Electronic supplementary material

Evaluating the impact of AMPK activation, a target of metformin, on risk of cardiovascular diseases and cancer in the UK Biobank: A Mendelian randomization study

Running title: The impact of metformin on CVD and cancer

Shan Luo, MPH (0000-0003-3368-9935), Catherine Mary Schooling, PhD (0000-0001-9933-5887), Ian Chi Kei Wong, PhD (0000-0001-8242-0014), Shiu Lun Au Yeung, PhD (0000-0001-6136-1836)

Table of contents

ESM Fig. 1: Linkage disequilibrium matrix for variants included in the AMP-activated protein kinase score

ESM Fig. 2. Association of AMP-activated protein kinase score with risk of a) stroke, b) prostate cancer, c) breast cancer, d) colorectal cancer in the UK Biobank

ESM Fig. 3. Genetic associations instrumented by genetic variants with HbA_{1c} against a) type 2 diabetes, b) coronary artery disease, c) stroke, d) breast cancer, e) prostate cancer

ESM Table 1: Variants included in AMP-activated protein kinase score and association with glycated haemoglobin in the Meta-Analyses of Glucose and Insulin-related traits Consortium, restricted to participants of European descent

Variant	Effect allele	Other allele	EAF	Effect size	Standard error	P $_{MAGIC}$	P $_{UKB}$
rs11239944	A	G	0.825	-0.018	0.0061	3.05E-03	1.88E-04
rs2059409	T	C	0.045	-0.011	0.0047	2.17E-02	1.16E-02
rs1365964	G	A	0.801	-0.0057	0.0029	4.93E-02	6.91E-03
rs11884246	A	C	0.522	-0.0033	0.0017	4.64E-02	2.61E-02
rs6726126	A	G	0.518	-0.0038	0.0018	3.09E-02	3.38E-07
rs16858808	A	G	0.04	-0.025	0.012	3.22E-02	2.36E-03
rs17572109	A	G	0.208	-0.0055	0.002	5.18E-03	1.54E-05
rs7596500	G	T	0.982	-0.016	0.0074	3.16E-02	8.87E-03
rs3816560	C	T	0.221	-0.004	0.0019	3.19E-02	2.07E-05
rs10230736	G	A	0.862	-0.011	0.0038	4.44E-03	4.13E-04
rs1808593	T	G	0.827	-0.0049	0.0024	3.96E-02	3.83E-03
rs1563636	C	T	0.196	-0.0057	0.0025	2.18E-02	1.94E-03
rs7806203	C	T	0.345	-0.0039	0.002	4.74E-02	6.75E-03
rs7780461	T	C	0.118	-0.0063	0.0032	4.97E-02	1.16E-02
rs10259821	A	G	0.31	-0.0037	0.0019	4.71E-02	1.20E-03
rs1635527	C	G	0.525	-0.0038	0.0019	4.15E-02	4.35E-36
rs1859444	C	T	0.808	-0.0048	0.0024	4.39E-02	1.42E-22
rs1859443	G	A	0.856	-0.0058	0.0025	1.96E-02	6.27E-13
rs11168355	G	A	0.792	-0.0055	0.0022	1.14E-02	2.63E-43
rs11168359	A	G	0.119	-0.015	0.0032	1.61E-06	6.94E-91
rs12297820	A	G	0.119	-0.017	0.0029	1.26E-08	1.09E-43
rs10492081	A	G	0.81	-0.0071	0.0024	2.87E-03	9.20E-21
rs17614932	A	C	0.975	-0.012	0.0051	1.45E-02	2.70E-11
rs10875764	C	T	0.584	-0.0077	0.0038	4.45E-02	1.32E-13
rs7134565	C	T	0.469	-0.0035	0.0018	4.68E-02	3.79E-37
rs2732480*	A	C	0.425	-0.012	0.002	2.00E-09	1.07E-142
rs1489107	A	G	0.031	-0.014	0.0058	1.73E-02	2.16E-22
rs2932091	A	C	0.325	-0.0043	0.0021	3.79E-02	1.25E-03
rs10875801	C	T	0.825	-0.0089	0.0023	9.35E-05	5.95E-30
rs4760702	A	T	0.967	-0.011	0.0048	1.98E-02	1.16E-09
rs7959684	A	G	0.226	-0.005	0.0021	1.58E-02	1.30E-50
rs11168547	T	C	0.943	-0.0079	0.0026	2.28E-03	1.06E-26
rs12582586	T	C	0.212	-0.005	0.0024	3.58E-02	1.92E-04
rs10875814	G	A	0.733	-0.0051	0.0024	3.22E-02	1.34E-09
rs12582811	G	T	0.128	-0.0089	0.0033	7.23E-03	1.96E-06
rs11168643	A	G	0.034	-0.018	0.0061	2.89E-03	3.84E-08
rs10875843	A	G	0.033	-0.0093	0.0045	3.67E-02	4.79E-08
rs7975821	A	G	0.724	-0.0041	0.0021	4.77E-02	2.81E-07
rs17834622	A	G	0.346	-0.0064	0.0019	5.95E-04	1.10E-17
rs12830014	G	A	0.914	-0.0074	0.0037	4.71E-02	8.62E-04
rs10783277	C	T	0.674	-0.0045	0.0019	1.58E-02	4.35E-05
rs12322783	A	G	0.189	-0.0045	0.0023	4.82E-02	1.80E-08
rs17197593	T	C	0.053	-0.014	0.0062	2.23E-02	5.50E-05
rs1050187	C	T	0.704	-0.0042	0.0019	2.42E-02	1.71E-03

Glycated haemoglobin (HbA_{1c}) measured in percentage. For each variant, the effect allele is the allele associated with lower HbA_{1c}. Association are taken from Wheeler E *et al. PLoS Med.* EAF, effect allele frequency.
*Variant used in the sensitivity analysis, chosen based on the stringent variant selection criteria.

ESM Table 2: UK biobank's algorithmically defined disease outcomes

Data fields	Type 2 diabetes	Coronary artery disease	Stroke	Overall cancer	Breast cancer	Colorectal cancer	Prostate cancer
ICD-9 40013, 41203, 41205	250	410-414	430, 431, 434, 436	140-239	174-175	153.0-154.1	185
ICD-10 40001, 40002, 40006, 41202, 41204	E11	I20-I25	I60, I61, I63, I64	C00-C97	C50	C18-C20	C61
20001				data coding3	1002	1020	1044
20002	1223	1074, 1075	1081, 1086, 1491, 1583				
2453				1			
6150		1, 2	3				
20003	1140884600, 1140874686, 1141153254, 1141171646, 1141177600, 1141189090, 1140857584, 1140874706, 1140874664, 1140874674, 1140857494, 1140874744, 1140874646, 1141157284, 1140874658, 1141152590, 1141168660, 1141173882						
30750	≥ 6.5%, or 48 mmol/mol						
30740	≥ 11.1 mol/L						

ESM Table 3: The variants associated with glycated haemoglobin in the Meta-Analyses of Glucose and Insulin-related traits Consortium, restricted to participants of European descent

Variant	Effect allele	Other allele	Effect allele frequency	Beta	Standard error	P value	Sample size
rs267738	T	G	0.7701	0.011	0.0019	2.59E-09	118146
rs857691	T	C	0.2715	0.019	0.0019	3.97E-25	121554
rs17509001	C	T	0.1576	0.018	0.0023	1.94E-15	121575
rs12621844	T	C	0.5999	0.01	0.0018	1.87E-08	88288
rs560887	C	T	0.6843	0.028	0.0018	1.48E-58	109489
rs7616006	A	G	0.5744	0.01	0.0017	5.07E-10	121507
rs9818758	A	G	0.2028	0.012	0.002	7.74E-10	121581
rs11708067	A	G	0.7542	0.013	0.0019	1.42E-12	119780
rs8192675	T	C	0.6906	0.011	0.0017	1.38E-11	119841
rs13134327	A	G	0.3335	0.013	0.0017	2.64E-15	119717
rs7756992	G	A	0.2862	0.012	0.0018	2.80E-12	118856
rs198846	G	A	0.8293	0.022	0.0022	1.18E-23	120479
rs11964178	A	G	0.5666	0.01	0.0016	6.38E-10	121505
rs592423	A	C	0.4566	0.009	0.0017	3.96E-08	107880
rs4607517	A	G	0.2017	0.031	0.0024	8.76E-38	86837
rs6474359	T	C	0.953	0.044	0.0053	1.50E-16	95687
rs4737009	A	G	0.2531	0.021	0.002	4.48E-27	120823
rs6980507	A	G	0.4014	0.01	0.0018	3.58E-08	108044
rs11558471	A	G	0.6745	0.015	0.0017	1.38E-19	121354
rs2383208	A	G	0.7992	0.014	0.0021	7.04E-12	113265
rs7040409	C	G	0.8953	0.028	0.0037	2.56E-14	106582
rs579459	C	T	0.2389	0.011	0.0019	9.42E-09	120555
rs4745982	T	G	0.8726	0.095	0.0056	2.87E-65	69523
rs17747324	C	T	0.2489	0.015	0.0023	6.12E-11	87696
rs3782123	C	A	0.3205	0.013	0.002	1.51E-10	105906
rs11603334	G	A	0.815	0.012	0.0021	6.85E-09	120520
rs10830963	G	C	0.2938	0.02	0.002	2.23E-23	100954
rs2110073	T	C	0.1097	0.015	0.0028	4.44E-08	119835
rs10774625	G	A	0.5056	0.009	0.0016	1.46E-08	121429
rs282587	G	A	0.1513	0.019	0.0027	1.70E-12	88316
rs9604573	A	G	0.2738	0.01	0.0018	9.60E-09	115096
rs11248914	T	C	0.647	0.014	0.0019	2.56E-14	85606
rs1558902	A	T	0.4128	0.01	0.0019	3.27E-08	88319
rs837763	T	C	0.5548	0.017	0.0016	1.68E-28	111180
rs9914988	A	G	0.7877	0.013	0.002	2.77E-11	121502
rs1046896	T	C	0.3162	0.028	0.0017	4.46E-64	123491
rs17533903	A	G	0.2428	0.015	0.0022	5.27E-12	119537
rs4820268	G	A	0.4606	0.016	0.0017	1.40E-22	109500

ESM Table 4: The impact of genetically predicted reduction in HbA_{1c} (%) instrumented by rs2732480 on type 2 diabetes, coronary artery disease and overall cancer in the UK Biobank

Outcome	**No. of case (%)**	**OR (95% CI)**	**P value**
Type 2 diabetes	26,690 (6.9)	0.11 (0.02 to 0.50)	4.08×10^{-3}
Coronary artery disease	38,098 (9.7)	0.22 (0.06 to 0.81)	0.02
Overall cancer	80,941(20.7)	0.45 (0.17 to 1.14)	0.09

OR, odds ratio; CI, confidence interval

ESM Table 5: The impact of genetically predicted reduction in HbA_{1c} (%) instrumented by AMPK variants on risk of type 2 diabetes, cardiovascular diseases and cancers

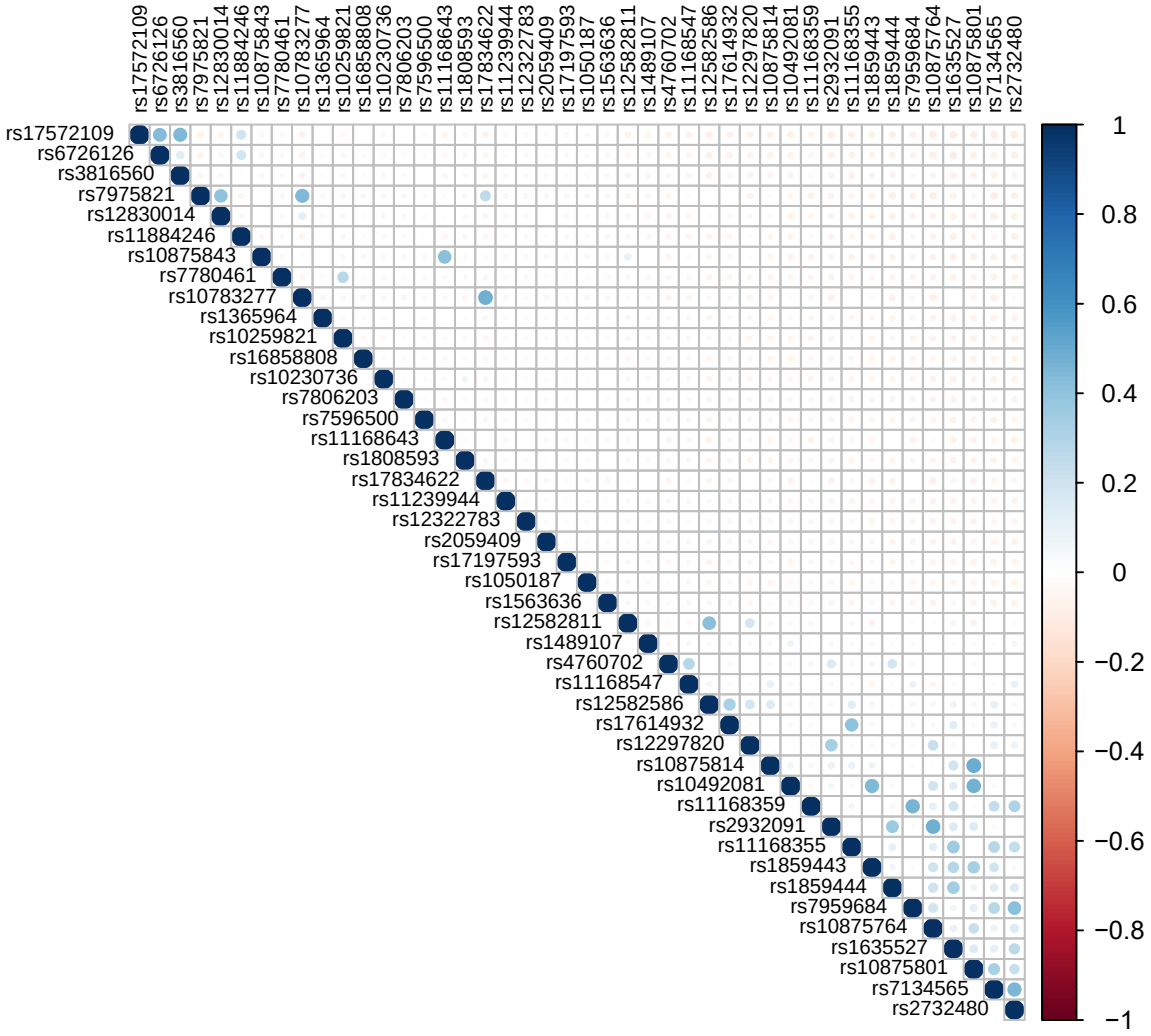
Outcome	Consortium	No. of case	No. of control	No. of variants	Inverse variance weighting			Q statistics	
					OR	95 % CI	P value	Heterogeneity	P value
Type 2 diabetes	DIAGRAM	12,171	56,862	39	0.11	0.04 to 0.35	**1.78×10^{-4}**	123	**6.09×10^{-11}**
Coronary artery disease	CARDIoGRAMplusC4D	60,801	123,504	42	0.48	0.33 to 0.72	**2.89×10^{-4}**	80	**2.45×10^{-4}**
Stroke	MEGASTROKE	40,585	406,111	43	0.95	0.56 to 1.61	0.84	216	**3.07×10^{-25}**
Breast cancer	BCAC	122,977	105,974	15	0.61	0.26 to 1.43	0.25	41	**2.12×10^{-4}**
Prostate cancer	PRACTICAL	79,148	61,106	13	0.85	0.41 to 1.78	0.66	7	0.85

OR, odds ratio, CI, confidence interval.

ESM Table 6: The impact of genetically predicted HbA_{1c} (%) on risk of cardiovascular diseases and cancers in the UK Biobank

Outcome	No. of case (%)	Method	OR (95% CI)	P value
Coronary artery disease	38,098 (9.7)	Inverse variance weighted	1.41 (1.03 to 1.93)	0.03
		Weighted median	1.17 (0.91 to 1.50)	0.23
		MR Egger	1.38 (0.76 to 2.50)	0.29
Stroke	11,358 (2.9)	Inverse variance weighted	1.24 (0.86 to 1.81)	0.25
		Weighted median	1.09 (0.72 to 1.65)	0.67
		MR Egger	0.99 (0.49 to 2.00)	0.98
Overall cancer	80,941(20.7)	Inverse variance weighted	0.84 (0.70 to 1.01)	0.07
		Weighted median	0.80 (0.67 to 0.96)	0.02
		MR Egger	0.77 (0.54 to 1.10)	0.16
Prostate cancer	8970 (2.2)	Inverse variance weighted	0.81 (0.52 to 1.28)	0.37
		Weighted median	0.96 (0.58 to 1.58)	0.87
		MR Egger	1.06 (0.45 to 2.46)	0.90
Breast cancer	9251 (2.4)	Inverse variance weighted	1.01 (0.61 to 1.67)	0.97
		Weighted median	1.41 (0.87 to 2.28)	0.16
		MR Egger	1.46 (0.57 to 3.73)	0.43
Colorectal cancer	5861 (1.5)	Inverse variance weighted	1.09 (0.67 to 1.79)	0.73
		Weighted median	0.82 (0.47 to 1.45)	0.50
		MR Egger	0.96 (0.38 to 2.44)	0.93

OR, odds ratio; CI, confidence interval

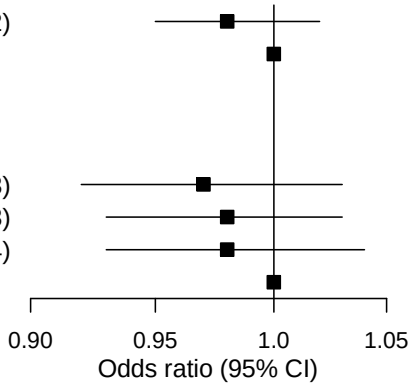


ESM Fig. 1: Linkage disequilibrium matrix for variants included in the AMP-activated protein kinase score

Values represent r^2 values, a measure of linkage disequilibrium. R^2 values range from 0 to1; with 0 representing complete equilibrium and 1 representing complete disequilibrium. Variants were included in the score if they had an r^2 value < 0.3 with other variants included in the score.

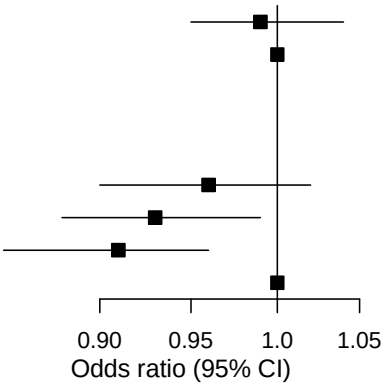
a) Stroke

Group	No. of cases (%)	OR (95% CI)
Primary analysis		
AMPK score above median	5659 (1.4)	0.98 (0.95, 1.02)
AMPK score below median	5699 (1.5)	Reference
Dose−response analysis		
Quartile of AMPK score		
4	2858 (0.7)	0.97 (0.92, 1.03)
3	2801 (0.7)	0.98 (0.93, 1.03)
2	2818 (0.7)	0.98 (0.93, 1.04)
1	2881 (0.7)	Reference



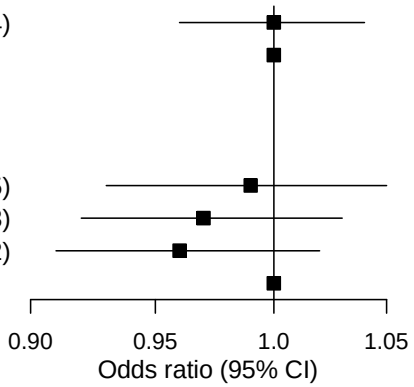
b) Prostate cancer

Group	No. of cases (%)	OR (95% CI)
Primary analysis		
AMPK score above median	4476 (1.1)	0.99 (0.95, 1.04)
AMPK score below median	4494 (1.1)	Reference
Dose−response analysis		
Quartile of AMPK score		
4	2289 (0.6)	0.96 (0.90, 1.02)
3	2187 (0.6)	0.93 (0.88, 0.99)
2	2138 (0.5)	0.91 (0.85, 0.96)
1	2356 (0.6)	Reference

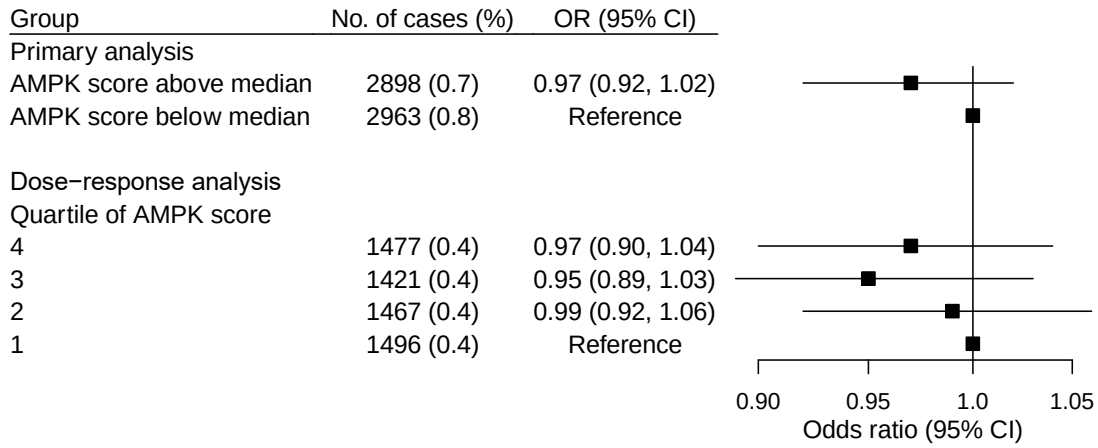


c) Breast cancer

Group	No. of cases (%)	OR (95% CI)
Primary analysis		
AMPK score above median	4656 (1.2)	1.00 (0.96, 1.04)
AMPK score below median	4595 (1.2)	Reference
Dose−response analysis		
Quartile of AMPK score		
4	2373 (0.6)	0.99 (0.93, 1.05)
3	2283 (0.6)	0.97 (0.92, 1.03)
2	2253 (0.6)	0.96 (0.91, 1.02)
1	2342 (0.6)	Reference



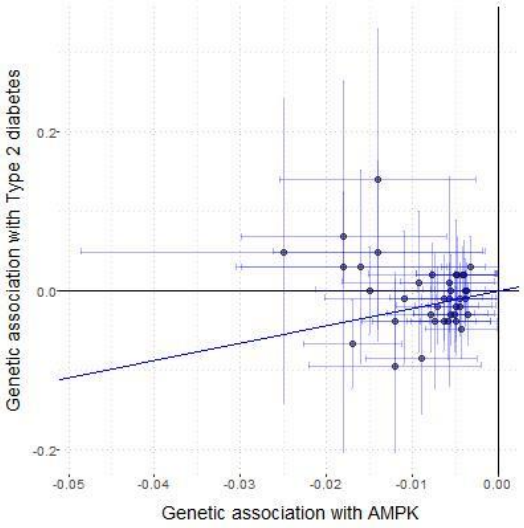
d) Colorectal cancer



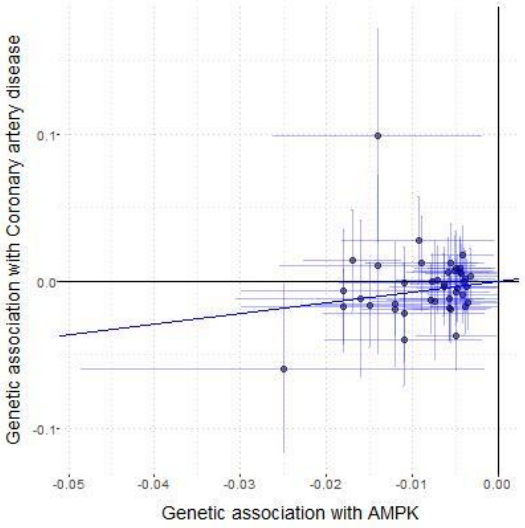
ESM Fig. 2. Association of AMP-activated protein kinase score with risk of a) stroke, b) prostate cancer, c) breast cancer, d) colorectal cancer in the UK Biobank.

Boxes represent odds ratios (OR) and lines represent 95% confidence intervals (CI).

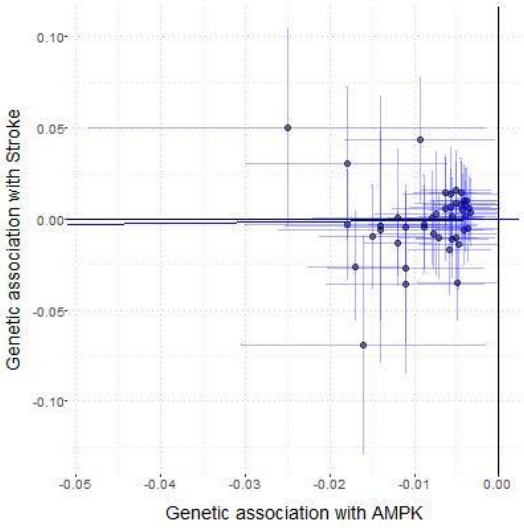
a) Type 2 diabetes



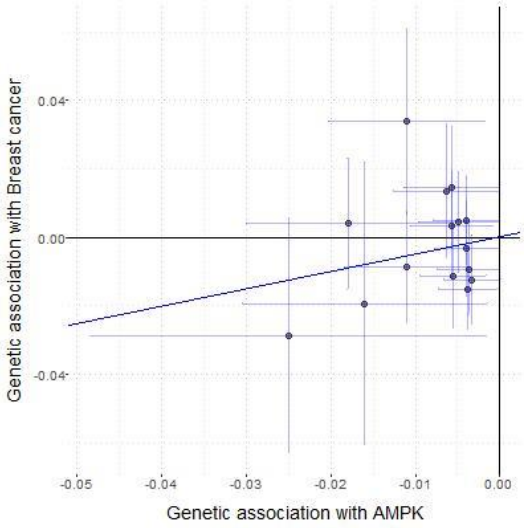
b) Coronary artery disease

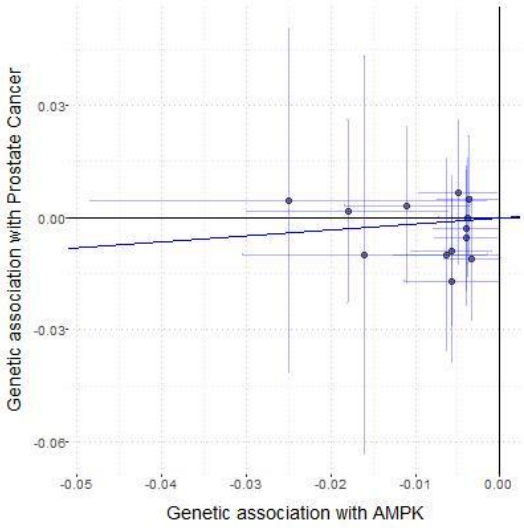


c) Stroke



d) Breast caner





e) Prostate cancer

ESM Fig. 3. Genetic associations instrumented by genetic variants with HbA_{1c} against a) type 2 diabetes, b) coronary artery disease, c) stroke, d) breast cancer, e) prostate cancer.

Genetic associations with HbA_{1c} (per allele decrease in HbA_{1c}) were estimated in participant of European ancestry only and obtained from the Meta-Analyses of Glucose and Insulin-related traits consortium, Wheeler et al (2017);[1] Genetic associations with type 2 diabetes risk (per allele log odds ratio) were obtained from the Diabetes Genetics Replication And Meta-analysis consortium, Morris et al (2012).[2] Genetic associations with coronary artery disease risk (per allele log odds ratio) were obtained the CARDIoGRAMplusC4D 1000 Genomes based genome wide association study, Nikpay et al (2015).[3] Genetic associations with stroke risk (per allele log odds ratio) were obtained from the MEGASTROKE consortium, Malik et al (2018);[4] Genetic associations with breast cancer risk (per allele log odds ratio) were obtained from the Breast Cancer Association Consortium, Michailidou et al (2017).[5] Genetic associations with prostate cancer risk (per allele log odds ratio) were obtained from the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome consortium, Schumacher et al (2018).[6]

Supplementary References

1. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. PLoS Med 2017;14:e1002383.

2. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 2012;44:981-+.

3. Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015;47:1121-30.

4. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 2018;50:524-37.

5. Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. Nature 2017;551:92-+.

6. Schumacher FR, Al Olama AA, Berndt SI, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat Genet 2018;50:928-+.