

## Supplementary Material

### Table of contents

eMethods.....	3
Supplementary Figure 1. Flow chart.....	8
Supplementary Table 1. Hazard ratios of incident SLD according to lifestyle categories.....	10
Supplementary Table 2. Hazard ratios of incident SLD according to healthy lifestyle categories within each genetic risk category.....	11
Supplementary Table 3. Hazard ratios of incident SLD according to genetic and lifestyle risk.....	12
Supplementary Table 4. Hazard ratios of liver biomarkers according to healthy lifestyle categories within each genetic risk category.....	13

### eMethods

#### *Data collection on lifestyle factors*

Individuals were classified by smoking status as never, occasional, former regular, or current regular smokers. Never smokers were defined as those who reported not smoking at baseline and had smoked <100 cigarettes (or equivalent) in their lifetime. Former regular smokers were defined as those who had smoked  $\geq 100$  cigarettes (or equivalent) but had quit smoking by choice for  $\geq 6$  months before baseline. Occasional smokers were defined as those who did not meet the criteria for never smokers, and who had not stopped smoking completely for  $\geq 6$  months before baseline. Current regular smokers were defined as those who reported having ever smoked  $\geq 1$  cigarette (or equivalent) daily for at least 6 months. Approximately 50% of former regular smokers stopped smoking because of physical illness that they already had, and they were still considered as current regular smokers in the main analyses.

The information on alcohol drinking included whether the participant had drunk alcohol regularly (i.e., drank at least once a week on a regular basis) during the past year, and if so, the age at which drinking began, the type (beer, wine, or spirits) and the amount consumed on a typical drinking day. Abstainers were defined as those who had never or almost never drunk alcohol in the past year and had not drunk weekly in the past. Occasional drinkers were defined as those who in the past year had drunk alcohol occasionally, during certain seasons, or monthly but less than weekly, and had

not drunk weekly in the past. Reduced-intake drinkers were those who in the past year had drunk alcohol occasionally, during certain seasons, or monthly but less than weekly, but had drunk weekly in the past. Ex-weekly drinkers were those who had drunk weekly in the past but had never or almost never drunk alcohol in the past year. Weekly drinkers were those who usually drank at least once a week during the past year.

Participants were asked about their usual type and duration of activities related to work, commuting, household chores, and leisure-time exercise during the past year. To quantify the amount of physical activity, metabolic equivalent of tasks (METs) from the 2011 update of a major compendium of physical activities were used.<sup>[1]</sup> The MET value for a particular type of physical activity represents the ratio of the energy expended per kilogram of body weight per hour during that activity to that expended when sitting quietly. The number of hours spent per day participating in each activity was multiplied by the MET score for that activity, and the daily amount of total physical activity was obtained by summing the MET-hours for activities related to occupation and non-occupational activities (i.e., commuting, housework, and non-sedentary leisure-time activities).

Standing height was measured using a stadiometer. Weight was measured using a body composition analyzer (TANITA-TBF-300GS; Tanita Corporation), with subtraction of weight of clothing according to season (ranging from 0.5 kg in summer to 2.0–2.5 kg in winter). Waist circumference (WC) and hip circumference (HC) were measured using a soft non-stretchable tape, with HC measured at the maximum circumference around the buttocks. BMI at baseline was calculated as the measured weight in kilograms divided by the square of the measured height in meters. WHR is the ratio of WC to HC.

### ***Nested case-control study***

The nested case-control study involved stroke and coronary heart disease (CHD) subtypes among 18,183 China Kadoorie Biobank (CKB) individuals, and included genome-wide array data and standard clinical biochemistry analysis. Cases were identified as those that had an incident fatal or non-fatal event coded as ICD-10: I21-23 for myocardial infarction (MI,  $n = 1273$ ); I63 and I69.3 for ischemic stroke (IS,  $n = 5447$ ); and I61 for intracerebral hemorrhage (ICH,  $n = 5150$ ) at the censoring date of January 1, 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, triglycerides [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, TGs, 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), an inverse probability of sampling weights (i.e., inclusion in the nested case-control study) was developed to ensure that the analyses considered the inclusion and exclusion criteria and sampling scheme for a nested case-control study.<sup>[2]</sup> 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 80.0 for controls, 4.1 for MI cases, 5.8 for IS cases, and 1.6 for ICH cases.

### ***Liver biomarkers***

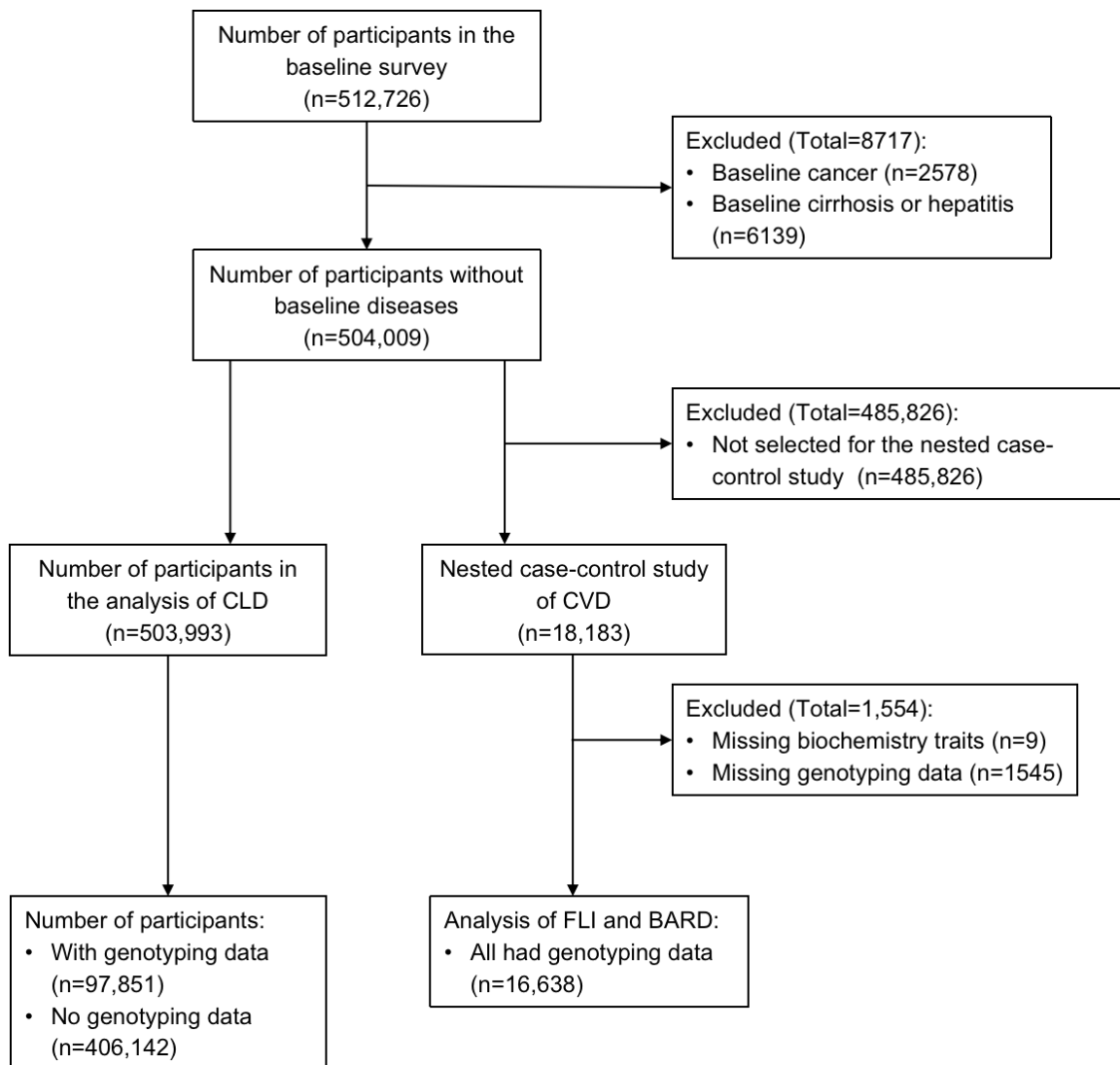
A total of 17 biomarkers were measured by standard clinical biochemistry assays in a nested case-control study of stroke and CHD subtypes among 18,183 CKB individuals with genome-wide array data, including ALT, aspartate transaminase (AST),  $\gamma$ -glutamyl transferase (GGT), total cholesterol, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and TGs.

### ***Statistical analysis***

In the analysis of healthy lifestyle factors and liver biomarkers, an inverse probability of sampling weights (i.e., inclusion in the nested case-control study) was developed to ensure that our analysis accounted for the inclusion/exclusion criteria and sampling scheme for the nested case-control study. 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. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, *et al.* 2011 compendium of physical activities: A second update of codes and MET values. *Med Sci Sports Exerc* 2011;43:1575–1581. doi: 10.1249/MSS.0b013e31821ece12.
2. Pang Y, Kartsonaki C, Du H, Millwood IY, Guo Y, Chen Y, *et al.* Physical activity, sedentary leisure time, circulating metabolic markers, and risk of major vascular diseases. *Circ Genom Precis Med* 2019;12:386–396. doi: 10.1161/CIRCGEN.118.002527.



Supplementary Figure 1: Flow chart.

**Supplementary Table 1: HR of incident SLD according to lifestyle categories.**

Healthy lifestyle category	No. of participants	No. of cases	HR (95% CI)*	HR (95% CI)†
----------------------------	---------------------	--------------	--------------	--------------

			1.00 (0.94,	1.00 (0.88,
0–1	109,337	1619	1.06)	1.14)
2 vs. 0–1			0.88 (0.85,	0.87 (0.79,
	196,608	2190	0.92)	0.95)
3 vs. 0–1			0.74 (0.69,	0.77 (0.66,
	164,553	1238	0.79)	0.88)
4 vs. 0–1			0.56 (0.48,	0.64 (0.45,
	42,228	194	0.65)	0.89)
Per 1 unit			0.85 (0.82,	0.87 (0.80,
			0.88)	0.95)

GWAS: Genome-wide association studies; HBV: Hepatitis B virus; HR: Hazard ratio; SLD: Severe liver disease.

\*The model was adjusted for age, sex, region, HBV, education, and self-rated health.

†The model was restricted to participants with GWAS data and additionally adjusted for 12 national PCs and GRS.

**Supplementary Table 2: HRs of incident SLD according to healthy lifestyle categories within each genetic risk category.**

Subgroup	Liver cancer		HR (95% CI)*	Cirrhosis	
	No. of participants	No. of cases		No. of cases	HR (95% CI)*
Low HBV-GRS					
0–1	12,071	98	1.00 (0.78, 1.27)	72	1.00 (0.74, 1.34)
2 vs. 0–1	19,677	136	0.96 (0.81, 1.13)	91	0.79 (0.64, 0.97)
3 vs. 0–1	15,298	54	0.66 (0.49, 0.88)	51	0.69 (0.51, 0.94)
4 vs. 0–1	3741	9	0.53 (0.26, 1.08)	10	0.54 (0.26, 1.10)
Per 1 unit			0.82 (0.69, 0.96)		0.83 (0.68, 0.99)
High HBV-GRS					
0–1	11,721	121	1.00 (0.80, 1.25)	103	1.00 (0.77, 1.29)
2 vs. 0–1	19,106	133	0.85 (0.72, 1.00)	116	0.81 (0.68, 0.97)

3 vs. 0–1	14,504	79	0.85 (0.66, 1.09)	80	0.81 (0.63, 1.05)
4 vs. 0–1	3521	16	0.95 (0.56, 1.60)	10	0.39 (0.18, 0.82)
Per 1 unit			0.95 (0.82, 1.10)		0.85 (0.72, 1.01)

HBV: Hepatitis B virus; HBV-GRS: Hepatitis B virus genetic risk score; HRs: Hazard ratios; SLD: Severe liver disease.

*P*-value for interaction was 0.46 for liver cancer and 0.73 for SLD.

\*The model was adjusted for age, sex, region, HBV, education, self-rated health, and 12 national PCs.

**Supplementary Table 3: HRs of incident SLD according to genetic and lifestyle risk.**

Subgroup	No. of participants	No. of cases	HR (95% CI)*
Low HBV-GRS			
0–1	12,071	149	1.00 (0.84, 1.20)
2 vs. 0–1	19,677	212	0.96 (0.83, 1.11)
3 vs. 0–1	15,298	95	0.70 (0.57, 0.87)
4 vs. 0–1	3741	18	0.63 (0.38, 1.03)
High HBV-GRS			
0-1	11,721	201	1.27 (1.08, 1.49)
2 vs. 0–1	19,106	222	1.02 (0.88, 1.17)
3 vs. 0–1	14,504	143	1.04 (0.87, 1.25)
4 vs. 0–1	3521	23	0.82 (0.52, 1.28)

HBV: Hepatitis B virus; HBV-GRS: Hepatitis B virus genetic risk score; HR: Hazard ratio; SLD: Severe liver disease.

\*The model was adjusted for age, sex, region, HBV, education, self-rated health, and 12 national PCs.

**Supplementary Table 4: HRs of liver biomarkers according to healthy lifestyle categories within each genetic risk category.**

Subgroup	FLI	BARD	AST	ALT	GGT
	SD difference (95% CI)*	Unit difference (95% CI)*	SD difference (95% CI)*	SD difference (95% CI)*	SD difference (95% CI)*
Low HBV-					

GRS					
0–1	0 (-0.03, 0.03)	0 (-0.03, 0.03)	0 (-0.03, 0.03)	0 (-0.04, 0.04)	0 (-0.03, 0.03)
2 vs. 0–1	-0.24 (-0.26, -0.22)	-0.17 (-0.19, -0.15)	-0.06 (-0.08, -0.03)	-0.12 (-0.14, -0.09)	-0.15 (-0.17, -0.13)
3 vs. 0–1	-0.38 (-0.41, -0.36)	-0.74 (-0.79, -0.70)	-0.02 (-0.05, -0.1)	-0.06 (-0.10, -0.03)	-0.21 (-0.23, -0.18)
4 vs. 0–1	-0.51 (-0.56, -0.46)	-0.96 (-1.06, -0.85)	-0.08 (-0.14, -0.02)	-0.19 (-0.27, -0.12)	-0.27 (-0.32, -0.22)
High HBV-GRS					
0–1	0 (-0.03, 0.03)	0 (-0.03, 0.03)	0 (-0.04, 0.04)	0 (-0.04, 0.04)	0 (-0.04, 0.04)
2 vs. 0–1	-0.23 (-0.24, -0.21)	-0.09 (-0.12, -0.07)	-0.01 (-0.04, -0.2)	-0.03 (-0.06, -0.0)	-0.11 (-0.13, -0.08)
3 vs. 0–1	-0.32 (-0.34, -0.29)	-0.14 (-0.18, -0.10)	-0.06 (-0.10, -0.03)	-0.12 (-0.16, -0.08)	-0.15 (-0.18, -0.12)
4 vs. 0–1	-0.39 (-0.45, -0.34)	-0.31 (-0.41, -0.21)	-0.08 (-0.16, -0.01)	-0.10 (-0.18, -0.01)	-0.22 (-0.29, -0.16)
<i>P</i> -value for interaction	0.11	<0.001	0.34	0.90	<0.001

ALT: Alanine aminotransferase; AST: Aspartate transaminase; FLI: Fatty liver index; GGT:  $\gamma$ -glutamyl transferase; HBV: Hepatitis B virus; HBV-GRS: Hepatitis B virus genetic risk score; HRs: Hazard ratios. \*The model was adjusted for age, sex, region, HBV, education, self-rated health, 12 national PCs, and fasting hours.