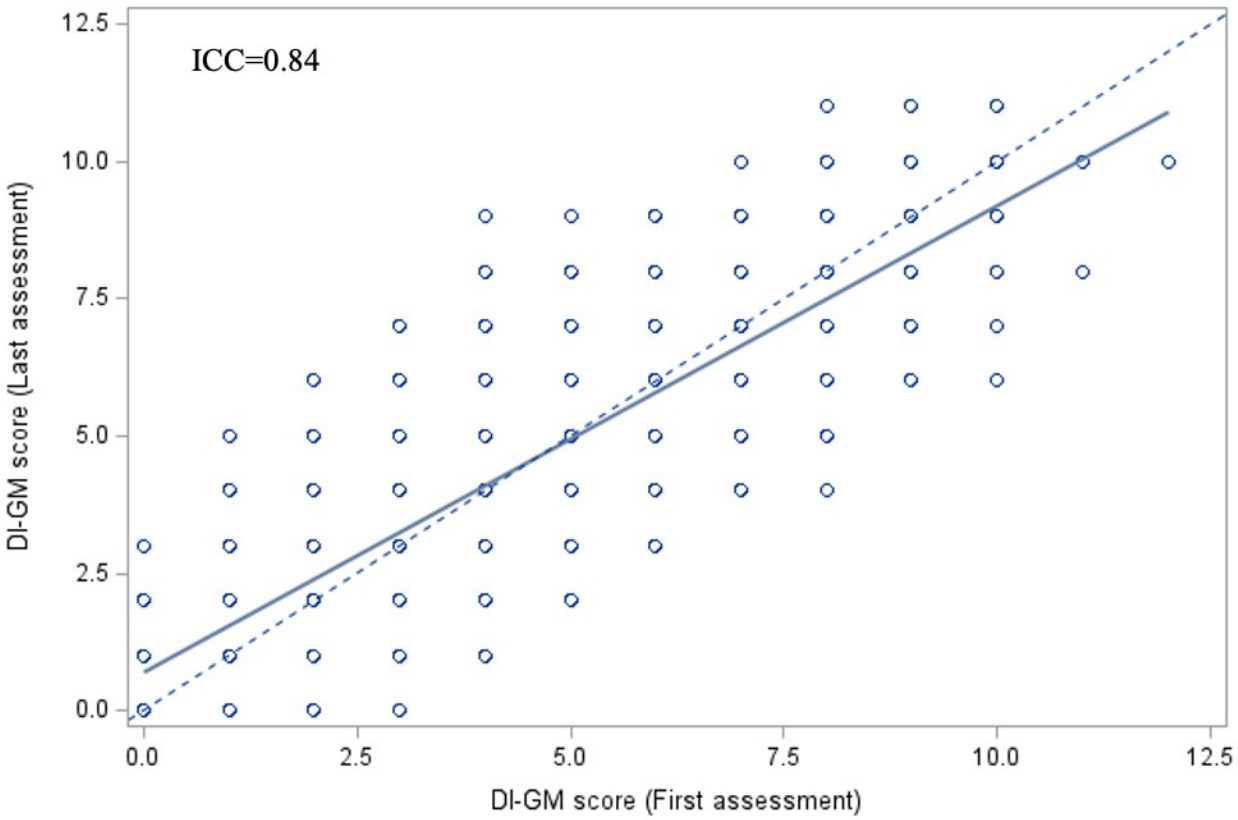
[f] Model 3 was further adjusted for ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, the genotype measurement batch, history of antibiotic use, history of bowel resection surgery, and IBD.



Supplementary Figure 5. Associations of the DI-GM with the risk of CKD according to genetically predicted gut microbial abundance, additionally adjusted for antibiotic use, bowel resection surgery, and IBD history.

Abbreviations: DI-GM, dietary index for gut microbiota; CKD, chronic kidney disease; HR, hazard ratio; CI, confidence interval; BMI, body mass index; CVD, cardiovascular disease; GRS, genetic risk score; CRP, C-reactive protein; IBD, inflammatory bowel disease.

The model was adjusted for age, sex, BMI, ethnicity, education levels, Townsend deprivation index, smoking status, drinking status, physical activity, total energy intake, history of diseases (hypertension, diabetes, CVD, and cancer), use of anti-hypertensive and cholesterol-lowering medications, CKD-GRS, baseline CRP, the first 10 principal components of ancestry, the genotype measurement batch, genetically predicted gut microbial abundance, history of antibiotic use, history of bowel resection surgery, and IBD.



Supplementary Figure 6. Reproducibility of the DI-GM score across repeated dietary assessments
Abbreviations: DI-GM, dietary index for gut microbiota; intraclass correlation coefficient, ICC.
Each point represents an individual participant with at least two dietary assessments. The DI-GM scores were calculated separately based on the first and the last dietary questionnaires. The dashed line indicates the line of identity (y = x), and the solid line represents the fitted regression line.