On-line Supplementary Material

Metabolic risk factors, genetic predisposition, and risk of severe liver disease in Chinese: a prospective study of 0.5 million people

American Journal of Clinical Nutrition

Yuanjie Pang, DPhil[1], Jun Lv, PhD[1], Christiana Kartsonaki, DPhil[2,3], Canqing Yu, PhD[1], Yu Guo, MSc[4], Yiping Chen, DPhil[2,3], Ling Yang, PhD[2,3] Iona Y. Millwood, DPhil[2,3], Robin Walters, PhD[2,3], Sisi Wang, BM[5], Junshi Chen, MD[6], Zhengming Chen, DPhil[2,3], Liming Li, MD[1]

1. Department of Epidemiology & Biostatistics, School of Public Health, Peking University, 38 Xueyuan Road, Beijing 100191, China
2. Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, Big Data Institute Building, Roosevelt Drive, University of Oxford, UK
3. Medical Research Council Population Health Research Unit (MRC PHRU) at the University of Oxford, Nuffield Department of Population Health, University of Oxford, UK
4. Chinese Academy of Medical Sciences, 9 Dongdan San Tiao, Beijing 100730, China
5. Liuzhou Center for Disease Prevention and Control, 1 Tanzhongxi Road, Liuzhou 545007, China
6. National Center for Food Safety Risk Assessment, 37 Guangqu Road, Beijing 100021, China

Address for correspondence:

Prof. Liming Li
Department of Epidemiology and Biostatistics
School of Public Health
Peking University
38 Xueyuan Road
Beijing, 100191, China
Tel: 86-010-8280-1528
Fax: 86-010-82801530
lmleeph@vip.163.com

Table of content

Supplementary methods

Nested case-control study

The nested case-control study involved stroke and coronary heart disease (CHD) subtypes among 18,183 CKB individuals, and included genome-wide array data and standard clinical biochemistry analysis. Cases were identified as those that had an incident fatal or nonfatal event coded as ICD-10: I21-23 for myocardial infarction (MI, n=1273); I63 and I69.3 for ischemic stroke (IS, n=5447); I61 for intracerebral hemorrhage (ICH, n=5150) at the censoring date of 1 January 2015. Cases were selected from among the larger study as those with the youngest age at event. Common controls (n=6313) were frequency matched to cases by age, sex and area. Cases and controls were free of prior vascular disease (including absence of statin therapy) and cancer.

Measurement of clinical chemistry

Eight traits (total cholesterol, LDL-C, HDL-C, TGs, apolipoprotein B, apolipoprotein-A1, albumin and creatinine) were quantified using standard clinical biochemistry assays at the Wolfson Laboratory, CTSU, University of Oxford, UK. Plasma concentrations of total cholesterol, triglycerides, apolipoprotein B and apolipoprotein A1 in EDTA plasma were measured by Beckman-Coulter AU680 clinical chemistry analyzers using manufacturers' reagents, calibrators and settings (Beckman-Coulter, UK), except LDL-C and HDL-C which used N-geneous reagents, calibrators and settings (Genzyme Diagnostics, UK). In the analysis of BMI genetic score and elevated alanine aminotransferase (ALT), inverse probability of sampling weights (i.e. inclusion in the nested case-control study) were developed to ensure that the analyses took account of the inclusion and exclusion criteria

and sampling scheme for a nested case-control study.[1] Cases and controls were assigned different weights to reflect the different proportions of cases and controls from eligible participants in the entire CKB cohort.

Assessment of liver biomarkers

17 biomarkers were measured by standard clinical biochemistry assays in 18,181 participants from a nested case-control study of stroke and CHD (5486 cases of IS, 5067 of ICH, 1008 of MI, 277 of fatal ischemic heart disease, 6343 controls; all free of prior vascular disease and cancer, and not on statin therapy), at the Wolfson Laboratory, CTSU, University of Oxford, UK. The biochemistry measurements included the liver enzymes alanine aminotransferase (ALT), aspartate transaminase (AST), and gamma-glutamyl transferase (GGT), and triglycerides (TGs) used to assess liver steatosis. Liver function was measured by ALT, AST, and GGT. Steatosis was measured by the fatty liver index (FLI) using the following formula:[2]

$$e^{0.953\times log_e TG+0.139\times BMI+0.718\times log_e GGT+0.052\times WC-15.745}/$$

$$(1+e^{0.953\times log_e TG+0.139\times BMI+0.718\times log_e GGT+0.052\times WC-15.745}) \text{ x } 100.$$

Genotyping

Genotyping was conducted using a custom-designed 800K-SNP array (Axiom; Affymetrix) with imputation to 1000 Genomes Phase 3. For this study, Genotype data were available for samples from 100,408 participants passing QC (overall call rate >99.97% across all variants). This included a population-based sample of 75,736 participants randomly selected from the total CKB cohort. The remaining 24,672 participants were genotyped as part of nested case-

control studies of incident stroke, coronary heart disease (CHD), or chronic obstructive pulmonary disease. To avoid potential ascertainment bias, only the 75,736 population-representative subset of participants were used for genetic analyses of hepatobiliary outcomes. All participants with clinical biochemistry measures were genotyped.

Statistical analysis

In the analysis of BMI and liver biomarkers, inverse probability of sampling weights (i.e. inclusion in the nested case-control study) were developed to ensure that our analysis accounted for the inclusion/exclusion criteria and sampling scheme for the nested case-control study.[3] Cases and controls were assigned different weights to reflect the different proportions of cases and controls from eligible participants in the entire CKB cohort. The weights were calculated separately for controls and cases as the number of eligible participants divided by the number selected in the nested case-control study. The weights were 307.35 for controls, 4.47 for MI cases, 27.82 for IS cases, and 6.78 for ICH cases.

References

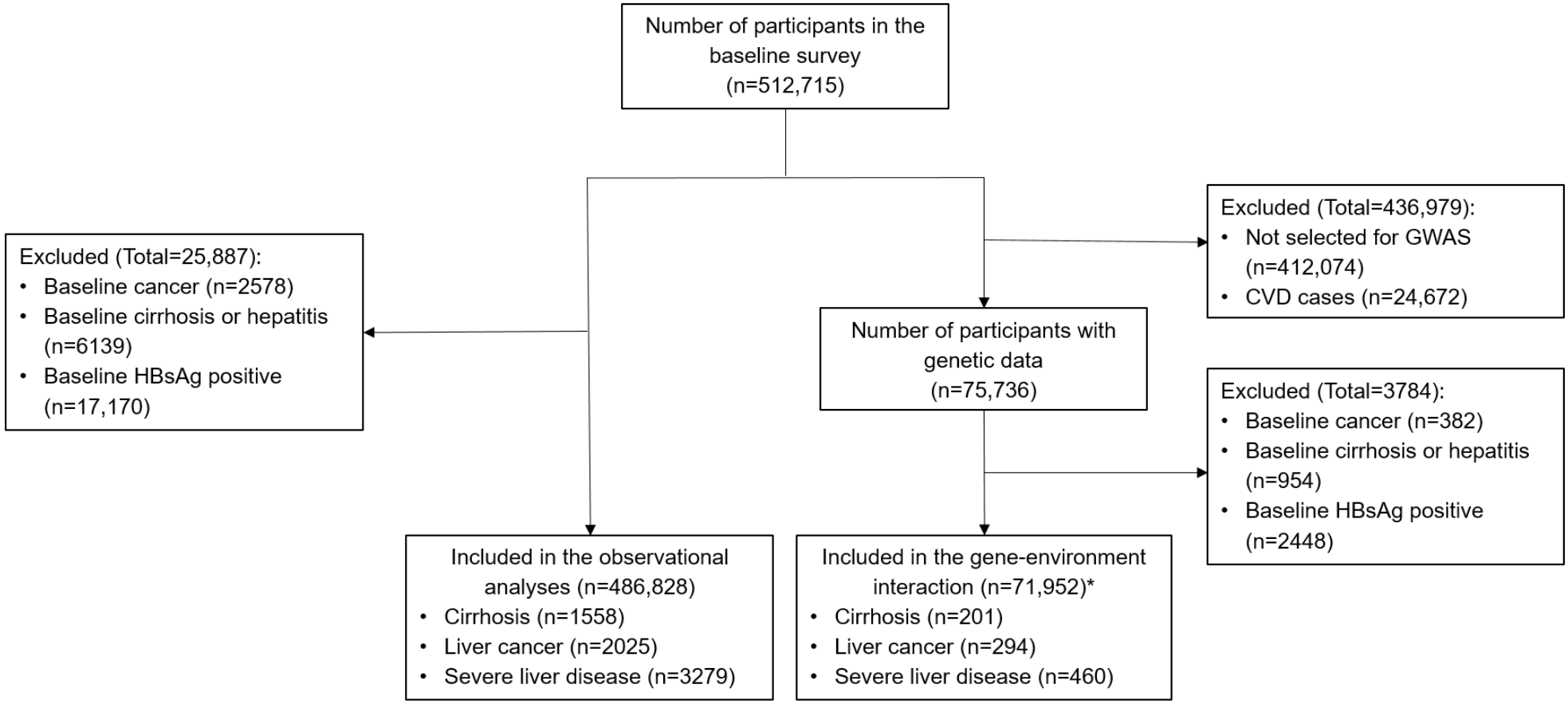
1. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, Li L, China Kadoorie Biobank (CKB) collaborative group. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011;40(6):1652-1666.

2. Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, Tiribelli C. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterol.* 2006;6:33. doi:10.1186/1471-230X-6-33.

3. Pang Y, Kartsonaki C, Du H, Millwood IY, Guo Y, Chen Y, Bian Z, Yang L, Walters R, Bragg

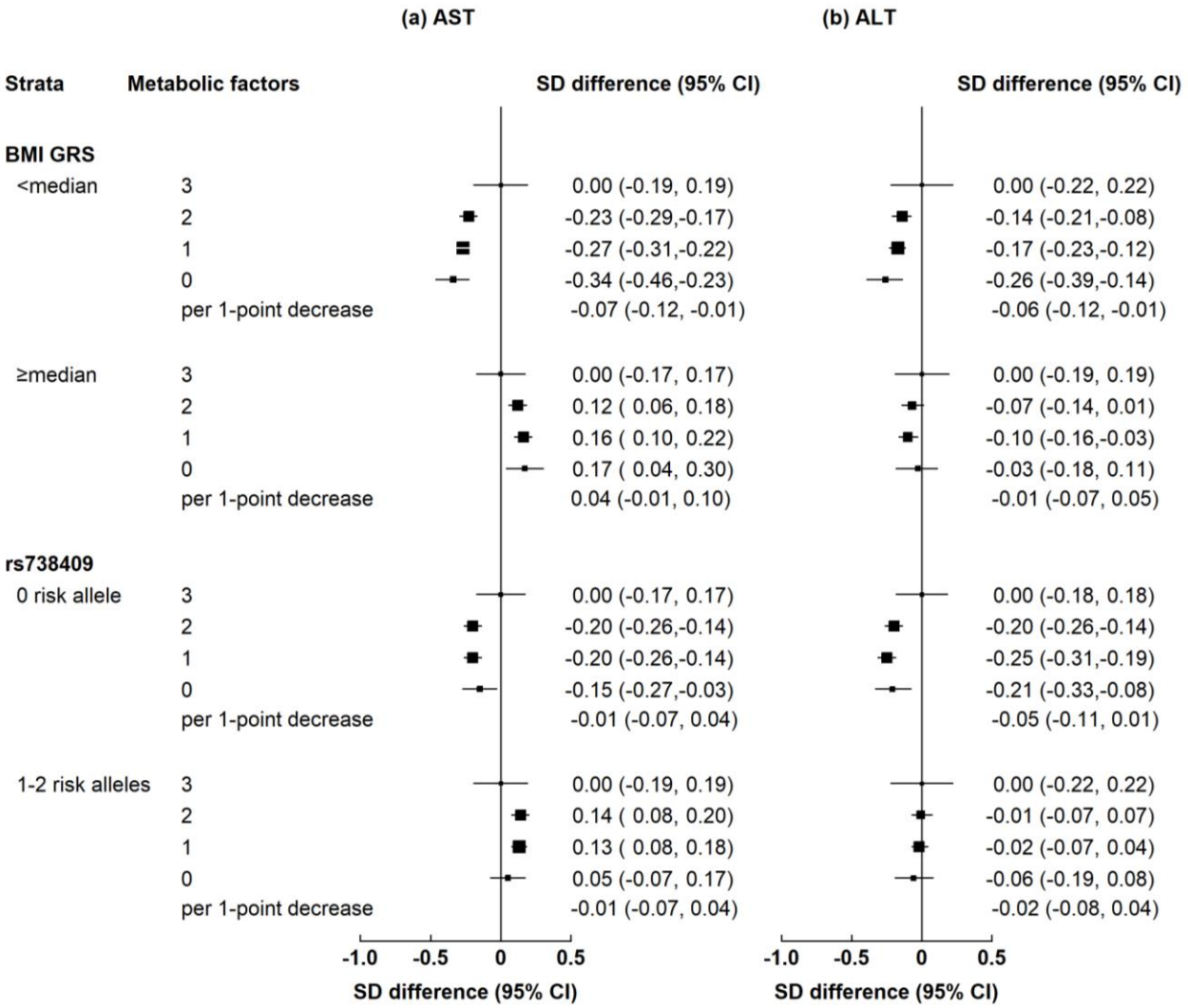
F, et al. Physical activity, sedentary leisure time, circulating metabolic markers, and risk of major vascular diseases. *Circ Genom Precis Med*. 2019; 12(9): 386-396.

Supplementary Figure 1. Flow diagram



*6023 of these participants had biomarker data.

Supplementary Figure 2. Associations between combined metabolic factors and liver biomarkers by genetic risk



Boxes represent the SD differences of liver biomarkers associated with combined metabolic risk factors in participants by strata of BMI GRS and rs738409, separately, with the size of the box inversely proportional to the variance of the SD difference. Linear regression models were used in the analysis. The model was adjusted for age at baseline, age squared, 10 regions, education, smoking, alcohol, and self-rated health. *P*-values for interaction were obtained from the likelihood ratio test. *P*-values for interaction: for AST, BMI GRS 0.17 and rs738409 0.09; for ALT, BMI GRS 0.81 and rs738409 0.14.

Supplementary Table 1. Baseline characteristics by genetic risk

Variable[1]	BMI GRS		rs738409	
	Low risk (n=36,257)	High risk (n=35,695)	Low risk (n=28,740)	High risk (n=43,212)
Age (SD), year	53.8 (11.0)	53.7 (11.0)	53.8 (10.6)	53.7 (8.3)
Female, %	57.3	57.7	57.3	57.6
Socioeconomic and lifestyle factors				
Urban region, %	43.3	43.3	44.8	42.3
≥9 years of education, %	20.4	20.2	20.1	20.4
Household income ≥35 000 RMB/year, %	16.0	16.2	16.0	16.2
Ever regular smoking, %				
Male	68.1	68.9	68.0	68.8
Female	3.1	3.1	3.0	3.1
Weekly drinking, %				
Male	34.4	33.7	34.5	33.9
Female	2.3	2.0	2.2	2.1
Total physical activity (SD), MET-h/day	20.8 (13.7)	20.8 (13.8)	20.8 (13.6)	20.8 (13.8)
Sedentary leisure time (SD), h/day	3.0 (1.5)	3.0 (1.5)	3.0 (1.5)	3.0 (1.5)
Blood pressure and anthropometry				
SBP (SD), mmHg	132.2 (22.7)	133.0 (22.8)	132.7 (22.8)	132.6 (22.7)
RPG (SD), mmol/L	6.0 (2.4)	6.1 (2.6)	6.0 (2.4)	6.1 (2.6)
BMI (SD), kg/m^2	23.5 (3.4)	23.9 (3.5)	23.7 (3.5)	23.7 (3.5)
Waist circumference (SD), cm	79.4 (9.8)	80.6 (10.0)	80.1 (9.9)	80.0 (10.0)
Hip circumference (SD), cm	90.3 (7.0)	91.0 (7.3)	90.7 (7.2)	90.6 (7.1)
Waist-to-hip ratio (SD)	0.88 (0.07)	0.88 (0.07)	0.88 (0.07)	0.88 (0.07)
Percent body fat (SD), %	27.6 (8.5)	28.4 (8.7)	28.1 (8.7)	28.0 (8.6)
Height (SD), cm	158.5 (8.3)	158.5 (8.3)	158.5 (8.4)	158.5 (8.3)
Prior disease history, %				
Diabetes	5.8	6.6	5.9	6.4
Coronary heart disease	2.5	2.7	2.6	2.6
Stroke or TIA	1.2	1.3	1.2	1.2
Hypertension	11.9	12.8	12.4	12.3
Family history of diabetes	4.8	5.0	4.9	4.9
Family history of cancer	13.9	13.7	13.9	13.8

Abbreviations: BMI=body mass index, GRS=genetic risk score, MET=metabolic equivalent of task, RPG=random plasma glucose, SBP=systolic blood pressure, TIA=transient ischemic attack.

[1] Results were standardized by age, sex, and region (where appropriate). Values are means unless otherwise stated.

Supplementary Table 2. Associations between individual metabolic risk factors and risk of severe liver disease

	Severe liver disease	Liver cancer	Cirrhosis
Metabolic factor	**HR (95% CI)**	**HR (95% CI)**	**HR (95% CI)**
Physical activity			
Quintile 1 (<8.72 MET-h/d)	1.00 (0.93, 1.08)	1.00 (0.91, 1.09)	1.00 (0.89, 1.12)
Quintile 2 (8.72-<14.2 MET-h/d)	0.90 (0.84, 0.98)	0.89 (0.81, 0.98)	0.91 (0.81, 1.02)
Quintile 3 (14.2-<21.8 MET-h/d)	0.85 (0.79, 0.92)	0.87 (0.78, 0.96)	0.85 (0.76, 0.96)
Quintile 4 (21.8-<33.2 MET-h/d)	0.94 (0.87, 1.03)	0.96 (0.86, 1.07)	0.90 (0.80, 1.01)
Quintile 5 (≥33.2 MET-h/d)	0.75 (0.68, 0.83)	0.76 (0.67, 0.87)	0.70 (0.61, 0.80)
WHR			
Quintile 1 (M <0.85 / F <0.81)	1.00 (0.91, 1.10)	1.00 (0.88, 1.13)	1.00 (0.87, 1.15)
Quintile 2 (M 0.85-<0.89 / F 0.81-<0.85)	1.10 (1.01, 1.19)	1.09 (0.98, 1.21)	1.12 (0.99, 1.26)
Quintile 3 (M 0.89-<0.92 / F 0.85-<0.88)	1.13 (1.04, 1.23)	1.15 (1.04, 1.28)	1.22 (1.08, 1.38)
Quintile 4 (M 0.92-<0.96 / F 0.88-<0.92)	1.26 (1.17, 1.36)	1.21 (1.10, 1.33)	1.35 (1.21, 1.50)
Quintile 5 (M ≥0.96 / F ≥0.92)	1.44 (1.34, 1.56)	1.32 (1.19, 1.46)	1.74 (1.56, 1.95)
Blood glucose			
RPG ≤5.5 mmol/L, no diabetes	1.00 (0.94, 1.06)	1.00 (0.93, 1.07)	1.00 (0.92, 1.09)
RPG 5.5-<6.5 mmol/L, no diabetes	1.12 (1.04, 1.19)	1.06 (0.97, 1.15)	1.18 (1.07, 1.30)
RPG ≥6.5 mmol/L, no diabetes	1.27 (1.18, 1.36)	1.14 (1.04, 1.25)	1.48 (1.33, 1.64)
Diabetes	1.65 (1.47, 1.84)	1.49 (1.29, 1.72)	2.04 (1.75, 2.39)

Abbreviations: WHR=waist-to-hip ratio, MET=metabolic equivalent of task, RPG=random plasma glucose.

Cox proportional hazards regression models were used in the analysis. The model was adjusted for age at baseline, age squared, 10 regions, education, smoking, alcohol, and self-rated health.

Supplementary Table 3. Joint associations between combined metabolic factors and genetic factors on severe liver disease

	No. cases	SLD HR (95% CI)	No. cases	Cirrhosis HR (95% CI)	No. cases	Liver cancer HR (95% CI)
BMI GRS						
High GRS & 3 factors	24	1.00 (0.66, 1.51)	16	1.00 (0.60, 1.66)	11	1.00 (0.54, 1.84)
High GRS & 2 factors	116	0.57 (0.47, 0.69)	75	0.57 (0.45, 0.72)	51	0.51 (0.39, 0.68)
High GRS & 0-1 factor	86	0.41 (0.32, 0.51)	47	0.39 (0.29, 0.52)	42	0.33 (0.24, 0.46)
Low GRS & 3 factors	16	0.74 (0.45, 1.22)	10	0.69 (0.37, 1.29)	6	0.60 (0.27, 1.35)
Low GRS & 2 factors	111	0.55 (0.45, 0.66)	75	0.57 (0.45, 0.71)	47	0.47 (0.35, 0.63)
Low GRS & 0-1 factor	107	0.44 (0.36, 0.54)	71	0.51 (0.39, 0.66)	44	0.30 (0.22, 0.41)
rs738409						
High GRS & 3 factors	27	1.00 (0.68, 1.48)	18	1.00 (0.62, 1.61)	12	1.00 (0.56, 1.79)
High GRS & 2 factors	138	0.60 (0.50, 0.71)	96	0.65 (0.53, 0.79)	58	0.52 (0.40, 0.68)
High GRS & 0-1 factor	115	0.43 (0.35, 0.53)	70	0.47 (0.36, 0.60)	54	0.35 (0.26, 0.47)
Low GRS & 3 factors	13	0.77 (0.45, 1.34)	8	0.72 (0.36, 1.46)	5	0.66 (0.27, 1.60)
Low GRS & 2 factors	89	0.56 (0.46, 0.70)	54	0.52 (0.40, 0.69)	40	0.54 (0.39, 0.74)
Low GRS & 0-1 factor	78	0.45 (0.36, 0.57)	48	0.49 (0.36, 0.65)	32	0.33 (0.23, 0.47)

Cox proportional hazards regression models were used in the analysis. The model was adjusted for age at baseline, age squared, 10 regions, education, smoking, alcohol, and self-rated health.

Supplementary Table 4. Association between combined metabolic risk factors and liver biomarkers by genetic risk

	Metabolic score	FLI	AST	ALT	GGT
		Effect size (95% CI)	Effect size (95% CI)	Effect size (95% CI)	Effect size (95% CI)
BMI GRS					
<median	0	0.00 (-0.14, 0.14)	0.00 (-0.19, 0.19)	0.00 (-0.22, 0.22)	0.00 (-0.13, 0.13)
	1	-0.31 (-0.36, -0.27)	-0.23 (-0.29, -0.17)	-0.14 (-0.21, -0.08)	-0.11 (-0.15, -0.07)
	2	-0.56 (-0.59, -0.52)	-0.27 (-0.31, -0.22)	-0.17 (-0.23, -0.12)	-0.19 (-0.23, -0.16)
	3	-0.76 (-0.84, -0.68)	-0.34 (-0.46, -0.23)	-0.26 (-0.39, -0.14)	-0.27 (-0.34, -0.19)
	per 1-point decrease	*-0.24 (-0.28, -0.20)*	*-0.07 (-0.12, -0.01)*	*-0.06 (-0.12, 0.00)*	*-0.08 (-0.12, -0.04)*
⩾median	0	0.00 (-0.13, 0.13)	0.00 (-0.17, 0.17)	0.00 (-0.19, 0.19)	0.00 (-0.12, 0.12)
	1	-0.44 (-0.49, -0.40)	0.12 (0.06, 0.18)	-0.07 (-0.14, 0.01)	-0.37 (-0.41, -0.33)
	2	-0.68 (-0.72, -0.64)	0.16 (0.10, 0.22)	-0.10 (-0.16, -0.03)	-0.40 (-0.44, -0.37)
	3	-0.84 (-0.93, -0.75)	0.17 (0.04, 0.30)	-0.03 (-0.18, 0.11)	-0.47 (-0.55, -0.39)
	per 1-point decrease	*-0.24 (-0.28, -0.20)*	*0.04 (-0.01, 0.10)*	*-0.01 (-0.07, 0.05)*	*-0.09 (-0.13, -0.05)*
	P for interaction	*0.90*	*0.17*	*0.81*	*0.65*
rs738409					
0 risk allele	0	0.00 (-0.13, 0.13)	0.00 (-0.17, 0.17)	0.00 (-0.18, 0.18)	0.00 (-0.12, 0.12)
	1	-0.43 (-0.47, -0.38)	-0.20 (-0.26, -0.14)	-0.20 (-0.26, -0.14)	-0.33 (-0.37, -0.29)
	2	-0.67 (-0.72, -0.63)	-0.20 (-0.26, -0.14)	-0.25 (-0.31, -0.19)	-0.36 (-0.40, -0.32)
	3	-0.91 (-1.00, -0.81)	-0.15 (-0.27, -0.03)	-0.21 (-0.33, -0.08)	-0.49 (-0.57, -0.41)
	per 1-point decrease	*-0.27 (-0.31, -0.23)*	*-0.01 (-0.07, 0.04)*	*-0.05 (-0.11, 0.01)*	*-0.10 (-0.14, -0.06)*
1-2 risk alleles	0	0.00 (-0.13, 0.13)	0.00 (-0.19, 0.19)	0.00 (-0.22, 0.22)	0.00 (-0.13, 0.13)
	1	-0.31 (-0.35, -0.27)	0.14 (0.08, 0.20)	0.01 (-0.07, 0.07)	-0.17 (-0.21, -0.14)
	2	-0.55 (-0.59, -0.52)	0.13 (0.08, 0.18)	-0.02 (-0.07, 0.04)	-0.26 (-0.29, -0.23)
	3	-0.70 (-0.78, -0.62)	0.05 (-0.07, 0.17)	-0.06 (-0.19, 0.08)	-0.29 (-0.37, -0.21)
	per 1-point decrease	*-0.22 (-0.26, -0.19)*	*-0.01 (-0.07, 0.04)*	*-0.02 (-0.08, 0.04)*	*-0.08 (-0.12, -0.04)*
	P for interaction	*0.02*	*0.09*	*0.14*	*0.62*

Linear regression models were used in the analysis. The model was adjusted for age at baseline, age squared, 10 regions, education, smoking, alcohol, and self-rated health.

Supplementary Table 5. Population attributable fractions according to individual metabolic factor and the combined metabolic score in all participants

Metabolic factor	Severe liver disease	
	Proportion of cases with risk factor, n (%)	PAF, %
Central adiposity	2059 (62.8)	6.6
Physical activity	2828 (86.3)	3.7
Diabetes	2949 (90.0)	23.7
Metabolic factors ≥2	1391 (42.4)	10.0

Abbreviation: PAF, population attributable fraction.