Association between metabolic dysfunction-associated steatotic liver disease and risk of cardiovascular disease

BAEK, JI WOO

Department of Public Health
Graduate School
Yonsei University

Association between metabolic dysfunction-associated steatotic liver disease and risk of cardiovascular disease

Advisor Kimm, Heejin

A Dissertation Submitted
to the Department of Public Health
and the Committee on Graduate School
of Yonsei University in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Public Health

BAEK, JI WOO

June 2025

Association between metabolic dysfunction-associated steatotic liver disease and risk of cardiovascular disease

This Certifies that the Dissertation of BAEK, JI WOO is Approved

Committee Chair	**Kimm, Heejin**
Committee Member	**Jee, Sun Ha**
Committee Member	**Jung, Keum Ji**
Committee Member	**Lee, Hokyou**
Committee Member	**Ha, Min Jin**

Department of Public Health
Graduate School
Yonsei University
June 2025

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

Association between metabolic dysfunction -associated steatotic liver disease and risk of cardiovascular disease

BAEK, JI WOO

Department of Public Health
The Graduate School
Yonsei University

(Directed by professor Heejin Kimm)

Background: Metabolic dysfunction-associated steatotic liver disease (MASLD) is a common liver condition associated with obesity and metabolic syndrome. Although MASLD has been associated with increased cardiovascular disease (CVD) risk, the causal nature of this association remains uncertain owing to the limitations of observational studies. Although genetic approaches such as genome-wide association studies (GWAS) and Mendelian randomization (MR) have been applied using large biobanks, the findings have been inconsistent. This study investigated the association between MASLD and CVD through observational and genetic analyses, utilizing individual-level data from Korean and UK populations, as well as summary-level data from Japan.

Methods: A multistep analytical strategy using data from the Korean Cancer Prevention Study (KCPS-II) cohort was applied. Steatotic Liver Disease (SLD) was defined using a fatty liver index (FLI) threshold of 30 or greater. First, a prospective observational analysis was performed using the Cox proportional hazards model to evaluate the association between MASLD and the incidence of CVD and its subtypes. Second, one-sample and bidirectional two-sample MR analyses were conducted to assess the causal effect of the MASLD on CVD using large-scale cohort data from the KCPS-II, Biobank Japan (BBJ), and the UK Biobank (UKB). Third, GWAS was conducted to identify MASLD-related genetic variants, followed by gene-based and tissue-specific expression analyses. Fourth, gene–smoking interaction analyses were performed to investigate how smoking modifies the genetic effects of MASLD.

Results: In total, 111,637 participants were included (median age 39, 35.7% women). At baseline, 32,018 (28.7%) patients were diagnosed with MASLD. During the median 10.0-

year follow-up, 3,926 incident CVD events occurred. Multivariable-adjusted hazard ratio (HR) of CVD was 1.69 (95% confidence interval (CI) 1.57–1.82) for MASLD. In the one-sample MR analysis using individual-level data from KCPS-II, genetically predicted MASLD was associated with a modest but statistically significant increase in overall CVD risk (HR 1.05, 95% CI 1.03–1.08) in the fully adjusted model.

To further examine causality, two-sample MR analyses were conducted using summary-level data from BBJ and individual-level data from UKB. The inverse variance-weighted (IVW) method demonstrated a significant positive association between MASLD and coronary artery disease (CAD). The odds ratio (OR) was 1.08 (95% CI 1.05–1.13, $p = 1.83 \times 10^{-5}$) when using BBJ outcomes and 1.04 (95% CI 1.01–1.08, $p = 6.03 \times 10^{-3}$) based on UKB outcomes. In the reverse direction, genetic liability to CVD derived from the KCPS-II cohort was also significantly associated with increased MASLD risk when MASLD was defined in UKB (OR 1.15, 95% CI 1.08−1.24, $p = 4.39 \times 10^{-5}$).

GWAS using KCPS-II data identified multiple loci associated with MASLD. GWAS identified numerous loci associated with MASLD, including FTO and CUX2. Expression profiling revealed liver-specific enrichment of GCKR, LIPC, APOA5, APOA4, APOA1, APOC3, APOE, and HNF1A, whereas RPH3A was enriched in brain tissue—functional enrichment analyses implicated lipid metabolism and coronary disease-related pathways. Variance decomposition revealed that MASLD had a total heritability of 38.6%, with 6.5% attributed to genetic factors, 5.3% to gene–smoking interactions, and 26.7% to environmental noise components. In gene–environment interaction analyses, several loci, including rs671 in ALDH2, exhibited enhanced effects in the presence of smoking.

Conclusion: This study provided comprehensive evidence linking MASLD to an increased risk of CVD through both observational and genetic analyses. The MR findings support the potential causal role of MASLD in the development of CVD. The identification of MASLD-associated loci and gene–environment interactions highlights the complex genetic architecture and modifying effects of smoking. These results underscore the importance of metabolic liver health in CVD prevention and suggest that MASLD may represent a modifiable target for reducing the cardiovascular burden.

Keywords: Prospective cohort study, genetic study, Mendelian randomization, metabolic dysfunction-associated steatotic liver disease, cardiovascular disease

I. INTRODIUCTION

1. Study background

Metabolic dysfunction-associated steatotic liver disease (MASLD) is a chronic liver disease marked by hepatic fat accumulation that occurs without excessive alcohol intake or other identifiable liver diseases.[1] In recent years, the prevalence of MASLD has increased rapidly, driven mainly by the global increase in obesity and metabolic syndrome. The global prevalence of MASLD is estimated to be approximately 30%, with an average annual growth of approximately 1% over the past three decades.[2,3] In 2019, the estimated global incidence of MASLD was 4.9%, increasing from 3.7% in 1994–2006 to 5.9% in 2010–2014, representing an approximate increase of 60%.[4]

Nonalcoholic fatty liver disease (NAFLD), a condition strongly associated with metabolic dysfunction, has recently been reclassified and renamed MASLD. The traditional term 'NAFLD' has several limitations, such as the need to exclude other potential causes during diagnosis and the potentially stigmatizing connotation of the word 'fatty'.[5] To address these concerns, an international expert panel proposed the term metabolic dysfunction-associated fatty liver disease (MAFLD), which emphasizes the inclusion of metabolic dysfunction in diagnosis, rather than excluding other etiologies as in NAFLD. However, the MAFLD classification has also faced criticism for excluding NAFLD patients with a lower metabolic risk profile and for failing to account for dual etiologies.[6] In response, an international expert panel representing the American Association for the

Study of Liver Diseases (AASLD), the European Association for the Study of the Liver (EASL), and the Latin American Association for the Study of the Liver (ALEH) proposed a new nomenclature: MASLD.[6]

MASLD, a representative subtype of steatotic liver disease (SLD), is defined by the presence of hepatic steatosis along with at least one cardiometabolic risk factor, while also meeting exclusion criteria for significant alcohol intake and other liver diseases.[7] Although liver biopsy remains the gold standard for diagnosing hepatic steatosis, its invasive nature, high cost, potential risks, and patient discomfort limit its use in routine clinical settings.[8] In clinical practice, hepatic fat accumulation is commonly assessed through imaging techniques such as magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS), with MRI considered the gold standard due to its high accuracy in quantifying liver fat content.[9] Magnetic resonance imaging-derived proton density fat fraction (MRI-PDFF) offers quantitative data and is particularly useful for evaluating the degree of steatosis and disease progression. However, the high cost and technical complexity of MRI-PDFF limit its widespread application in large-scale cohort studies and clinical research.[10] Consequently, non-invasive scoring systems (NSS) have been developed to identify patients with steatotic liver disease.[11] Commonly used indices for assessing the likelihood of NAFLD include the fatty liver index (FLI), hepatic steatosis index (HSI), lipid accumulation product (LAP), and Framingham steatosis index (FSI).[12] The FLI is a validated and widely accepted marker recommended by the European Association for the Study of the Liver (EASL), the European Association for the Study of

Diabetes (EASD), the European Association for the Study of Obesity (EASO), and the Asia-Pacific Association for the Study of the Liver (APASL).[13] The FLI is a simple and non-invasive index calculated using waist circumference, body mass index (BMI), triglyceride levels, and gamma-glutamyl transferase (GGT), making it highly applicable in both large-scale cohort studies and clinical practice.[14] In addition to MASLD, the broader category of SLD includes subtypes such as metabolic and alcohol-related liver disease (MetALD), alcohol-related liver disease (ALD), and other, which accounts for liver disease with mixed or unidentified etiologies.[6]

MASLD encompasses a wide range of pathological stages, from asymptomatic hepatic steatosis to metabolically dysfunctional steatohepatitis, and potentially progressing to liver cirrhosis and hepatocellular carcinoma.[3,15] In addition to its impact on the liver, MASLD is closely associated with an increased risk of cardiovascular and metabolic disorders, including myocardial infarction, stroke, coronary artery disease, and type 2 diabetes mellitus (T2DM).[16] These associations can be attributed to pathophysiological mechanisms commonly observed in MASLD, such as chronic inflammation, insulin resistance, and dysregulated lipid metabolism.[17] Several observational studies have reported an increased risk of cardiovascular disease (CVD) among individuals with MASLD. In Korean cohort studies, participants with MASLD had a significantly higher risk of developing CVD than those without MASLD, even after adjusting for multiple confounders.[18,19] However, given the observational nature of these studies, it was difficult to eliminate the influence of selection bias and confounding factors. Although randomized controlled trials (RCTs) are

considered the gold standard for establishing causal associations, they are often impractical or infeasible in this context.[20]

With advances in genomic analysis technologies, genome-wide association studies (GWAS) have been used to uncover the associations between various diseases and genetic variants.[21] GWAS is now an essential tool for investigating how specific genetic variants influence disease onset and progression by utilizing large-scale, population-based data. GWAS have identified several genetic variants involved in the regulation of hepatic lipid metabolism and inflammatory responses in fatty liver disease. Notable examples include PNPLA3, TM6SF2, MBOAT7, and HSD17B13, which have been reported to influence the development and progression of fatty liver disease, as well as the risk of liver fibrosis and hepatocellular carcinoma.[22] To further elucidate the biological mechanisms by which these genetic variants contribute to disease, functional annotation tools have been increasingly utilized. Functional Mapping and Annotation (FUMA) integrates public biological databases to provide insights into tissue-specific expression and biological pathways associated with GWAS-identified variants.[23] In this study, FUMA was initially used to annotate significant loci. Subsequently, a series of custom analyses, including differential expression analysis, heatmap visualization, and network mapping, were performed to gain deeper insights into the functional relevance of MASLD and CVD-associated variants.

Based on the genetic variants identified through GWAS, there has been a growing interest in the polygenic risk score (PRS), which analyzes multiple variants collectively rather than focusing on individual variants. PRS is a score calculated by summing the weighted effects

of various genetic variants and is used to quantify an individual's genetic predisposition to disease.[24] In the context of MASLD research, PRS can be utilized to identify high-risk individuals at an early stage and to develop preventive and therapeutic strategies based on genetic risk. PRS offers the advantage of integrating multiple genetic factors associated with MASLD, thereby overcoming the limitations of single-gene analyses and improving the accuracy of risk prediction.[25]

Genetic variants identified through GWAS serve as the foundation for Mendelian Randomization (MR) studies.[26] MR is a method that utilizes genetic variants as instrumental variables to assess causal associations, based on the principle that genetic variants are randomly inherited from parents according to Mendel's laws.[26] This random allocation helps minimize confounding factors and selection bias commonly encountered in observational studies, thereby allowing clearer inference of causality. Owing to these advantages, MR has become a valuable tool in epidemiological research for investigating associations and providing evidence to guide disease etiology and prevention strategies. MR is well-suited for evaluating the long-term effects of exposure and is especially useful in cases where clinical interventions are challenging or unethical. However, most MR studies have focused predominantly on populations of European ancestry, which limits the generalizability of their findings to other populations.[27,28]

In addition to the main genetic effects, there is growing recognition that the impact of genetic variants may vary depending on environmental exposure, a phenomenon known as gene–environment interaction (GxE).[29,30] While GWAS typically focus on average genetic

effects across the population, GxE analysis allows for the assessment of whether these effects differ across levels of environmental factors such as smoking, diet, and physical activity.[29,31] This is particularly relevant for complex, multifactorial diseases, such as MASLD, which arise from both genetic predisposition and modifiable lifestyle risk factors. Recent studies have begun to explore GxE interactions in the context of liver diseases; however, most have been limited in scope or have been conducted in homogeneous populations. By incorporating GxE analysis in this study, we aimed to evaluate whether genetic variants associated with MASLD demonstrate differential effects depending on smoking status. This approach enables a more nuanced understanding of genetic susceptibility, and may help identify individuals for whom environmental interventions are particularly beneficial.

This study aimed to elucidates the association and potential causal pathways between MASLD and cardiovascular disease development. Although numerous studies have examined the link between MASLD and CVD, few have simultaneously applied observational and genetic-epidemiological approaches. Previous MR studies have explored the causal associations between MASLD and outcomes such as myocardial infarction, coronary artery disease, and hyperlipidemia; however, the findings remain controversial. These studies often lack detailed definitions of MASLD and specific cardiovascular outcomes, limiting the accuracy of interpretation.

Therefore, this study aimed to clarify the causal effect of MASLD on the incidence of CVD by conducting survival analyses using the KCPS-II dataset, followed by genetic analyses,

including GWAS, MR, and GxE. Post-GWAS analyses including gene-based testing, tissue-specific expression profiling, and functional enrichment were performed to explore the biological mechanisms underlying MASLD. By integrating observational, genetic, and GxE approaches, this study aimed to move beyond simple associations and provide more nuanced and robust evidence on the association between MASLD and CVD.

2. Objectives

The objectives of this study are as follows:

(1) To evaluate the association between the MASLD and CVD using prospective cohort data from the KCPS-II.

(2) To investigate the causal association between MASLD and CVD using MR, employing both one and two-sample MR approaches.

(3) To explore the biological relevance of the genetic variants associated with MASLD and CVD by performing expression quantitative trait locus (eQTL) analysis and functional annotation.

(4) To assess GxE by evaluating whether genetic effects on MASLD differ with environmental exposure, particularly smoking status.

II. MATERIALS AND METHODS

1. Data source and study population

1-1. Korean Cancer Prevention Study-II

This study examined data from the Korean Cancer Prevention Study-II, comprising 153,950 participants with written consent from 18 institutions for health check-ups nationwide from 2004 to 2013. This data source was described in a previous study.[32] Participants with incomplete information on key variables for this analysis, those with a history of CVD, and those with a follow-up period of less than one year were excluded; finally, 111,637 participants were selected for analysis.

Data were collected in a standardized manner using questionnaires and examination data conducted every two years at local health check-ups. The questionnaire inquired about age, sex, smoking status (never, former, or current), alcohol consumption (g/day), exercise status (never, former, or current), and insurance coverage. The participants' body mass index (BMI), waist circumference, systolic blood pressure (SBP), diastolic blood pressure (DBP), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglyceride (TG), total cholesterol, alanine transaminase (ALT), aspartate transaminase (AST), gamma glutamyl transferase (GGT), fasting blood glucose (FBS), platelet counts (10^9/L), and glomerular filtration rate (eGFR(mL/min/1.73m^2)) were

included in the analysis. Blood sampling was performed in the morning following overnight fasting, and automatic analysis devices (Hitachi 737) were used for the laboratory tests.[32] Each medical institution that performed the tests was accredited for internal and external quality assessment using the Korean Association of External Quality Assessment Service, as to maintain the accuracy of the laboratory tests.

The FLI is an alternative to imaging diagnostic methods to distinguish SLD; an FLI value higher than or equal to 30 has been defined as SLD.[6] In the group defined as SLD, those who met at least one of the following cardiometabolic risk factors (CMRFs) were distinguished as MASLD: (i) body mass index (BMI) ≥ 23 kg/m^2 or waist circumference > 90 cm for males and > 80 cm for females; (ii) fasting serum glucose ≥ 100 mg/dL (iii) blood pressure ≥ 130/85 mmHg (iv) plasma triglycerides ≥ 150 mg/dL and (v) plasma high-density lipoprotein (HDL) cholesterol ≤ 40 mg/dL for males and ≤ 50 mg/dL for females. Alcoholic liver disease (ALD) was characterized by notable alcohol intake (>60 g/day for males and >50 g/day for females) leading to hepatic steatosis, irrespective of metabolic indicators. MetALD was defined as ALD accompanied by moderate alcohol consumption (30–60 g/day for males and 20–50 g/day for females). The other was delineated as liver disease lacking cardiometabolic risk factors and significant alcohol intake. The FLI is endorsed by international guidelines as a non-invasive proxy for imaging and is, particularly suitable for use in population-level studies.[33,34] The area under the receiver operating characteristic curve (AUROC) of the FLI was 0.87 in the Korean population.[35]

The Tyg index was calculated using the formula Ln [TG (mg/dL) × FBG (mg/dL)/2].[36] The fibrosis 4-score was calculated using the following formula age (years) x AST (U/L) / [platelet count ($\times 10^9$/L) x square root (ALT) (U/L)].[37]

T2DM and hepatitis B or C were classified based on the International Classification of Diseases, 10th Revision (ICD-10) codes: E10–E14 for diabetes and B16 and B18.2 for hepatitis B or C. The primary outcome, atherosclerotic vascular diseases, was defined by I10–I15, I20–I25, I44–I52, I60–I69, and I70–I74 [diseases of arteries], ischemic heart diseases (I20–I25), total stroke (I60–I69), hemorrhagic stroke (I60–I62), thrombotic stroke (I63.0, I63.3), and myocardial infraction (I21–I23) codes tracked in the national health insurance claims data.[38] The secondary outcomes were the specific events included in the primary outcome.

1-2. Biobank Japan summary statistics

GWAS summary statistics of coronary artery disease (CAD), myocardial infarction (MI), and ischemic stroke (IS) outcomes were obtained from the BioBank Japan (BBJ) population of East Asian ancestry[39] listed in the publicly available BBJ PheWeb (https://pheweb.jp/). The BBJ project enrolled patients diagnosed with one or more of the 47 designated common diseases between June 2003 and March 2008 in collaboration with 12 medical institutions across Japan. Clinical information was collected annually through interviews and review of medial records until 2013. BBJ collected DNA from all participants at baseline and collected annual serum samples until 2013. In addition, the registry followed patients who reported a history of 32 of the 47 target diseases to collect survival data, including the causes of death. Details of the BBJ study design, genotyping and quality control have been described previously.[15,40] CAD (ICD-10 code as I20-I25, 32,512 cases and 146,214 controls), MI (ICD-10 code as I21, 14,992 cases, 146,214 controls), and IS (ICD-10 code as I63, 22,664 cases, 152,022 controls) were included as outcomes.

1-3. United Kingdom Biobank

The UK Biobank (UKB) is a large-scale, population-based prospective cohort comprising approximately 500,000 participants aged 40–69 years between 2006 and 2010 across the United Kingdom.[41] It provides extensive phenotypic data, lifestyle information, health records, and genetic data, making it a valuable resource for investigating complex disease etiologies and GxE. Participants were linked to national health registries for the long-term follow-up of incident diseases, including cardiovascular outcomes. CVD (72,008 cases and 430,473 controls), IHD (38,597 cases and 462,884 controls), MI (14,232 cases and 488,249 controls), total stroke (11,762 cases and 490,719 controls), thrombotic stroke (669 cases and 501,812 controls), and hemorrhagic stroke (3,350 cases and 499,131 controls) were included as outcomes. For the two-sample MR, the MASLD and CVD phenotypes were defined in the UKB using the same criteria applied to the KCPS-II dataset. The ICD-10 codes used to determine CVD outcomes were identical to those used in the KCPS-II. MASLD was defined based on the fatty liver index (FLI $\geq$ 60). In the UKB, alcohol intake is calculated by summing drink-specific variables (e.g., red wine, beer, and spirits) to approximate the total daily alcohol consumption, enabling consistent exclusion criteria with the KCPS-II. Patients with viral hepatitis were excluded using ICD-10 codes (B15–B19).

GWAS was performed using logistic regression, adjusting for age, sex, assessment center, genotyping array, and the first 10 principal components of ancestry. Quality control

measures included filtering out SNPs with call rate < 0.95, Hardy-Weinberg equilibrium p < 1.0×10^{-7}, and minor allele frequency (MAF) < 0.01. Linkage disequilibrium (LD) clumping was conducted with an r^2 threshold of 0.3 and a 1,000-kb window to select independent genome-wide significant SNPs ($p < 5.0 \times 10^{-8}$).

2. Selection of genetic instruments: KCPS-II

Approximately 50% of the participants in this study were genotyped using the Global Screening Array (GSA) chip, while the remaining half were genotyped using the Korea Biobank Array.[42] Subsequently, genotype imputation was performed using IMPUTE5, based on the 1,000 Genomes Project reference panel, to construct an integrated dataset. IMPUTE5 is a widely used software that infers unobserved or missing genotypes by leveraging known haplotype reference panels and recombination maps. A GWAS was performed using the Affymetrix Genome-Wide Human SNP array 5.0 (Affymetrix, Santa Clara, CA, USA) to identify SNPs for MASLD, which was measured as an independent variable. Logistic regression, including sex, age, and chip type, was conducted. For quality control, monomorphic variants were excluded based on the following criteria: (1) call rate < 0.95, (2) Hardy-Weinberg equilibrium (HWE) ($p < 1.0 \times 10^{-7}$), and (3) minor allele frequencies (MAF) < 0.01. In total, 6,809,738 single-nucleotide polymorphisms (SNPs) were analyzed. To calculate the weighted genetic risk score (WGRS), LD clumping was performed using an r^2 threshold of 0.3, a window size of 1,000 kb, and a p-value of 5.0×10^{-8}. After excluding correlated SNPs using the clumping algorithm, 82 variants were identified as MASLD (Table 1). GWAS analysis was performed using PLINK 2.0.

Table 1. Selected genetic variants for MASLD (N=82)

No	CHR	SNP	BP	RA	EA	EAF	BETA	SE	P-value
1	1	rs12562924	75790373	T	C	0.1566	0.0715	0.0129	2.72E-08
2	1	rs543874	177889480	A	G	0.2459	0.0733	0.0109	1.74E-11
3	2	rs1260326	27730940	T	C	0.4504	-0.0914	0.0095	4.62E-22
4	2	rs145508558	27783195	TTTAA	T	0.3287	-0.0690	0.0100	5.77E-12
5	2	rs1728918	27635463	G	A	0.2459	0.0622	0.0109	1.01E-08
6	2	rs35142762	636790	T	C	0.0921	-0.1193	0.0164	3.79E-13
7	4	rs10006310	146809998	T	G	0.4336	0.0594	0.0095	3.46E-10
8	4	rs13130484	45175691	C	T	0.2826	0.0719	0.0104	5.15E-12
9	4	rs60142704	146789479	C	T	0.2678	-0.0584	0.0106	3.74E-08
10	5	rs34566853	95850866	C	CA	0.3198	0.0569	0.0101	1.53E-08
11	6	rs2523655	31444604	G	A	0.0953	0.1021	0.0159	1.37E-10
12	6	rs2744475	50784880	C	G	0.3887	0.0607	0.0096	2.86E-10
13	6	rs35366046	33731989	T	C	0.1145	0.0909	0.0147	6.03E-10
14	6	rs72896150	33793096	G	A	0.0491	0.1206	0.0215	2.17E-08
15	6	rs9273704	32629297	C	T	0.3149	0.0655	0.0101	9.03E-11
16	7	rs10245965	73063515	T	C	0.2487	-0.0656	0.0109	1.73E-09
17	7	rs370621425	99007392	CG	C	0.1780	-0.0782	0.0123	2.16E-10
18	7	rs3812316	73020337	C	G	0.1016	-0.1580	0.0158	1.35E-23
19	8	rs112875651	126506694	G	A	0.2053	-0.0644	0.0116	3.22E-08
20	8	rs2954029	126490972	T	A	0.4417	0.0621	0.0095	5.52E-11
21	8	rs6999813	19863471	T	A	0.1232	-0.0968	0.0144	1.59E-11
22	8	rs7829886	95497388	A	T	0.4928	0.0532	0.0094	1.45E-08
23	11	rs11216118	116596395	C	T	0.0657	0.1828	0.0188	2.12E-22
24	11	rs11602073	116646858	C	T	0.2178	-0.1274	0.0115	1.60E-28
25	11	rs117010832	116732519	G	A	0.0878	-0.0960	0.0168	1.05E-08
26	11	rs118175510	116532548	T	C	0.1069	0.0899	0.0151	2.61E-09
27	11	rs12293222	116565309	G	A	0.2693	0.0639	0.0106	1.43E-09
28	11	rs1240772	116519129	G	C	0.3329	0.0544	0.0099	4.62E-08

29	11	rs1263056	116576415	G	A	0.3099	-0.0737	0.0102	4.12E-13
30	11	rs144026079	116879777	C	T	0.0236	0.1847	0.0305	1.46E-09
31	11	rs180322	116626735	T	C	0.4057	-0.0898	0.0096	7.17E-21
32	11	rs180346	116612659	A	C	0.2271	-0.0905	0.0113	9.69E-16
33	11	rs180360	116598988	A	G	0.1788	-0.0988	0.0123	1.14E-15
34	11	rs184616707	116510558	C	G	0.0163	0.2279	0.0363	3.46E-10
35	11	rs200818218	116638691	AAAT	A	0.4657	-0.0593	0.0094	3.10E-10
36	11	rs2070665	116707684	G	A	0.3538	0.0981	0.0098	1.06E-23
37	11	rs2075291	116661392	C	A	0.0772	0.2633	0.0173	3.23E-52
38	11	rs34488176	116705206	C	CT	0.0288	0.2085	0.0278	6.18E-14
39	11	rs4938304	116588425	C	T	0.0761	-0.1031	0.0179	8.29E-09
40	11	rs60954647	116566933	C	T	0.4484	-0.0529	0.0095	2.32E-08
41	11	rs61697392	116893550	TA	T	0.4382	-0.0561	0.0095	3.32E-09
42	11	rs6265	27679916	C	T	0.4582	-0.0675	0.0095	9.38E-13
43	11	rs651821	116662579	T	C	0.2945	0.1998	0.0103	3.73E-84
44	11	rs66505542	116623213	T	TA	0.4255	0.0782	0.0095	1.62E-16
45	11	rs79538491	116681919	A	G	0.0901	0.0966	0.0164	3.79E-09
46	12	rs11065933	111942493	T	C	0.3900	-0.0988	0.0097	2.29E-24
47	12	rs11066525	113713062	A	C	0.0850	-0.1502	0.0170	9.05E-19
48	12	rs11614295	113196733	G	A	0.3948	-0.0540	0.0096	2.04E-08
49	12	rs1169289	121416622	C	G	0.4652	-0.0677	0.0095	8.66E-13
50	12	rs117949785	113153857	C	T	0.4013	-0.0715	0.0096	1.13E-13
51	12	rs141965732	110582338	C	T	0.0959	-0.1380	0.0161	9.78E-18
52	12	rs1628251	113520696	A	C	0.2454	-0.0605	0.0109	3.34E-08
53	12	rs2339905	112378365	G	A	0.3677	0.0722	0.0097	1.16E-13
54	12	rs34749124	111292097	C	CTT	0.4874	-0.0547	0.0094	6.83E-09
55	12	rs35065054	121200397	C	CT	0.3892	0.0583	0.0097	1.53E-09
56	12	rs35236844	111730156	G	GA	0.3565	0.0542	0.0098	3.64E-08
57	12	rs4766552	111631765	G	A	0.4430	-0.0637	0.0095	1.88E-11
58	12	rs4767014	113264755	T	C	0.2813	-0.0646	0.0105	6.88E-10

59	12	rs561500595	111704999	C	CAAAAAAAAAAA	0.4848	-0.0561	0.0094	2.71E-09
60	12	rs57606492	111382452	CA	C	0.4362	0.0653	0.0095	5.33E-12
61	12	rs671	112241766	G	A	0.1740	-0.1940	0.0126	1.60E-53
62	12	rs71445573	112829103	CAT	C	0.4106	0.0812	0.0096	2.01E-17
63	12	rs7294578	112197837	C	A	0.1184	0.0840	0.0145	6.98E-09
64	12	rs7298118	111837285	G	A	0.1591	0.0703	0.0128	3.61E-08
65	12	rs7299183	111771763	A	G	0.4158	0.0564	0.0096	3.69E-09
66	12	rs73199895	111866087	G	A	0.3560	0.0601	0.0098	8.89E-10
67	12	rs77346308	110153675	C	T	0.0850	-0.1015	0.0170	2.14E-09
68	12	rs886477	113319308	T	A	0.2033	-0.1063	0.0117	1.29E-19
69	15	rs3837737	68017574	CT	C	0.3778	0.0550	0.0097	1.36E-08
70	15	rs588136	58730498	T	C	0.3863	0.0530	0.0096	3.56E-08
71	16	rs11642015	53802494	C	T	0.1249	0.1386	0.0141	7.71E-23
72	16	rs12599076	20254901	G	C	0.2091	-0.0669	0.0116	8.71E-09
73	16	rs62033406	53824226	G	A	0.4889	-0.0539	0.0094	9.61E-09
74	17	rs113960551	66113360	C	T	0.0955	-0.0974	0.0162	1.78E-09
75	17	rs12449442	65947640	A	G	0.3686	-0.0572	0.0097	4.44E-09
76	17	rs67093103	34907050	CT	C	0.3556	-0.0589	0.0098	2.03E-09
77	18	rs558522	57861449	A	T	0.2648	0.0828	0.0106	5.35E-15
78	19	rs3852860	45382966	T	C	0.2200	0.0627	0.0113	3.14E-08
79	19	rs429358	45411941	T	C	0.0954	0.1303	0.0159	2.39E-16
80	19	rs584007	45416478	A	G	0.3862	0.0566	0.0096	4.57E-09
81	22	rs2330805	24998619	A	G	0.3476	0.1176	0.0098	3.41E-33
82	22	rs545591444	24989369	T	TG	0.0964	0.1283	0.0158	4.02E-16

CHR, chromosome; SNP, single nucleotide polymorphism; BP, base position; RA, reference allele; EA, effective allele; EAF, effective allele frequency

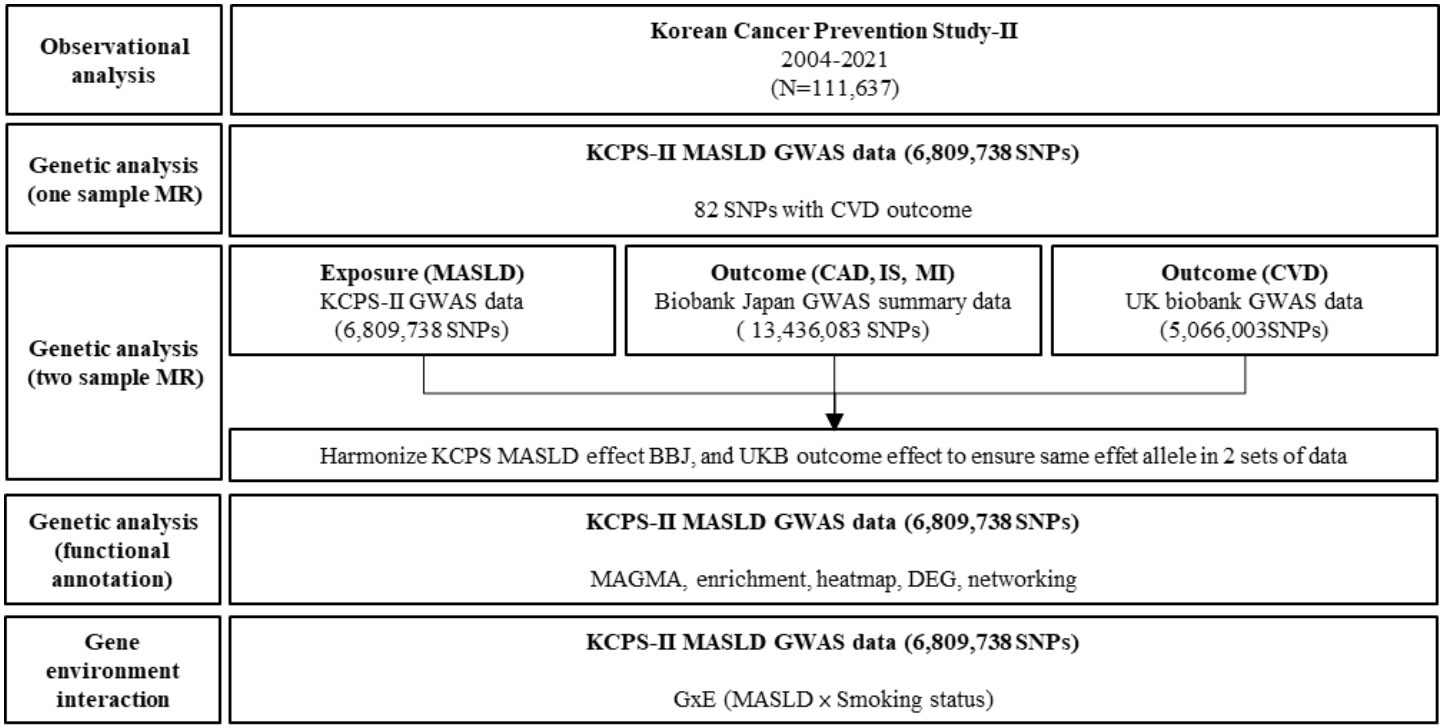


Figure 1. Overall study design and analytical workflow

3. Statistical analysis

3-1. Observational analysis

Continuous variables are presented as mean (standard deviation) if normally distributed or median (interquartile range) if not normally distributed. Differences between groups were evaluated using either the independent t-test or the Mann-Whitney U test. Categorical data were expressed as numbers (%), and differences between groups were determined using the chi-squared test. The cumulative incidence of CVD events was estimated using the Kaplan-Meier method. Incidence was calculated as cases per 100,000 individuals. The HRs and 95% confidence intervals (CI) of the outcomes were calculated using Cox proportional hazards regression analysis adjusted for age, sex, exercise status, smoking status, alcohol consumption, insurance, and glomerular filtration rate. The variables used were associated with the occurrence of SLD and CVD in previous studies.[43,44] CVD risk factors, alcohol consumption, and accompanying variables were not adjusted in the analysis because liver disease was included in the diagnostic criteria for SLD.

The first sensitivity analysis was performed by stratifying the main analysis according to age, sex, smoking status, obesity, TyG index, T2DM, hypertension, and dyslipidemia. A higher cut-off of FLI ≥ 60 was adopted to define SLD in the second analysis. CVD cases that occurred within three years of the first follow-up date were excluded to preclude incident cases before the start of follow-up in the third analysis. The E-value was estimated to assess the potential influence of unmeasured confounding factors, indicating that the

minimum strength of association of an unmeasured confounder would be needed with both MASLD and CVD to account for the observed association fully. The fifth analysis evaluated the robustness of the findings and included a sensitivity analysis using the concordance index (C-index) to assess the predictive performance and discrimination ability.

All statistical analyses were conducted using R version 4.3.2 (R Foundation for Statistical Computing, Vienna, Austria), and the significance level was set at a p-value of < 0.05.

3-2. Genetic analysis

3-2-1. Functional mapping and network-based characterization of MASLD-associated genetic variants

Significant genome-wide variants were functionally annotated using multiple bioinformatics tools and public databases. Initially, functional mapping was performed using FUMA, which applies positional mapping and quantitative expression trait loci (eQTL) mapping based on the Ensembl gene builder and the GTEx v8 reference panel. Tissue-specific expression enrichment was assessed using GTEx data.

To complement the web-based annotation, custom analyses were conducted using R-based pipelines that incorporated gene-level information from Ensembl, eQTL data from GTEx, and transcriptomic profiles. In addition, protein–protein interaction networks were constructed using the STRING DB to identify functionally connected gene modules associated with MASLD. This multilevel integrative approach enabled a more comprehensive biological interpretation of GWAS findings.

3-2-2. One-sample Mendelian randomization study

Deviation of the genetic variants from the Hardy-Weinberg equilibrium was investigated using Pearson's chi-square test. To assess the strength of the genetic variants as instrumental variables, F-values and variations in MASLD explained by the genetic variants were calculated using logistic regression. Associations between genetic MASLD scores and potential confounders were investigated using logistic regression analysis. Associations between weighted genetic MASLD scores and clinical outcomes were examined using multivariate logistic regression analysis. Instrumental variable estimates of causal risk ratios were calculated using the Wald-type estimator and internally weighted genetic MASLD score to estimate the influence of genetically determined MASLD on the risk of CVD, ischemic heart disease, myocardial infarction, total stroke, thrombotic stroke, and hemorrhagic stroke. All single-sample genetic analyses were adjusted for age and sex.

The weighted genetic risk score (WGRS) was calculated as follows: For each SNP with alleles A and T, where T is considered the risk allele, the number of risk alleles was coded as 0, 1, or 2 for the genotypes AA, AT, and TT, respectively. Assuming an additive genetic model, the regression coefficients corresponding to each additional risk allele were calculated. To assign weights, the coefficient was multiplied by the number of risk alleles for each SNP. The WGRS for each individual was obtained by summing the weighted values of all selected SNPs. Formally, the weighted GRS (GRS i) is defined as the weighted

sum of the number of risk alleles (coded as 0, 1, and 2) of the k considered SNPs (g i1, …, g ik) for n subjects (i = 1, …, n):[45]

$$WGRS_i = w_1 g_{i1} + \cdots + w_k g_{ik}$$

3-2-3. Two-sample Mendelian randomization

This study investigated the causal associations between exposure (MASLD) and outcomes (CVD) using a two-sample MR analysis. In two-sample MR, the inverse variance-weighted (IVW) method estimates the causal effect by combining SNP-specific Wald ratios using a weighted average. Each ratio was calculated by dividing the impact of the SNP on the outcome by its effect on the exposure. The weights are the inverse of the variance of each ratio, giving more influence to precise estimates. The IVW assumes that all genetic variants are valid instruments with no horizontal pleiotropy. It is the most commonly used MR method because of its high statistical power.[46]

The MR-Egger regression extends the IVW method by allowing for an intercept term that captures the unbalanced horizontal pleiotropy across SNPs. Unlike the IVW, it does not assume that all genetic variants are valid instruments. The slope of the MR-Egger regression represents the causal effect, whereas the intercept represents the presence of directional pleiotropy. It is more robust to pleiotropy but, typically has lower statistical power. MR-Egger requires a stronger assumption called Instrument Strength Independent of Direct Effect (InSIDE). This is useful when pleiotropy is suspected based on genetic instruments.[47]

The weighted-median method provides a causal estimate by taking the median of the individual SNP ratio estimates weighted by the inverse variance of their effects. It provides a valid causal estimate even if up to 50% of the instruments are invalid, as long as the

majority of the weight comes from the valid instruments. This makes it more robust than IVW in horizontal pleiotropy. The method does not require the InSIDE assumption, similar to the MR-Egger method. This is useful when genetic variants violate instrumental variable assumptions. It balances robustness and efficiency, offering a compromise between IVW and MR-Egger.[48]

The simple mode method estimates the causal effect in MR by identifying the most common (modal) value among the individual SNP ratio estimates. We assumed that the largest group of SNPs with similar effects would be a valid instrument. Unlike IVW or MR-Egger, it does not rely on all SNPs being valid or balanced. This method is robust to pleiotropy if the most valid instruments cluster around the actual effect.[48]

The weighted-mode method estimates the causal effect by identifying the most frequent value (mode) among the SNP-specific causal estimates weighted by their precision. This method assumes that the largest weighted cluster of SNPs was valid. Compared with the simple mode, it provides greater efficiency by assigning more weight to precise estimates. It is robust against horizontal pleiotropy if valid instruments comprise the largest weighted group. This method does not require all SNPs to be valid instruments, offering a robustness similar to that of the median method. This is useful for sensitivity analyses alongside IVW and MR-Egger.[49]

Radial MR is a framework that reformulates MR using a radial plot in which, the contribution of each SNP is visualized in terms of its precision and influence. This allows better detection of outliers and influential SNPs that may distort the causal estimate. The

method can be applied to various MR models, including the IVW and MR-Egger, in a radial form. Radial MR uses a modified regression approach that helps in variance stabilization and improves model fit, enabling a more lucid interpretation of heterogeneity and visual assessment of pleiotropy. Radial plots are often used for diagnostic and sensitivity purposes in MR studies.[50]

3-2-4. Gene environment interaction

To estimate the proportion of phenotypic variance attributable to additive genetic effects and GxE, the Gene–Environment Interaction Estimator (GENIE) framework was applied.[51] GENIE is based on a linear mixed model, in which phenotype y is modeled as:

$$y = g + g \times e + \varepsilon$$

where:

$g \sim N(0, \sigma_g^2 K)$: $additive\ genetic\ effects$

$g \times e \sim N(0, \sigma_{g\times e}^2 K_{g\times e})$: $gene\ environment\ ineraction\ effects$

$\varepsilon \sim N(0, \sigma_e^2 I)$: $reidual\ error.$

This model partitions total phenotypic variance into three components:

$\sigma_g^2 =$ variance explained by additive genetic effects

$\sigma_{g\times e}^2 =$ variance explained by GxE interactions

$\sigma_e^2 =$ residual variance

The proportion of trait variance explained by each component is then calculated as:

$$h_g^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{g\times e}^2 + \sigma_e^2}, \qquad h_{g\times e}^2 = \frac{\sigma_{g\times e}^2}{\sigma_g^2 + \sigma_{g\times e}^2 + \sigma_e^2}$$

This allowed us to quantify the extent to which the total variability of the trait could be attributed to genetic factors and their interactions with the environment.

Genome-wide analysis was performed using the GxEScanR package to identify specific loci involved in GxE.[52] For each SNP, a linear regression model was fitted, including the main genetic effect, environmental variables of interest (e.g., smoking status), and their interaction term. The model can be expressed as follows:

$$logit(\Pr(D = 1|G, E, C) = \beta_0 + \beta_G G + \beta_E E + \beta_{GXE} GxE + \beta_c C$$

Where D is the binary trait (e.g., disease status), G is the genotype (e.g., SNP dosage), E is the environmental exposure (e.g., smoking), G×E is the gene–environment interaction term, and C denotes covariates such as age, sex, and ancestry principal components. The primary parameter of interest is $\beta_{G\times E}$, which quantifies whether the effect of the genotype on the trait differs depending on the environmental exposure. Statistical significance of the interaction was evaluated using Wald tests under the null hypothesis $H_0: \beta_{G\times E} = 0$.

The proportion of phenotypic variance explained by GxE effects was estimated using GENIE, a scalable linear mixed model–based method for partitioning variance into additive genetic and GxE components (https://github.com/younglululu/GENIE). Genome-wide scans were conducted using GxEScanR, an R package that implements several methods for detecting GxE (https://CRAN.R-project.org/package=GxEScanR).

III. RESULTS

PART I. Association between MASLD and CVD: observational study

1. General characteristics of the study population

A total of 111,637 participants were included in the analytic cohort, with a median age of 39 years (interquartile range: 33–46 years) and a female proportion of 36.1%. Table 2 shows the baseline characteristics of 32,023 (29.1%) patients with MASLD: 5,602 (5.1%) had MetALD, 3,778 (3.4%) had ALD, and 303 (0.2%) had other combined etiologies. Compared with the non-SLD group, participants with MASLD were more likely to be older, male, have a higher body mass index (BMI), and be current smokers. Biomarkers such as FBS, SBP, DBP, total cholesterol, TG, platelet count, Tyg index, and liver enzyme levels (ALT, AST, and GGT) were significantly elevated in the MASLD and MetALD groups. In contrast, HDL cholesterol and eGFR levels were significantly lower.

Table 2. Baseline characteristics of participants by subcategories of SLD

Variables		No-SLD N=68,492 (61.3%)	MASLD N= 32,018 (28.7%)	MetALD N= 5,602 (5.0%)	ALD N=3,778 (3.4%)	Other N=1,747 (1.5%)
Age, years		39.0 (9.6)	42.1 (9.1)	41.3 (8.6)	42.6 (8.9)	42.0 (8.6)
BMI, kg/m^2		21.8 (2.3)	26.1 (2.5)	26.1 (2.5)	25.9 (2.6)	25.4 (2.7)
Waist circumference, cm		75.2 (7.0)	88.8 (6.3)	89.1 (6.4)	89.2 (6.6)	88.1 (6.6)
Systolic BP, mm/Hg		114 (13.2)	123 (12.9)	125 (12.9)	125 (12.8)	122 (12.5)
Diastolic BP, mm/Hg		71.4 (9.2)	77.9 (9.5)	78.9 (9.6)	79.3 (9.7)	77.2 (9.6)
LDL-cholesterol, mg/dl		108 (28.7)	121 (33.9)	118 (34.9)	117 (34.6)	118 (30.2)
HDL-cholesterol, mg/dl		55.0 (10.3)	46.4 (8.0)	48.1 (9.0)	48.5 (9.6)	48.8 (9.3)
Total cholesterol, mg/dl		181 (30.1)	202 (33.0)	201 (33.9)	202 (34.5)	195 (30.7)
Triglyceride, mg/dl		85.0 (65.0-113.0)	178.0 (132.0-241.0)	181.0 (132.0-252.0)	185.0 (134.0-261.0)	136.0 (110.0-179.0)
Fasting glucose, mg/dl		86.0 (80.0-92.0)	91.0 (84.0-100.0)	93.0 (85.0-101.0)	93.0 (86.0-103.0)	89.0 (82.0-97.0)
GGT, IU/L		18.0 (14.0-25.0)	42.0 (29.0-64.0)	56.0 (38.0-86.0)	64.0 (42.0-102.0)	48.0 (32.0-77.0)
AST, IU/L		19.0 (16.0-22.0)	24.0 (20.0-30.0)	25.0 (21.0-31.0)	26.0 (22.0-32.0)	28.0 (22.0-37.0)
ALT, IU/L		16.0 (12.0-21.0)	30.0 (22.0-42.0)	30.0 (22.0-41.0)	30.0 (22.0-43.0)	35.0 (25.0-51.0)
Platelet (1000cells/uL)		244.0 (212.0-281.0)	253.0 (220.0-290.0)	248.0 (216.0-283.0)	246.0 (213.0-282.0)	227.0 (191.0-265.0)
Tyg index		8.23 (0.4)	9.05 (0.5)	9.08 (0.5)	9.12 (0.5)	8.77 (0.4)
eGFR, mL/min/1.73m^2		87.6 (13.9)	84.7 (12.8)	86.0 (12.6)	86.1 (13.4)	84.8 (13.3)
Sex	Male	32,254 (47.1)	28,781 (89.9)	5,371 (95.9)	3,694 (97.8)	1,645 (94.2)
	Female	36,238 (52.9)	3,237 (10.1)	231 (4.12)	84 (2.2)	102 (5.8)
Alcohol	Never	20,622 (30.1)	5,436 (17.0)	0 (0.0)	0 (0.0)	282 (16.1)
consumption	0-30	42,337 (61.8)	26,582 (83.0)	143 (2.5)	0 (0.0)	1,404 (80.4)
(g/day)	30-60	3,699 (5.4)	0 (0.0)	5,459 (97.4)	19 (0.5)	61 (3.5)
	60>	1,834 (2.6)	0 (0.0)	0 (0.0)	3,759 (99.5)	0 (0.0)
Smoking status	Never	41,779 (61.0)	9,043 (28.2)	841 (15.0)	432 (11.4)	424 (24.3)
	Former	12,742 (18.6)	9,439 (29.5)	1,667 (29.8)	1,158 (30.7)	506 (29.0)
	Current	13,971 (20.4)	13,536 (42.3)	3,094 (55.2)	2,188 (57.9)	817 (46.8)
Exercise status	Never	39,337 (57.4)	21,284 (66.5)	3,721 (66.4)	2,497 (66.1)	1,144 (65.5)
	Former	13,217 (19.3)	5,254 (16.4)	981 (17.5)	551 (14.6)	313 (17.9)
	Current	15,938 (23.3)	5,480 (17.1)	900 (16.1)	730 (19.3)	290 (16.6)
Insurance	Q1	23,147 (33.8)	8,329 (26.0)	1,499 (26.8)	1,225 (32.4)	475 (27.2)
	Q2	23,509 (34.3)	11,188 (34.9)	1,930 (34.5)	1,208 (32.0)	597 (34.2)
	Q3	21,836 (31.9)	12,501 (39.0)	2,173 (38.8)	1,345 (35.6)	675 (38.6)
Viral Hepatitis		2,874 (4.20)	0 (0.0)	198 (3.53)	142 (3.76)	1,383 (79.2)

Continuous data are presented as mean (standard deviation) (normally distributed) or medians (interquartile ranges) (not normally distributed). Categorical data are expressed as the number (%). *all p-values <0.001 for difference. MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology; BMI, body mass index; LDL, low density lipoprotein; HDL, high density lipoprotein; TG, Triglyceride; FBS, fasting blood glucose; CRP, c-reactive Protein; GGT, Gamma Glutamyl Transferase; ALT, alanine transaminase; AST, aspartate transaminase; Viral Hepatitis, hepatitis B or C;

2. Cumulative incidence of CVD according to different subtypes of SLD

In total, 111,637 participants without prior CVD were included in this longitudinal analysis. During a median follow-up of 10.0 years, 3,926 incident CVD events occurred. After adjusting for sociodemographic factors (sex, age, exercise status, smoking status, and insurance) and eGFR, the risk of CVD was 1.69 times higher (95% CI 1.58–1.82) in participants with SLD than in those without. HR was 1.69 (95% CI 1.57–1.82) for MASLD, 1.73 (95% CI 1.52–1.97) for MetALD, 1.76 (95% CI 1.52–2.03) for ALD, and 1.59 (95% CI 1.27–1.98) for Other (Table 3).

MASLD was also associated with higher risks of all secondary outcomes including IHD (HR 1.74, 95% CI 1.54–1.96), MI (HR 2.53, 95% CI 1.86–3.45), total stroke (HR 1.43, 95% CI 1.25–1.65), thrombotic stroke (HR 1.63, 95% CI 1.31−2.03), and hemorrhagic stroke (HR 1.40, 95% CI 1.01–1.92) but not with all-cause mortality (HR 0.87, 95% CI 0.71–1.08) (Table 4).

Table 3. CVD risk according to SLD subtypes

Group	Events	Person-years	Rate[a]	HR (95% CI)		
				Model 1	Model 2	Model 3
No-SLD	1,655 / 68,492	677,208.1	244.4	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	2,271 / 43,145	421,232.7	539.1	2.20 (2.10-2.40)	1.74 (1.63-1.86)	1.70 (1.58-1.82)
MASLD	1,679 / 32,018	312,838.1	536.7	2.20 (2.10-2.40)	1.72 (1.60-1.85)	1.69 (1.57-1.82)
MetALD	285 / 5,602	54,568.7	522.2	2.10 (1.90-2.40)	1.82 (1.60-2.07)	1.73 (1.52-1.97)
ALD	223 / 3,778	36,743.0	606.9	2.50 (2.20-2.90)	1.88 (1.63-2.17)	1.76 (1.52-2.03)
Other	84 / 1,747	17,082.9	491.7	2.00 (1.60-2.50)	1.64 (1.31-2.05)	1.59 (1.27-1.98)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.
[a]Rate per 100,000 person-years.
Model 1 was unjusted.
Model 2 was adjusted for age and sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR.

Table 4. Risk of secondary outcomes according to SLD subtypes

Group	Events	Person-years	Rate[a]	HR (95% CI)		
				Model 1	Model 2	Model 3
Ischemic heart disease						
No-SLD	539 / 68,492	683,253.7	78.9	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	886 / 43,145	428,611.8	206.7	2.60 (2.40-2.90)	1.78 (1.59-2.00)	1.72 (1.53-1.93)
MASLD	657 / 32,018	318,342.3	206.4	2.60 (2.30-2.90)	1.78 (1.58-2.01)	1.74 (1.54-1.96)
MetALD	109 / 5,602	55,480.1	196.5	2.50 (2.00-3.10)	1.77 (1.43-2.19)	1.65 (1.34-2.04)
ALD	83 / 3,778	37,454.0	221.6	2.80 (2.20-3.50)	1.75 (1.38-2.22)	1.61 (1.27-2.04)
Other	37 / 1,747	17,335.4	213.4	2.70 (1.90-3.80)	1.88 (1.35-2.63)	1.80 (1.29-2.52)
Myocardial infarction						
No-SLD	64 / 68,492	685,866.6	9.3	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	162 / 43145	432,578.2	37.5	4.00 (3.50-5.40)	2.45 (1.81-3.31)	2.27 (1.68-3.10)
MASLD	132 / 32,018	321,222.1	41.1	4.70 (3.30-6.00)	2.70 (1.98-3.70)	2.53 (1.86-3.45)
MetALD	14 / 5,602	55,994.8	25.0	2.70 (1.50-4.80)	1.64 (0.91-3.00)	1.46 (0.81-2.64)
ALD	8 / 3,778	37,860.4	21.1	2.30 (1.10-4.70)	1.23 (0.58-2.60)	1.06 (0.51-2.24)
Other	8 / 1,747	17,500.8	45.7	4.90 (2.40-10.20)	2.97 (1.42-6.20)	2.72 (1.30-5.71)
Total stroke						
No-SLD	484 / 68,492	683,688.7	70.8	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	549 / 43,145	430594.7	127.5	1.80 (1.60-2.00)	1.50 (1.30-1.70)	1.44 (1.26-1.64)
MASLD	406 / 32,018	319,779.3	126.9	1.80 (1.60-2.10)	1.50 (1.28-1.70)	1.43 (1.25-1.65)
MetALD	60 / 5,602	55,763.8	107.6	1.50 (1.20-2.00)	1.40 (1.04-1.80)	1.27 (0.96-1.68)
ALD	63 / 3,778	37,612.2	167.5	2.40 (1.80-3.10)	1.90 (1.46-2.50)	1.72 (1.31-2.26)
Other	20 / 1,747	17,439.4	114.7	1.60 (1.00-2.50)	1.40 (0.88-2.20)	1.31 (0.84-2.06)
Thrombotic stroke						
No-SLD	166 / 68,492	685,293.5	24.2	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	244 / 43,145	432173.4	56.5	2.30 (1.90-2.80)	1.66 (1.35-2.05)	1.61 (1.31-1.99)
MASLD	182 / 32,018	320,956.1	56.7	2.30 (1.90-2.90)	1.66 (1.33-2.07)	1.63 (1.31-2.03)
MetALD	27 / 5,602	55,925.0	48.3	2.00 (1.30-3.00)	1.53 (1.01-2.33)	1.44 (0.95-2.18)
ALD	26 / 3,778	37,785.7	68.8	2.80 (1.90-4.30)	1.89 (1.24-2.88)	1.72 (1.12-2.63)
Other	9 / 1,747	17,506.6	51.4	2.10 (1.10-4.20)	1.55 (0.79-3.05)	1.47 (0.75-2.89)
Hemorrhagic stroke						
No-SLD	91 / 68,492	685,695.5	13.3	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	104 / 43,145	432,841.2	24.0	1.80 (1.40-2.40)	1.35 (0.99-1.83)	1.30 (0.96-1.76)
MASLD	82 / 32,018	321,438.2	25.5	1.93 (1.43-2.60)	1.43 (1.04-1.96)	1.40 (1.01-1.92)
MetALD	6 / 5,602	56,041.4	10.7	0.81 (0.35-1.80)	0.62 (0.27-1.42)	0.57 (0.25-1.33)
ALD	11 / 3,778	37,842.5	29.1	2.20 (1.17-4.10)	1.53 (0.80-2.90)	1.36 (0.71-2.59)
Other	5 / 1,747	17,519.1	28.5	2.16 (0.88-5.30)	1.61 (0.65-3.98)	1.52 (0.61-3.77)

All cause mortality						
No-SLD	227 / 68,492	754,431.3	30.1	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	244 / 43145	476,312.9	51.2	1.70 (1.40-2.00)	1.04 (0.87-1.26)	1.00 (0.83-1.21)
MASLD	155 / 32,018	353,749.4	43.8	1.50 (1.20-1.80)	0.90 (0.73-1.11)	0.87 (0.71-1.08)
MetALD	36 / 5,602	61,638.5	58.4	1.90 (1.40-2.80)	1.24 (0.87-1.78)	1.13 (0.79-1.62)
ALD	30 / 3,778	41,650.1	72.0	2.40 (1.60-3.50)	1.31 (0.89-1.93)	1.15 (0.78-1.70)
Other	23 / 1,747	19,274.8	119.3	4.00 (2.60-6.10)	2.50 (1.62-3.85)	2.33 (1.51-3.59)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.
[a]Rate per 100,000 person-years.
Model 1 was unjusted.
Model 2 was adjusted for age sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR.

3. CVD risk according to SLD subtypes and advanced liver fibrosis

According to the subclassification of SLD, individuals with a FIB-4 index ≥ 2.67 were categorized into four groups. After adjusting for sociodemographic factors (i.e., sex, age, exercise status, smoking status, and insurance) and eGFR, compared to the no-SLD group, the MASLD & FIB-4 ≥ 2.67 group (HR 2.14, 95% CI 1.42–3.25) and the ALD & FIB-4 ≥ 2.67 group (HR 2.47, 95% CI 1.32–4.62) showed significantly increased risks.

In contrast, the MetALD & FIB-4 ≥ 2.67 group (HR 0.63, 95% CI 0.16–2.53) and the Other & FIB-4 ≥ 2.67 group (HR 1.69, 95% CI 0.80–3.56) did not show statistical significance (Table 5).

Table 5. CVD risk according to SLD subtypes and advanced liver fibrosis

Group	Events	Person-years	Rate[a]	HR (95% CI)		
				Model 1	Model 2	Model 3
No-SLD	1,655 / 68,492	677,208.1	244.4	1.00 (ref)	1.00 (ref)	1.00 (ref)
MASLD	1,679 / 32,018	312,838.1	536.7	2.20 (2.10-2.40)	1.72 (1.60-1.85)	1.69 (1.57-1.81)
MASLD, FIB-4 <2.67	1,656 / 31,903	311,809.3	531.1	2.20 (2.00-2.30)	1.71 (1.59-1.84)	1.68 (1.57-1.81)
MASLD, FIB-4 ≥2.67	23 / 115	1,028.8	2,235.5	9.20 (6.10-13.90)	2.23 (1.47-3.37)	2.14 (1.42-3.25)
No-SLD	1,655 / 68,492	677,208.1	244.4	1.00 (ref)	1.00 (ref)	1.00 (ref)
MetALD	285 / 5,602	54,568.7	522.3	2.10 (1.90-2.40)	1.83 (1.60-2.09)	1.75 (1.53-2.00)
MetALD, FIB-4 <2.67	283 / 5,562	54,189.0	522.3	2.10 (1.88-2.40)	1.85 (1.62-2.11)	1.77 (1.55-2.03)
MetALD, FIB-4 ≥2.67	2 / 40	379.7	526.7	2.10 (0.53-8.50)	0.70 (0.17-2.79)	0.63 (0.16-2.53)
No-SLD	1,655 / 68,492	677,208.1	244.4	1.00 (ref)	1.00 (ref)	1.00 (ref)
ALD	223 / 3,778	36,743.0	606.9	2.50 (2.20-2.90)	1.90 (1.62-2.18)	1.76 (1.51-2.04)
ALD, FIB-4 <2.67	213 / 3,733	36,367.9	585.7	2.40 (2.10-2.80)	1.90 (1.60-2.15)	1.73 (1.49-2.01)
ALD, FIB-4 ≥2.67	10 / 45	375.1	2,666.1	10.90 (5.80-20.20)	2.70 (1.43-5.01)	2.47 (1.32-4.62)
No-SLD	1,655 / 68,492	677,208.1	244.4	1.00 (ref)	1.00 (ref)	1.00 (ref)
Other	84 / 1,747	17,082.9	491.7	2.00 (1.60-2.50)	1.60 (1.32-2.06)	1.60 (1.28-2.00)
Other, FIB-4 <2.67	77 / 1,682	16,479.3	467.3	1.90 (1.50-2.40)	1.60 (1.30-2.07)	1.59 (1.26-2.01)
Other, FIB-4 ≥2.67	7 / 65	603.7	1,159.6	4.70 (2.20-9.80)	1.70 (0.80-3.54)	1.69 (0.80-3.56)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.

[a]Rate per 100,000 person-years.

Model 1 was unjusted.

Model 2 was adjusted for age sex.

Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR.

4. Sensitivity analysis

First, MASLD, SLD, and individual entities were associated with a higher risk of CVD than the absence of MASLD and SLD, regardless of age, sex, smoking status, obesity, TYG index, T2DM, hypertension, or dyslipidemia (Table 6).

Second, using a higher FLI threshold of ≥60 to define SLD, the association between MASLD or SLD and CVD risk remained similar to previous findings. After multivariable adjustment, the HR for CVD was 1.63 (95% CI 1.51–1.76) in the SLD group and 1.65 (95% CI 1.52–1.80) in the MASLD group, compared with the no-SLD group (Table 7).

Third, to reduce the risk of reverse causation and exclude events potentially driven by subclinical disease present at baseline, a sensitivity analysis was performed, excluding outcomes that occurred within the first three years of follow-up. After multivariable adjustment, the hazard ratio for CVD was 1.67 (95% CI 1.54–1.81) in the SLD group and 1.66 (95% CI 1.53–1.81) in the MASLD group, compared with the no-SLD group (Table 8).

Fourth, the observed hazard ratios were smaller than the calculated E values across all groups (Table 9).

Finally, the predictive performances of the models were evaluated using a concordance index (C-index). The C-index improved from 0.598 (0.590–0.605) in Model 1 to 0.760 (0.752–0.768) in Model 2 and further to 0.763 (0.755–0.771) in Model 3 (Table 10).

Table 6. Subgroup analysis of the association of the CVD and the presence of SLD

Group	Events	Person-years	Rate	HR (95% CI)	Events	Person-years	Rate	HR (95% CI)
				BMI<23 kg/m^2				**BMI≥23 kg/m^2**
No-SLD	898 / 46,912	463,283.8	193.8	1.00 (ref)	757 / 21,580	213,924.3	353.8	1.00 (ref)
SLD	178 / 3,644	35480.7	501.7	1.51 (1.27-1.80)	2,093 / 39,501	385,752.1	542.6	1.49 (1.36-1.63)
MASLD	121 / 2,355	22,917.9	527.9	1.64 (1.35-2.00)	1,558 / 29,663	289,920.2	537.4	1.47 (1.34-1.62)
MetALD	18 / 419	4,089.3	440.2	1.25 (0.78-2.00)	267 / 5,183	50,479.4	528.9	1.55 (1.34-1.79)
ALD	24 / 423	4,090.2	586.7	1.45 (0.96-2.20)	199 / 3,355	32,652.7	609.4	1.58 (1.35-1.86)
Other	15 / 447	4,383.2	342.2	1.10 (0.66-1.80)	69 / 1,300	12,699.7	543.3	1.56 (1.22-2.00)
				AGE<60 years				**AGE≥60 years**
No-SLD	1,323 / 66,260	656,127.6	201.6	1.00 (ref)	332 / 2,232	21,080.5	1,574.9	1.00 (ref)
SLD	1,863 / 41,221	406,875.3	461.3	2.11 (1.95-2.29)	408 / 1,924	17,357.5	2,350.6	1.42 (1.23-1.60)
MASLD	1,360 / 30,485	298,946.6	454.9	2.10 (1.93-2.09)	319 / 1,533	13,891.5	2,296.4	1.40 (1.20-1.60)
MetALD	246 / 5,429	53,022.9	463.9	2.11 (1.83-2.44)	39 / 173	1,545.8	2,522.9	1.47 (1.05-2.10)
ALD	190 / 3,616	35,304.3	538.1	2.37 (2.02-2.78)	33 / 162	1,438.7	2,293.7	1.33 (0.92-1.90)
Other	67 / 1,691	16,601.3	403.5	1.85 (1.44-2.37)	17 / 56	481.5	3,530.3	2.13 (1.31-3.50)
				Female				**Male**
No-SLD	780 / 36,238	357,977.2	217.8	1.00 (ref)	875 / 32,254	319,230.9	274.1	1.00 (ref)
SLD	291 / 3,654	35,333.3	823.6	1.61 (1.40-1.86)	1,980 / 39,491	385899.5	513.1	1.70 (1.57-1.80)
MASLD	264 / 3,237	31,304.4	843.3	1.60 (1.38-1.85)	1,415 / 28,781	281,533.8	502.6	1.69 (1.55-1.80)
MetALD	10 / 231	2,248.6	444.7	1.37 (0.73-2.55)	275 / 5,371	52,320.1	525.6	1.75 (1.52-2.00)
ALD	7 / 84	803.8	870.8	2.11 (1.00-4.46)	216 / 3,694	35,939.2	601.0	1.76 (1.51-2.00)
Other	10 / 102	976.5	1,024.1	1.92 (1.03-3.59)	74 / 1,645	16,106.4	459.4	1.55 (1.22-2.00)
				Non-smoker				**Smoker**
No-SLD	909 / 41,779	411,619.0	220.8	1.00 (ref)	746 / 26,713	265,589.1	280.8	1.00 (ref)
SLD	581 / 10,740	104,162.4	557.8	1.65 (1.47-1.85)	1,690 / 32,405	317,070.4	533.0	1.72 (1.57-1.90)
MASLD	494 / 9,043	87,758.5	562.9	1.63 (1.45-1.83)	1,185 / 22,975	225,079.6	526.4	1.71 (1.56-1.90)
MetALD	42 / 841	8,122.3	517.1	1.86 (1.35-2.55)	243 / 4,761	46,446.4	523.1	1.73 (1.50-2.00)
ALD	22 / 432	4,172.3	527.3	1.57 (1.02-2.41)	201 / 3,346	32,570.6	617.1	1.82 (1.55-2.10)
Other	23 / 424	4,109.2	559.7	1.81 (1.20-2.75)	61 / 1,323	12,973.7	470.1	1.53 (1.18-2.00)
				Tyg index<8.84				**Tyg index≥8.84**
No-SLD	1,401 / 62,944	622,543.5	225.0	1.00 (ref)	254 / 5548	54,664.6	464.6	1.00 (ref)
SLD	775 / 15,751	153,983.9	503.3	1.71 (1.56-1.88)	1,496 / 27,394	267,248.8	559.8	1.32 (1.15-1.50)
MASLD	573 / 11,538	112,843.0	507.7	1.74 (1.57-1.92)	1106 / 20480	199,995.2	553.0	1.31 (1.13-1.50)
MetALD	97 / 1,916	18,622.2	520.8	1.86 (1.50-2.29)	188 / 3686	35,946.5	523.0	1.31 (1.07-1.60)
ALD	58 / 1,206	11,830.7	490.2	1.50 (1.15-1.96)	165 / 2572	24,912.3	662.3	1.46 (1.19-1.80)
Other	47 / 1,091	10,688.1	439.7	1.48 (1.10-1.98)	37 / 656	6,394.8	578.6	1.43 (1.01-2.00)

				No hypertension				Hypertension
No-SLD	1,342 / 64,925	642,491.9	208.8	1.00 (ref)	313 / 3,567	34,716.2	901.6	1.00 (ref)
SLD	1,538 / 35,267	345,434.8	445.2	1.65 (1.53-1.79)	733 / 7,878	75797.9	967.1	1.21 (1.05-1.40)
MASLD	1,148 / 26,400	258,728.0	443.7	1.65 (1.52-1.79)	531 / 5,618	54,110.1	981.3	1.22 (1.06-1.40)
MetALD	188 / 4,452	43,495.6	432.2	1.70 (1.45-1.99)	97 / 1,150	11,073.1	876.0	1.13 (0.89-1.40)
ALD	140 / 2,922	28,574.9	489.9	1.69 (1.41-2.02)	83 / 856	8,168.1	1,016.1	1.19 (0.92-1.50)
Other	62 / 1,493	14,636.2	423.6	1.60 (1.24-2.07)	22 / 254	2,446.6	899.1	1.14 (0.74-1.80)
				No type2 diabetes				**Type2 diabetes**
No-SLD	1,408 / 66,201	655,135.6	214.9	1.00 (ref)	247 / 2291	22,072.5	1,119.0	1.00 (ref)
SLD	1,661 / 37,726	369,413.1	449.6	1.68 (1.55-1.81)	610 / 5,419	51,819.7	1,177.2	1.16 (1.00-1.40)
MASLD	1,218 / 27,960	274,055.1	444.4	1.66 (1.53-1.81)	461 / 4058	38,783.0	1,188.6	1.18 (1.01-1.40)
MetALD	212 / 4,931	48,129.9	440.5	1.75 (1.50-2.03)	73 / 671	6,438.8	1,133.7	1.12 (0.85-1.50)
ALD	159 / 3,279	32,013.5	496.6	1.69 (1.42-2.00)	64 / 499	4,729.4	1,353.2	1.31 (0.99-1.70)
Other	72 / 1,556	15,214.5	473.2	1.76 (1.38-2.24)	12 / 191	1,868.4	642.2	0.69 (0.38-1.20)
				No dyslipidemia				**Dyslipidemia**
No-SLD	1,645 / 68,168	674,023.0	244.1	1.00 (ref)	10 / 324	3,185.1	313.9	1.00 (ref)
SLD	2,052 / 39,314	384,132.7	534.2	1.66 (1.55-1.78)	219 / 3,831	37,100.1	590.3	2.31 (1.21-4.40)
MASLD	1,505 / 29,008	283,683.9	530.5	1.65 (1.54-1.78)	174 / 3,010	29,154.2	596.8	2.33 (1.22-4.50)
MetALD	263 / 5,171	50,399.0	521.8	1.72 (1.50-1.96)	22 / 431	4,169.7	527.6	2.16 (1.00-4.60)
ALD	205 / 3,484	33,906.9	604.6	1.73 (1.49-2.01)	18 / 294	2,836.1	634.7	2.39 (1.09-5.30)
Other	79 / 1651	16,142.9	489.4	1.57 (1.25-1.97)	5 / 96	940.0	531.9	2.01 (0.68-5.90)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.

[a]Rate per 100,000 person-years.

Model was further adjusted for age, sex, smoking status, exercise status, insurance, and eGFR.

Table 7. CVD risk according to SLD subtypes using higher cut-off FLI≥60 to define SLD

Group	Events	Person-years	Rate[a]	HR (95% CI)		
				Model 1	Model 2	Model 3
No-SLD	2,957 / 94,618	933,328.8	316.8	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	969 / 17,019	165,112.1	586.9	1.80 (1.70-2.00)	1.70 (1.57-1.83)	1.63 (1.51-1.76)
MASLD	678 / 11,972	116,271.7	583.1	1.80 (1.70-2.00)	1.70 (1.56-1.85)	1.65 (1.52-1.80)
MetALD	141 / 2,642	25,614.7	550.5	1.70 (1.50-2.00)	1.65 (1.39-1.96)	1.56 (1.31-1.85)
ALD	128 / 1,972	19,007.3	673.4	2.10 (1.80-2.50)	1.72 (1.44-2.06)	1.61 (1.34-1.92)
Other	22 / 433	4,218.4	521.5	1.60 (1.10-2.50)	1.78 (1.17-2.70)	1.70 (1.12-2.59)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.
[a]Rate per 100,000 person-years.
Model 1 was unjusted.
Model 2 was adjusted for age and sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR.

Table 8. CVD risk according to SLD subtypes (follow up >3 years)

Group	Events	Person-years	Rate[a]	HR (95% CI)		
				Model 1	Model 2	Model 3
No-SLD	1,240 / 68,076	676,356.4	183.3	1.00 (ref)	1.00 (ref)	1.00 (ref)
SLD	1,674 / 42,548	419,976.9	398.6	2.20 (2.00-2.30)	1.72 (1.59-1.86)	1.67 (1.54-1.81)
MASLD	1,230 / 31,569	311,894.2	394.4	2.20 (2.00-2.30)	1.69 (1.60-1.84)	1.66 (1.53-1.81)
MetALD	214 / 5,531	54,416.3	393.3	2.20 (1.90-2.50)	1.83 (1.60-2.12)	1.72 (1.48-2.00)
ALD	171 / 3,726	36,638.5	466.7	2.60 (2.20-3.00)	1.93 (1.60-2.28)	1.79 (1.51-2.11)
Other	59 / 1,722	17,027.8	346.5	1.90 (1.50-2.50)	1.54 (1.20-2.01)	1.49 (1.14-1.94)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.
[a]Rate per 100,000 person-years.
Model 1 was unjusted.
Model 2 was adjusted for age and sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR.

Table 9. E-values for the CVD risk according to SLD subtypes

Group	HR (95% CI)	E-value (95% CI)
No-SLD	1.00 (ref)	1.00 (ref)
SLD	1.70 (1.58-1.82)	2.78 (2.54-NA)
MASLD	1.69 (1.57-1.82)	2.77 (2.52-NA)
MetALD	1.72 (1.82-1.97)	2.85 (2.40-NA)
ALD	1.75 (1.52-2.03)	2.91 (2.41-NA)
Other	1.58 (1.27-1.98)	2.55 (1.86-NA)

MASLD, metabolic dysfunction-associated steatotic liver disease; MetALD, MASLD with increased alcohol intake; ALD, alcohol-related liver disease; Other, MASLD with other combined aetiology.
Model was adjusted for age, sex, smoking status, exercise status, insurance, and eGFR.

Table 10. C-index for the CVD risk according to SLD

		C-index (95% CI)
CVD	Model 1	0.598 (0.590-0.605)
	Model 2	0.759 (0.752-0.768)
	Model 3	0.763 (0.755-0.771)

Model 1 was unjusted.
Model 2 was adjusted for age sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR.

PART II. Causal association between MASLD and CVD: genetic analysis

1. One-sample Mendelian randomization study

In the one-sample MR using individual-level data from the KCPS-II cohort, HRs were estimated for various CVD outcomes according to the genetically predicted MASLD. In the fully adjusted Model 3, the overall CVD risk was 1.03 (95% CI 1.00–1.07). No statistically significant associations were observed for IHD (HR 1.03, 95% CI 0.97–1.10), MI (HR 0.91, 95% CI 0.77–1.06), total stroke (HR 1.03, 95% CI 0.96–1.11), thrombotic stroke (HR 0.94, 95% CI 0.84−1.06), and hemorrhagic stroke (HR 1.03, 95% CI 0.87–1.22) (Table 11). This analysis categorized the participants into low-risk (quartiles 1–3) and high-risk (quartile 4) groups based on their MASLD WGRS scores. In the fully adjusted model, no significant associations were observed between the high- and low-risk groups for any primary or secondary cardiovascular outcomes (Table 12).

However, when the LD clumping threshold was set to a lower p-value of 5×10^{-5}, and 423 SNPs were selected to construct the WGRS, statistically significant associations emerged for CVD (HR 1.12, 95% CI 1.04–1.20) and IHD (HR 1.15, 95% CI 1.02–1.30) (Table 13 & Table 14).

Table 11. Associations between MASLD WGRS and risk of CVD and secondary outcomes using LD clumping thresholds of 5×10^{-8}

Group	Events	HR (95% CI)		
		Model 1	Model 2	Model 3
Cardiovascular disease	3,926 / 111,637	1.00 (0.98-1.10)	1.00 (0.99-1.07)	1.03 (1.00-1.07)
Ischemic heart disease	1,425 / 111,637	1.00 (0.95-1.10)	1.03 (0.97-1.10)	1.03 (0.97-1.10)
Myocardial infarction	226/ 111,637	0.88 (0.75-1.00)	0.90 (0.77-1.05)	0.91 (0.77-1.06)
Total stroke	1,033 / 111,637	1.00 (0.94-1.10)	1.03 (0.96-1.11)	1.03 (0.96-1.11)
Thrombotic stroke	410 / 111,637	0.91 (0.81-1.00)	0.94 (0.83-1.06)	0.94 (0.84-1.06)
Hemorrhagic stroke	195/ 111,637	1.00 (0.86-1.20)	1.00 (0.87-1.22)	1.03 (0.87-1.22)

Model 1 was unadjusted.
Model 2 was adjusted for age and sex.
Model 3 was further adjusted for smoking status, alcohol status, exercise status, insurance, and eGFR.

Table 12. Association of low vs. high MASLD WGRS groups with risk of CVD and secondary outcomes using LD clumping thresholds of 5×10^{-8}

Group	Events	HR (95% CI)		
		Model 1	Model 2	Model 3
Cardiovascular disease				
Low risk	2,959 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	967 / 27,907	0.98 (0.91-1.10)	1.00 (0.93-1.07)	1.00 (0.93-1.07)
Ischemic stroke				
Low risk	1,063 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	362 / 27,907	1.00 (0.91-1.20)	1.05 (0.93-1.18)	1.05 (0.93-1.20)
Myodardial infarction				
Low risk	172 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	54 / 27,907	0.94 (0.69-1.30)	0.97 (0.71-1.32)	0.98 (0.72-1.33)
Total stroke				
Low risk	784 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	249 / 27,907	0.95 (0.83-1.10)	0.97 (0.84-1.12)	0.97 (0.84-1.12)
Thrombotic stroke				
Low risk	321 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	89 / 27,907	0.83 (0.66-1.10)	0.86 (0.68-1.09)	0.86 (0.68-1.09)
Hemorrhagic stroke				
Low risk	150 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	45 / 27,907	0.90 (0.65-1.30)	0.92 (0.66-1.28)	0.92 (0.66-1.28)

Model 1 was unadjusted.
Model 2 was further adjusted for age, sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR

Table 13. Associations between MASLD WGRS and risk of CVD and secondary outcomes using LD clumping thresholds of 5×10^{-5}

Group	Events	HR (95% CI)		
		Model 1	Model 2	Model 3
Cardiovascular disease	3,926 / 111,637	1.00 (1.00-1.10)	1.10 (1.03-1.08)	1.05 (1.03-1.08)
Ischemic heart disease	1,397 / 111,637	1.00 (1.00-1.10)	1.06 (1.02-1.10)	1.06 (1.02-1.10)
Myocardial infarction	219/ 111,637	0.96 (0.86-1.10)	0.97 (0.88-1.07)	0.97 (0.88-1.08)
Total stroke	1,020 / 111,637	1.00 (0.97-1.10)	1.03 (0.98-1.08)	1.03 (0.98-1.08)
Thrombotic stroke	403 / 111,637	0.97 (0.90-1.00)	0.99 (0.92-1.07)	0.99 (0.92-1.07)
Hemorrhagic stroke	195/ 111,637	1.10 (0.96-1.20)	1.10 (0.97-1.21)	1.08 (0.97-1.21)

Model 1 was unadjusted.
Model 2 was adjusted for age and sex.
Model 3 was further adjusted for smoking status, alcohol status, exercise status, insurance, and eGFR.

Table 14. Association of low vs. high MASLD WGRS groups with risk of CVD and secondary outcomes using LD clumping thresholds of 5×10^{-5}

Group	Events	HR (95% CI)		
		Model 1	Model 2	Model 3
Cardiovascular disease				
Low risk	2,880 /83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	1,046 /27,907	1.10 (1.00-1.20)	1.10 (1.04-1.20)	1.12 (1.04-1.20)
Ischemic stroke				
Low risk	1,038 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	387 / 27,907	1.10 (1.00-1.30)	1.15 (1.02-1.29)	1.15 (1.02-1.30)
Myodardial infarction				
Low risk	171 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	55 / 27,907	0.97 (0.74-1.40)	0.99 (0.73-1.34)	0.99 (0.73-1.34)
Total stroke				
Low risk	761 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	272 / 27,907	1.20 (0.93-1.20)	1.10 (0.96-1.26)	1.10 (0.96-1.26)
Thrombotic stroke				
Low risk	303 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	107 / 27,907	1.10 (0.85-1.30)	1.10 (0.88-1.37)	1.10 (0.99-1.37)
Hemorrhagic stroke				
Low risk	137 / 83,730	1.00 (ref)	1.00 (ref)	1.00 (ref)
High risk	58 / 27,907	1.30 (0.93-1.70)	1.30 (0.95-1.75)	1.29 (0.99-1.76)

Model 1 was unadjusted.
Model 2 was further adjusted for age, sex.
Model 3 was further adjusted for smoking status, exercise status, insurance, and eGFR

2. Two-sample Mendelian randomization study

2-1. Univariable two-sample MR

MR analysis was conducted using multiple methods to explore the causal associations between MASLD and CAD. In the initial analysis using 139 SNPs, the IVW method did not show statistical significance (OR 1.06, 95% CI 0.99–1.14, p = 0.06) (Table 15). Based on the Radial MR, the analysis was repeated for the 120 SNPs. In this refined analysis, the IVW method showed a statistically significant positive association between MASLD and CAD (OR 1.08, 95% CI 1.05–1.13, $p = 1.83\times10^{-5}$) (Table 16).

Additional MR analyses evaluated the causal associations among MASLD, IS, and MI. For IS, the IVW method indicated a statistically significant association (OR 1.04, 95% CI 1.00–1.09, p = 0.032). In contrast, the MR-Egger method did not show a significant effect (OR 1.03, 95% CI 0.92–1.15, p = 0.563), suggesting limited evidence of directional pleiotropy. In the analysis of MI, the IVW method demonstrated a significant positive association between MASLD and MI risk (OR 1.12, 95% CI 1.06–1.18, $p = 2.32 \times 10^{-5}$), whereas the MR-Egger method showed no significant association (OR 1.06, 95% CI 0.92–1.23, p = 0.380) (Table 16). Notably, in all analyses, the tests for pleiotropy and heterogeneity yielded p-values $\geq$ 0.05, suggesting no evidence of directional pleiotropy or significant heterogeneity.

The robustness of the causal associations between MASLD and various CVD outcomes was further evaluated using bidirectional MR analyses. When MASLD was defined in KCPS-II and CVD outcomes were derived from UK Biobank data, the IVW method showed a statistically significant association with CVD (OR 1.04, 95% CI 1.01–1.08, p = 0.006), IHD (OR 1.08, 95% CI 1.02–1.14, p = 0.009), and total stroke (OR 1.10, 95% CI 1.04–1.18, p = 0.002). In contrast, no significant associations were observed for MI (OR 1.01, 95% CI 0.95–1.08, p = 0.642), thrombotic stroke (OR 0.98, 95% CI 0.77–1.24, p = 0.874), or hemorrhagic stroke (OR 1.11, 95% CI 1.00–1.24, p = 0.059) (Table 17 & Table 18). The MR-Egger method did not yield significant associations for these outcomes, suggesting no strong evidence of directional pleiotropy.

In the reverse direction, when MASLD was defined as in the UK Biobank and CVD outcomes were drawn from KCPS-II, the IVW method identified significant associations with CVD (OR 1.15, 95% CI 1.08–1.24, $p = 4.39\times10^{-5}$), IHD (OR 1.15, 95% CI 1.04–1.27, p = 0.007), MI (OR = 1.78, 95% CI 1.42–2.23, $p = 4.89\times10^{-7}$), and thrombotic stroke (OR = 1.26, 95% CI 1.05–1.52, p = 0.015). A nominally significant association was observed for total stroke (OR 1.15, 95% CI 1.01–1.31, p = 0.027), whereas no association was found for hemorrhagic stroke (OR 1.12, 95% CI 0.86–1.45, p = 0.372) (Table 19 & Table 20).

Table 15. Univariable two-sample MR analysis of MASLD with BBJ outcome

Outcome	Methods	nSNP	*B*	*SE*	OR (95% CI)	*P* value
Coronary artery disease	MR-Egger	139	-0.121	0.079	0.88 (0.75-1.03)	1.30E-01
	Weighted median	139	0.105	0.030	1.11 (1.04-1.18)	6.11E-04
	Inverse variance weighted	139	0.065	0.035	1.06 (0.99-1.14)	6.55E-02
	Simple mode	139	0.059	0.078	1.06 (0.91-1.24)	4.52E-01
	Weighted mode	139	0.133	0.056	1.14 (1.02-1.27)	2.00E-02
Ischemic stroke	MR-Egger	139	0.127	0.049	1.13 (1.03-1.25)	1.07E-02
	Weighted median	139	0.025	0.029	0.96 (0.97-1.09)	4.05E-01
	Inverse variance weighted	139	0.067	0.022	1.03 (1.02-1.11)	1.99E-03
	Simple mode	139	0.055	0.077	0.91 (0.91-1.23)	4.70E-01
	Weighted mode	139	-0.021	0.041	0.90 (0.90-1.06)	6.00E-01
Myocardial infarction	MR-Egger	139	-0.314	0.115	0.73 (0.58-0.91)	7.03E-03
	Weighted median	139	0.199	0.041	1.22 (1.12-1.32)	1.04E-06
	Inverse variance weighted	139	0.038	0.052	1.04 (0.94-1.15)	4.60E-01
	Simple mode	139	0.159	0.102	1.17 (0.95-1.43)	1.22E-01
	Weighted mode	139	0.245	0.054	1.27 (1.15-1.42)	1.19E-05

nSNP, the number of SNPs; *B*, beta; SE, standard error; and OR, odds ratio.

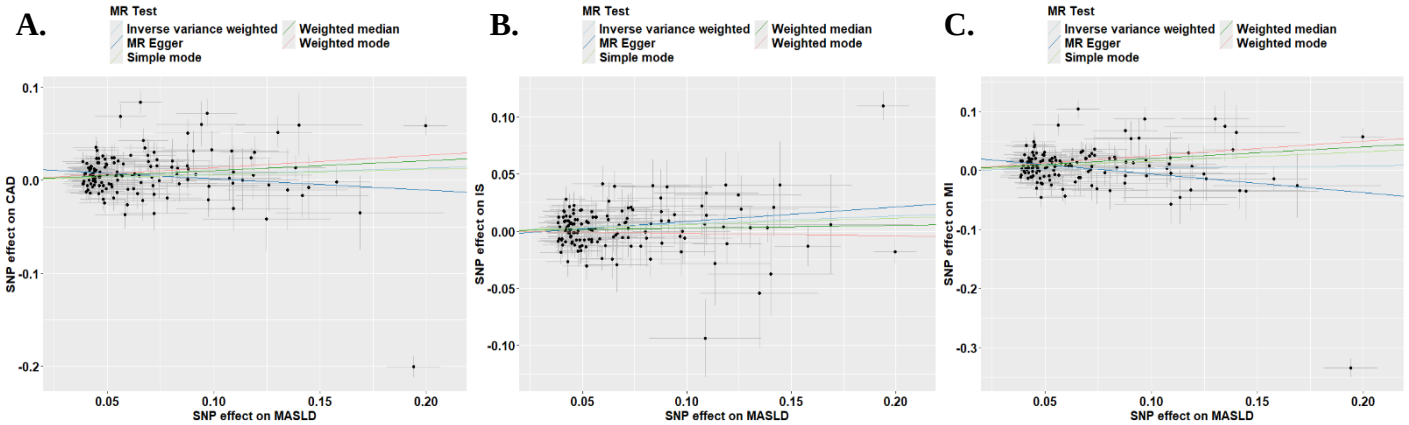


Figure 2. Scatter plots from two-sample MR analyses of MASLD (KCPS-II) on CAD, IS, and MI (BBJ)

Table 16. Radial MR analysis of MASLD with BBJ outcome

Outcome	Methods	nSNP	*B*	*SE*	OR (95% CI)	*P* value
Coronary artery disease	MR-Egger	120	0.063	0.053	1.06 (0.95-1.18)	2.40E-01
	Weighted median	120	0.076	0.031	1.08 (1.02-1.14)	1.22E-02
	Inverse variance weighted	120	0.083	0.019	1.08 (1.05-1.13)	1.83E-05
	Simple mode	120	0.063	0.079	1.06 (0.91-1.24)	4.25E-01
	Weighted mode	120	0.088	0.057	1.09 (0.97-1.22)	1.22E-01
Ischemic stroke	MR-Egger	130	0.032	0.055	1.03 (0.92-1.15)	5.63E-01
	Weighted median	130	0.032	0.029	1.03 (0.97-1.09)	2.86E-01
	Inverse variance weighted	130	0.042	0.019	1.04 (1.00-1.09)	3.23E-02
	Simple mode	130	0.059	0.078	1.06 (0.91-1.24)	4.46E-01
	Weighted mode	130	0.038	0.055	1.04 (0.93-1.15)	4.94E-01
Myocardial infarction	MR-Egger	123	0.064	0.073	1.06 (0.92-1.23)	3.80E-01
	Weighted median	123	0.136	0.041	1.15 (1.06-1.24)	1.07E-01
	Inverse variance weighted	123	0.111	0.026	1.12 (1.06-1.18)	2.32E-05
	Simple mode	123	0.139	0.107	1.15 (0.93-1.42)	1.96E-01
	Weighted mode	123	0.205	0.894	1.22 (1.05-1.43)	9.95E-03

nSNP, the number of SNPs; *B*, beta; SE, standard error; and OR, odds ratio.

Table 17. Univariable two-sample MR analysis using MASLD defined in KCPS-II with UKB

Outcome	Methods	nSNP	*B*	*SE*	OR (95% CI)	*P* value
Cardiovascular disease	MR-Egger	67	0.045	0.064	1.04 (0.92-1.18)	4.85E-01
	Weighted median	67	0.011	0.023	1.01 (0.96-1.05)	6.32E-01
	Inverse variance weighted	67	0.048	0.027	1.05 (0.99-1.11)	7.38E-02
	Simple mode	67	0.085	0.050	1.09 (0.98-1.20)	9.46E-02
	Weighted mode	67	-0.002	0.025	0.99 (0.95-1.05)	9.35E-01
Ischemic heart disease	MR-Egger	67	0.041	0.097	1.04 (0.96-1.26)	6.72E-01
	Weighted median	67	-0.031	0.033	0.96 (0.91-1.03)	3.42E-01
	Inverse variance weighted	67	0.062	0.041	1.06 (0.98-1.15)	1.25E-01
	Simple mode	67	-0.099	0.106	0.90 (0.73-1.11)	3.50E-01
	Weighted mode	67	-0.036	0.035	0.96 (0.89-1.03)	3.01E-01
Myocardial infarction	MR-Egger	67	0.065	0.109	1.06 (0.86-1.32)	5.52E-01
	Weighted median	67	-0.007	0.045	0.99 (0.91-1.08)	8.60E-01
	Inverse variance weighted	67	0.045	0.046	1.04 (0.95-1.14)	3.24E-01
	Simple mode	67	0.098	0.099	1.10 (0.91-1.34)	3.25E-01
	Weighted mode	67	-0.024	0.048	0.97 (0.88-1.07)	6.11E-01
Total stroke	MR-Egger	67	0.099	0.083	1.10 (0.94-1.30)	2.35E-01
	Weighted median	67	0.113	0.045	1.12 (1.02-1.22)	1.33E-02
	Inverse variance weighted	67	0.085	0.034	1.09 (1.02-1.17)	1.41E-02
	Simple mode	67	0.108	0.105	1.11 (0.91-1.37)	3.08E-01
	Weighted mode	67	0.122	0.070	1.13 (0.98-1.30)	8.73E-02
Thrombotic stroke	MR-Egger	67	0.058	0.285	1.06 (0.60-1.85)	8.37E-01
	Weighted median	67	0.092	0.194	1.09 (0.74-1.61)	6.35E-01
	Inverse variance weighted	67	0.022	0.121	1.02 (0.81-1.30)	8.50E-01
	Simple mode	67	-0.173	0.356	0.84 (0.42-1.69)	6.29E-01
	Weighted mode	67	0.070	0.258	1.07 (0.64-1.78)	7.87E-01
Hemorrhagic stroke	MR-Egger	67	0.046	0.149	1.05 (0.78-1.40)	7.54E-01
	Weighted median	67	0.121	0.092	1.13 (0.94-1.35)	1.88E-01
	Inverse variance weighted	67	0.112	0.063	1.12 (0.99-1.26)	7.46E-02
	Simple mode	67	0.332	0.201	1.39 (0.94-2.07)	1.02E-01
	Weighted mode	67	-0.001	0.156	1.00 (0.74-1.35)	9.97E-01

nSNP, the number of SNPs; *B*, beta; SE, standard error; and OR, odds ratio.

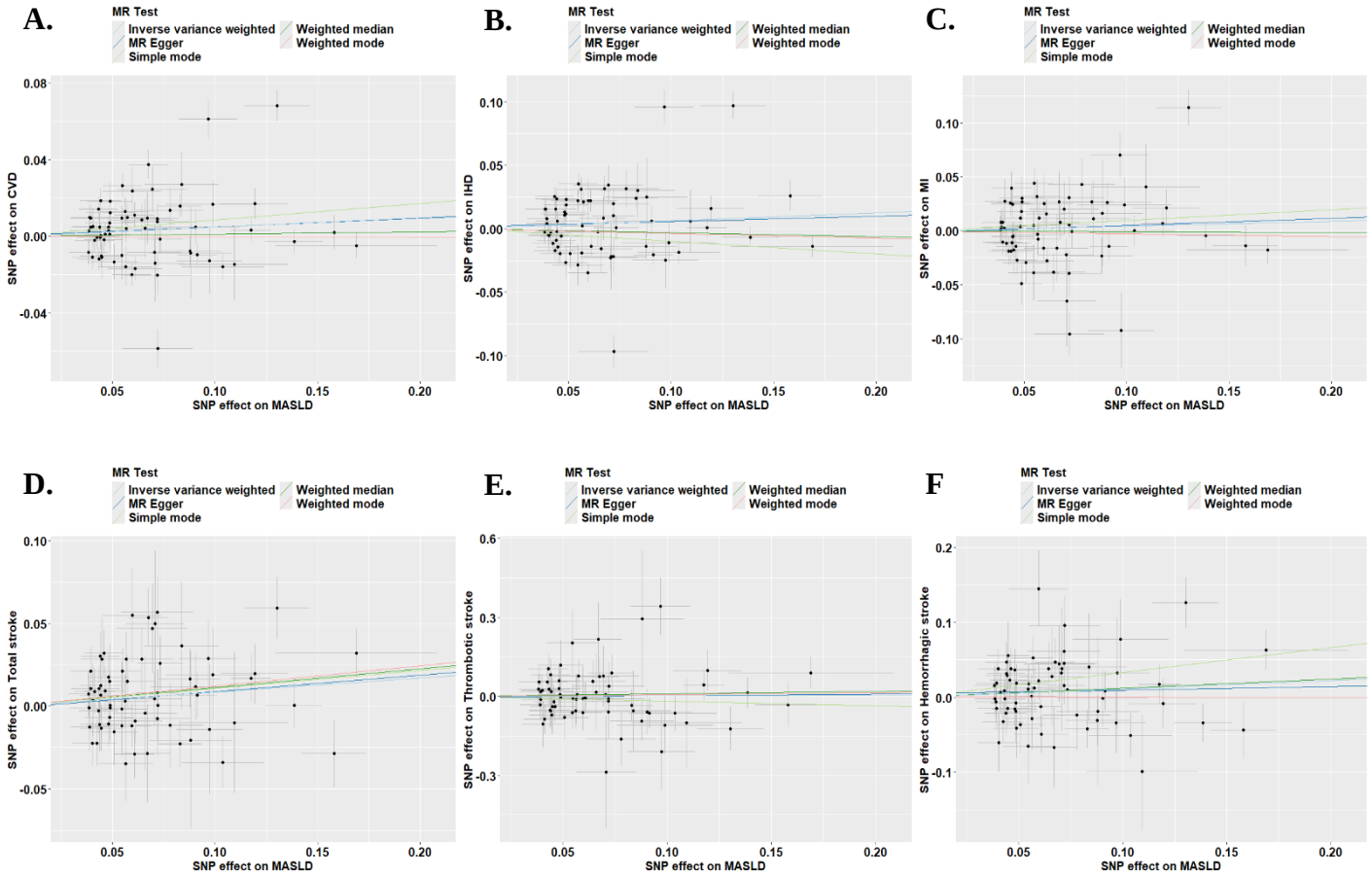


Figure 3. Scatter plots from two-sample MR analysis of MASLD (KCPS-II) on UKB outcomes

Table 18. Radial MR analysis using MASLD defined in KCPS-II with UKB

Outcome	Methods	nSNP	*B*	*SE*	OR (95% CI)	*P* value
Cardiovascular disease	MR-Egger	52	-0.014	0.035	0.98 (0.92-1.06)	6.93E-01
	Weighted median	52	0.017	0.023	1.02 (0.97-1.06)	4.65E-01
	Inverse variance weighted	52	0.044	0.016	1.04 (1.01-1.08)	6.03E-03
	Simple mode	52	0.104	0.057	1.11 (0.99-1.24)	7.45E-01
	Weighted mode	52	-0.003	0.026	0.99 (0.95-1.05)	9.19E-01
Ischemic heart disease	MR-Egger	43	0.088	0.072	1.09 (0.94-1.26)	2.34E-01
	Weighted median	43	0.077	0.038	1.08 (1.00-1.17)	4.44E-02
	Inverse variance weighted	43	0.075	0.029	1.08 (1.02-1.14)	8.90E-03
	Simple mode	43	0.107	0.083	1.11 (0.95-1.31)	2.04E-01
	Weighted mode	43	0.086	0.050	1.09 (0.99-1.20)	9.13E-02
Myocardial infarction	MR-Egger	54	-0.025	0.072	0.97 (0.85-1.12)	7.25E-01
	Weighted median	54	-0.014	0.046	0.98 (0.90-1.08)	7.61E-01
	Inverse variance weighted	54	0.015	0.031	1.01 (0.95-1.08)	6.42E-01
	Simple mode	54	0.121	0.089	1.13 (0.94-1.34)	1.83E-01
	Weighted mode	54	-0.037	0.051	0.96 (0.87-1.06)	4.63E-01
Total stroke	MR-Egger	59	0.169	0.077	1.18 (1.02-1.38)	3.29E-02
	Weighted median	59	0.134	0.049	1.14 (1.04-1.26)	6.79E-03
	Inverse variance weighted	59	0.099	0.032	1.10 (1.04-1.18)	2.28E-03
	Simple mode	59	0.132	0.097	1.14 (0.94-1.38)	1.76E-01
	Weighted mode	59	0.132	0.060	1.14 (1.01-1.28)	3.17E-02
Thrombotic stroke	MR-Egger	66	-0.020	0.286	0.98 (0.55-1.71)	9.44E-01
	Weighted median	66	0.081	0.190	1.08 (0.74-1.58)	6.70E-01
	Inverse variance weighted	66	-0.019	0.122	0.98 (0.77-1.24)	8.74E-01
	Simple mode	66	-0.173	0.369	0.84 (0.40-1.73)	6.40E-01
	Weighted mode	66	0.070	0.257	1.07 (0.64-1.77)	7.84E-01
Hemorrhagic stroke	MR-Egger	61	-0.005	0.133	0.99 (0.76-1.29)	9.64E-01
	Weighted median	61	0.139	0.092	1.15 (0.95-1.37)	1.31E-01
	Inverse variance weighted	61	0.109	0.057	1.11 (1.00-1.24)	5.91E-02
	Simple mode	61	0.358	0.193	1.43 (0.98-2.09)	6.79E-02
	Weighted mode	61	0.101	0.148	1.11 (0.83-1.48)	5.01E-01

nSNP, the number of SNPs; *B*, beta; SE, standard error; and OR, odds ratio.

Table 19. Univariable two-sample MR analysis using MASLD defined in UKB with KCPS-II outcomes

Outcome	Methods	nSNP	*B*	*SE*	OR (95% CI)	*P* value
Cardiovascular disease	MR-Egger	152	0.225	0.128	1.25 (0.97-1.61)	8.18E-02
	Weighted median	152	0.156	0.053	1.17 (1.05-1.30)	3.63E-03
	Inverse variance weighted	152	0.147	0.039	1.16 (1.07-1.25)	1.53E-04
	Simple mode	152	0.121	0.149	1.13 (0.84-1.51)	4.21E-01
	Weighted mode	152	0.187	0.124	1.21 (0.94-1.54)	1.35E-01
Ischemic heart disease	MR-Egger	152	0.579	0.180	1.78 (1.25-2.53)	1.55E-03
	Weighted median	152	0.258	0.074	1.30 (1.11-1.50)	5.47E-04
	Inverse variance weighted	152	0.172	0.055	1.19 (1.06-1.32)	1.89E-03
	Simple mode	152	0.220	0.205	1.25 (0.83-1.86)	2.83E-01
	Weighted mode	152	0.388	0.128	1.47 (1.15-1.89)	2.78E-03
Myocardial infarction	MR-Egger	152	0.823	0.373	2.28 (1.09-4.74)	2.90E-02
	Weighted median	152	0.628	0.181	1.87 (1.31-2.67)	5.28E-04
	Inverse variance weighted	152	0.448	0.113	1.56 (1.25-1.95)	7.96E-05
	Simple mode	152	0.148	0.409	1.16 (0.52-2.58)	7.17E-01
	Weighted mode	152	0.649	0.310	1.91 (1.04-3.51)	3.80E-02
Total stroke	MR-Egger	152	0.193	0.214	1.21 (0.79-1.84)	3.70E-01
	Weighted median	152	0.107	0.098	1.11 (0.92-1.35)	2.76E-01
	Inverse variance weighted	152	0.144	0.065	1.15 (1.01-1.31)	2.62E-02
	Simple mode	152	0.083	0.219	1.08 (0.71-1.67)	7.05E-01
	Weighted mode	152	0.129	0.162	1.13 (0.83-1.56)	4.24E-01
Thrombotic stroke	MR-Egger	152	0.325	0.326	1.38 (0.73-2.62)	3.19E-01
	Weighted median	152	0.145	0.142	1.16 (0.87-1.53)	3.05E-01
	Inverse variance weighted	152	0.156	0.099	1.17 (0.96-1.42)	1.14E-01
	Simple mode	152	0.064	0.365	1.07 (0.52-2.18)	8.60E-01
	Weighted mode	152	-0.047	0.327	0.95 (0.50-1.81)	8.85E-01
Hemorrhagic stroke	MR-Egger	152	0.014	0.425	1.01 (0.44-2.33)	9.74E-01
	Weighted median	152	0.067	0.197	1.07 (0.73-1.57)	7.33E-01
	Inverse variance weighted	152	0.170	0.129	1.18 (0.92-1.53)	1.89E-01
	Simple mode	152	-0.122	0.435	0.88 (0.37-2.08)	7.79E-01
	Weighted mode	152	-0.027	0.331	0.97 (0.51-1.86)	9.35E-01

nSNP, the number of SNPs; *B*, beta; SE, standard error; and OR, odds ratio.

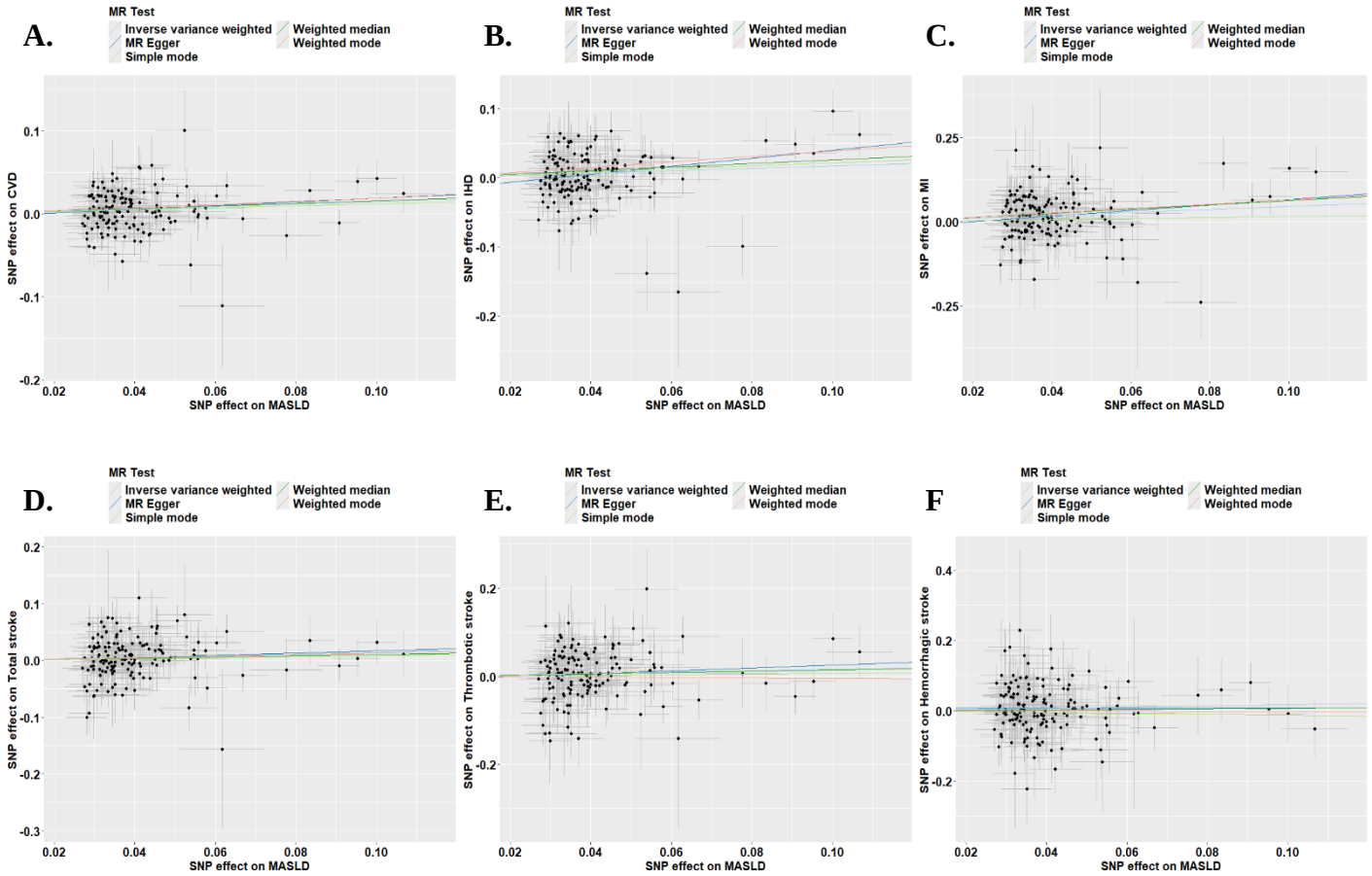


Figure 4. Scatter plots from two-sample MR analysis of MASLD (UKB) on KCPS-II outcomes

Table 20. Radial MR analysis using MASLD defined in UKB with KCPS-II outcomes

Outcome	Methods	nSNP	*B*	*SE*	OR (95% CI)	*P* value
Cardiovascular disease	MR-Egger	141	0.159	0.114	1.17 (0.93-1.46)	1.65E-01
	Weighted median	141	0.155	0.056	1.17 (1.04-1.31)	6.05E-03
	Inverse variance weighted	141	0.145	0.035	1.15 (1.08-1.24)	4.39E-05
	Simple mode	141	0.119	0.151	1.12 (0.84-1.52)	4.32E-01
	Weighted mode	141	0.194	0.131	1.21 (0.94-1.57)	1.43E-01
Ischemic heart disease	MR-Egger	140	0.590	0.176	1.80 (1.27-2.55)	1.07E-03
	Weighted median	140	0.247	0.077	1.28 (1.10-1.49)	1.47E-03
	Inverse variance weighted	140	0.139	0.051	1.15 (1.04-1.27)	6.97E-03
	Simple mode	140	0.247	0.215	1.28 (0.84-1.95)	2.53E-01
	Weighted mode	140	0.391	0.146	1.47 (1.11-1.97)	8.29E-03
Myocardial infarction	MR-Egger	144	0.952	0.381	2.59 (1.22-5.47)	1.35E-02
	Weighted median	144	0.684	0.182	1.98 (1.41-2.78)	7.12E-05
	Inverse variance weighted	144	0.579	0.115	1.78 (1.42-2.23)	4.89E-07
	Simple mode	144	0.001	0.463	1.00 (0.40-2.48)	9.98E-01
	Weighted mode	144	0.718	0.331	2.05 (1.07-3.92)	3.16E-02
Total stroke	MR-Egger	140	0.171	0.214	1.87 (0.78-1.81)	4.24E-01
	Weighted median	140	0.107	0.098	1.11 (0.92-1.35)	2.75E-01
	Inverse variance weighted	140	0.145	0.065	1.15 (1.01-1.31)	2.71E-02
	Simple mode	140	0.070	0.212	1.07 (0.71-1.62)	7.42E-01
	Weighted mode	140	0.119	0.171	1.12 (0.81-1.57)	4.88E-01
Thrombotic stroke	MR-Egger	140	0.047	0.312	1.05 (0.57-1.93)	8.80E-01
	Weighted median	140	0.164	0.145	1.17 (0.88-1.56)	2.61E-01
	Inverse variance weighted	140	0.234	0.095	1.26 (1.05-1.52)	1.45E-02
	Simple mode	140	0.048	0.373	1.05 (0.50-2.18)	8.97E-01
	Weighted mode	140	-0.055	0.294	0.94 (0.53-1.68)	8.50E-01
Hemorrhagic stroke	MR-Egger	146	0.116	0.428	1.12 (0.48-2.60)	7.86E-01
	Weighted median	146	0.055	0.193	1.06 (0.72-1.54)	7.76E-01
	Inverse variance weighted	146	0.117	0.131	1.12 (0.86-1.45)	3.72E-01
	Simple mode	146	-0.144	0.462	0.86 (0.35-2.14)	7.56E-01
	Weighted mode	146	-0.054	0.353	0.94 (0.47-1.89)	8.76E-01

nSNP, the number of SNPs; *B*, beta; SE, standard error; and OR, odds ratio.

3. Genetic association with MASLD

GWAS was conducted to identify genetic variants associated with MASLD using KCPS-II individual-level data. Figure 5 shows Manhattan plot KCPS-II individual data: (A) Stratified according to FLI ≤ 30, (B) Stratified according to FLI ≤ 60, (C) MASLD defined based on FLI ≤ 30, and (D) MASLD defined based on FLI ≤ 60. The simple stratification analysis based on FLI values (Figures 5A & 5B) showed higher–$\log_{10}$ (p-value) peaks compared to the GWAS based on MASLD (Figures 5C & 5D), indicating stronger statistical significance.

Figure 6 shows a gene-based Manhattan plot using KCPS-II individual data: (A) stratified according to FLI ≤ 30, (B) stratified according to FLI ≤ 60, (C) MASLD defined based on FLI ≤ 30, and (D) MASLD defined based on FLI ≤ 60. Similar patterns were observed in this study. Specifically, on chromosome 12, RPH3A exhibited the highest peak, followed by GCKR, CUX2, APOC1, and GGT1 (Figures 6A and 6B), whereas FTO and CUX2 were the most prominent genes (Figures 6C and 6D).

According to the GWAS Catalog, several genes overlapped with previously reported loci: BMI (51 genes), hypertriglyceridemia (17 genes), metabolic syndrome (22 genes), nonalcoholic fatty liver disease (7 genes), platelet count (19 genes), lipid traits (11 genes), response to alcohol consumption (flushing response) (8 genes), myocardial infarction (15 genes), and coronary heart disease (10 genes) (Table 21).

The results of the enrichment analyses performed on all-canonical pathways and cell-type signature databases for genes associated with MASLD are shown in Figures 8 and 9. A total of 23 genes were also involved in 24 canonical processes according to canonical pathways, including statin inhibition of cholesterol production, chylomicron remodeling, plasma lipoprotein remodeling, familial hyperlipidemia, lipid particle composition, HDL remodeling, and chylomicron clearance. A total of 13 genes are involved in four cell types: hepatocyte clusters and bone marrow-derived precursor B cells. These cell types are functionally relevant to hepatic lipid metabolism, lipoprotein transport, and immune metabolic regulation. Genes such as APOA1, APOC3, and MLXIPL are consistently enriched across hepatocyte subsets, suggesting their potential roles in hepatic triglyceride processing and MASLD pathogenesis.[53]

For MASLD genes, the heat map illustrating the expression profiles according to genes and tissues (Figure 10) showed that the genes GCKR, LIPC, APOA5, APOA4, APOC3, and HNF1A were highly expressed in the liver tissues. In contrast, the RPH3A gene was expressed in the brain tissues (cerebellar hemisphere, cerebellum, cortex, and spine).

The joint expression of MASLD-related genes differed among tissues, as shown in Figure 11. The genes involved in MASLD showed an up-regulated differential expression of the genes (DEG) in the liver, heart atrial appendage, heart left ventricle, and renal cortex. However, the enrichment in these tissues was not statistically significant. Downregulated DEGs were observed in the bladder, followed by the brain amygdala, liver, hippocampus, whole blood, and adipose visceral omentum; however, the p-values for these tissues were

not statistically significant. Only the liver showed statistically significant enrichment in two-sided DEG analysis.

Figure 12 shows the gene network derived from MASLD-associated genes. A tightly connected cluster consisting of FTO, APOA5, APOA4, APOA1, APOC3, APOE, LIPC, GCKR, CUX, and MLXIPL indicates a functional module related to lipid metabolism. Although the node positions in the network layout are arbitrary, the connectivity pattern highlights functionally relevant groupings among MASLD-associated genes.

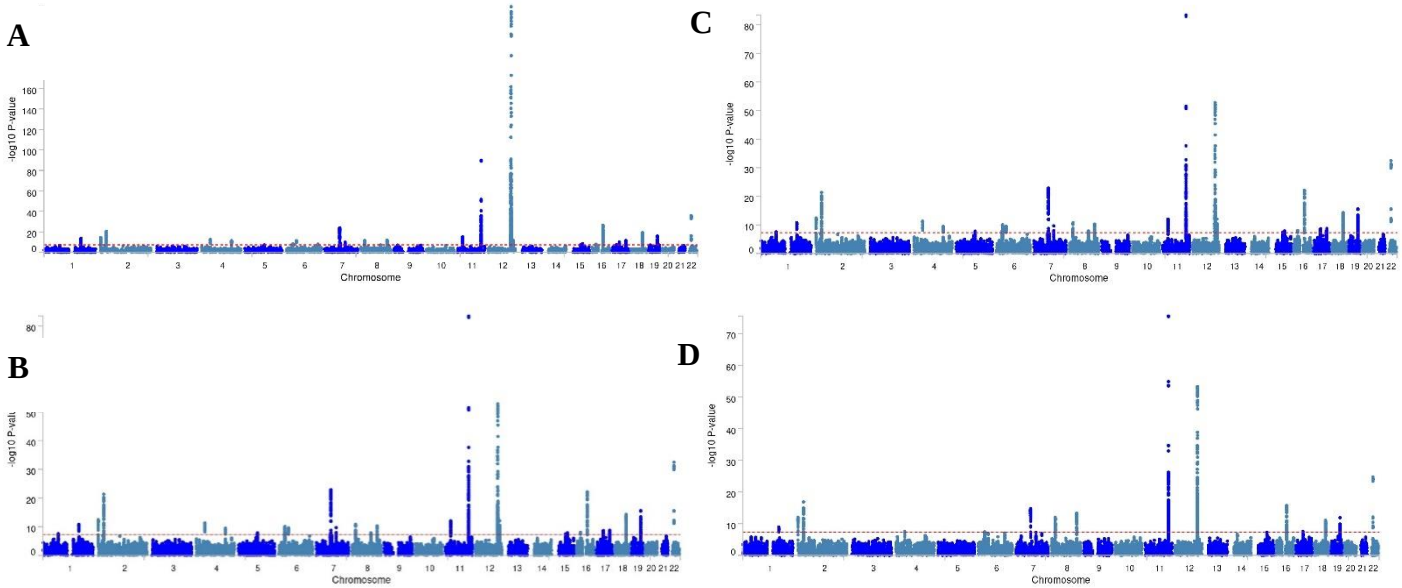


Figure 5. Manhattan plot KCPS-II individual data. (A) Stratified by FLI less than or equal to 30, (B) Stratified by FLI less than or equal to 60, (C) MASLD defined based on FLI less than or equal to 30, (D) MASLD defined based on FLI less than or equal to 60

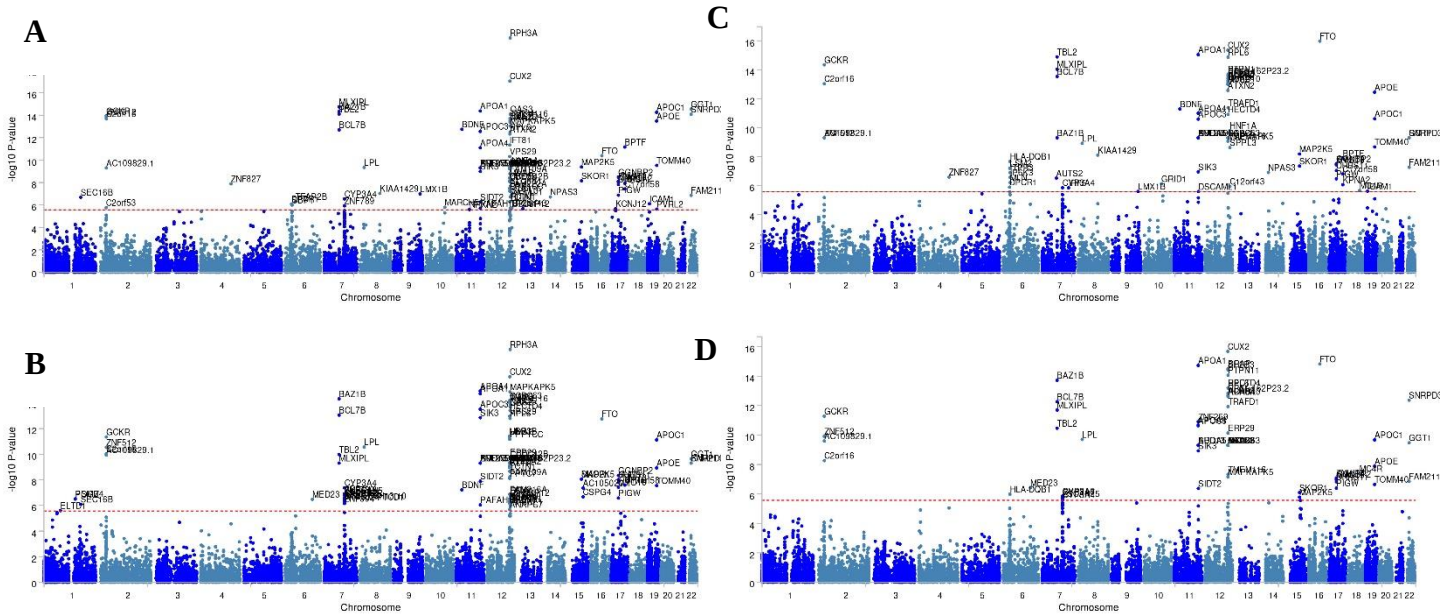


Figure 6. Gene based Manhattan plot using KCPS-II individual data. (A) Stratified by FLI less than or equal to 30, (B) Stratified by FLI less than or equal to 60, (C) MASLD defined based on FLI less than or equal to 30, (D) MASLD defined based on FLI less than or equal to 60

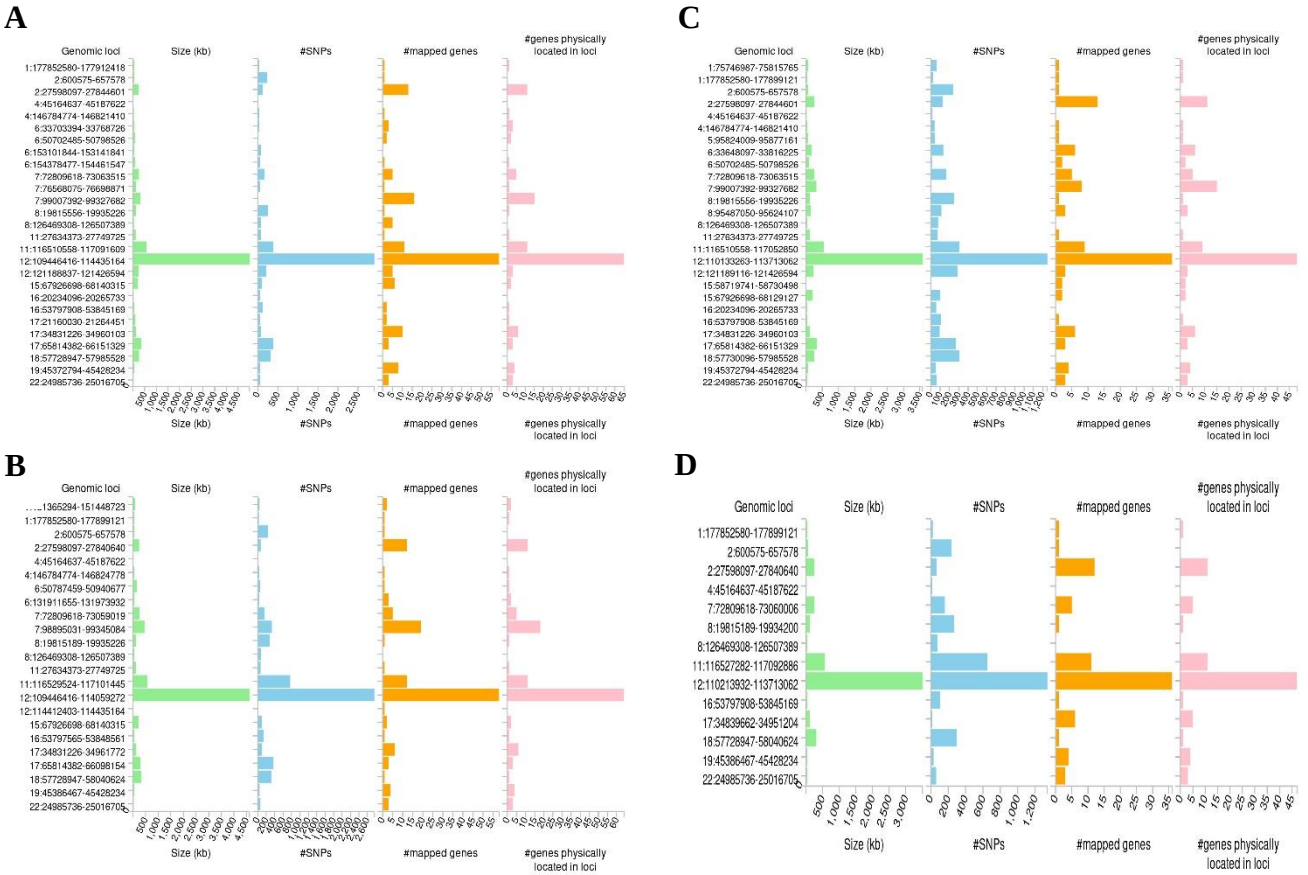


Figure 7. Genetic risk loci identified by FUMA analysis. (A) Stratified by FLI less than or equal to 30, (B) Stratified by FLI less than or equal to 60, (C) MASLD defined based on FLI less than or equal to 30, (D) MASLD defined based on FLI less than or equal to 60

Table 21. GWAS catalog reported genes compared with KCPS-II

Geneset	N	n	P-value	Adjusted P	Genes
Body mass index	1048	51	3.22e-35	1.43E-31	SEC16B, TMEM18, CAST, ITPR3, IP6K3, LEMD2, MLN, TFAP2D, TFAP2B, BDNF, BUD13, APOA5, APOA4, APOC3, APOA1, SIK3, MYL2, CUX2, SH2B3, ATXN2, BRAP, ACAD10, ALDH2, MAPKAPK5, TMEM116, ERP29, NAA25, TRAFD1, HECTD4, RPL6, PTPN11, RPH3A, OAS1, OAS3, OAS2, DTX1, RASAL1, DDX54, MAP2K5, FTO, ZNHIT3, MYO19, PIGW, GGNBP2, DHRS11, MRM1, BPTF, C17orf58, TOMM40, APOE, APOC1
Hypertriglyceridemia	37	17	3.15E-29	6.40E-26	PPM1G, NRBP1, KRTCAP3, IFT172, GCKR, C2orf16, ZNF512, BAZ1B, BCL7B, TBL2, MLXIPL, LPL, BUD13, APOA5, APOA4, APOC3, APOA1
Metabolic syndrome	102	22	4.34E-29	6.40E-26	SEC16B, TMEM18, GCKR, BAZ1B, MLXIPL, LPL, BDNF, BUD13, ZNF259, APOA5, SIK3, GLTP, GIT2, MYL2, ALDH2, HECTD4, LIPC, FTO, BPTF, TOMM40, APOE, APOC1
Lipid traits	21	11	4.51E-20	3.33E-17	GCKR, LPL, BUD13, APOA5, APOA4, APOC3, APOA1, LIPC, TOMM40, APOE, APOC1
Postprandial triglyceride levels	12	9	1.06E-18	5.86E-16	GCKR, MLXIPL, LPL, ZNF259, APOA5, APOA4, APOC3, APOA1, APOE
Response to alcohol consumption (flushing response)	9	8	8.21E-18	3.63E-15	CUX2, BRAP, ACAD10, ALDH2, NAA25, TRAFD1, HECTD4, PTPN11
Hematological and biochemical traits	33	10	2.13E-15	6.75eE-13	GCKR, LPL, APOA5, APOA4, APOC3, APOA1, BRAP, ALDH2, HECTD4, LIPC
Myocardial infarction	197	15	4.43E-13	9.33E-11	LPL, ZNF259, TCHP, CCDC63, SH2B3, ATXN2, BRAP, ACAD10, ALDH2, NAA25, HECTD4, OAS3, HNF1A, APOE, APOC1
Coronary heart disease	55	10	6.07E-13	1.22E-10	APOA5, APOA4, APOC3, APOA1, MYL2, CUX2, SH2B3, ACAD10, ALDH2, HNF1A
Nonalcoholic fatty liver disease	21	7	1.86E-11	2.65E-9	GCKR, C2orf16, ZNF512, GPN1, LPL, FTO, APOE
Coronary artery disease	482	18	2.79E-10	3.34E-8	LPL, BDNF, BUD13, ZNF259, APOA5, APOA4, APOA1, CUX2, SH2B3, ATXN2, ACAD10, ALDH2, NAA25, RPH3A, HNF1A, TOMM40, APOE, APOC1
Platelet count	551	19	3.23E-10	3.76E-8	GCKR, IP6K3, BAZ1B, MLXIPL, BUD13, APOA5, APOC3, SIK3, TCHP, SH2B3, ATXN2, BRAP, ACAD10, TMEM116, NAA25, TRAFD1, PTPN11, RPH3A, APOE

N, total number of genes included in the given geneset; n, Number of genes overlapping between the geneset and the list of input genes; Adjusted P, P-value corrected for multiple testing, typically using the Benjamini–Hochberg false discovery rate (FDR) method; Genes, list of overlapping genes

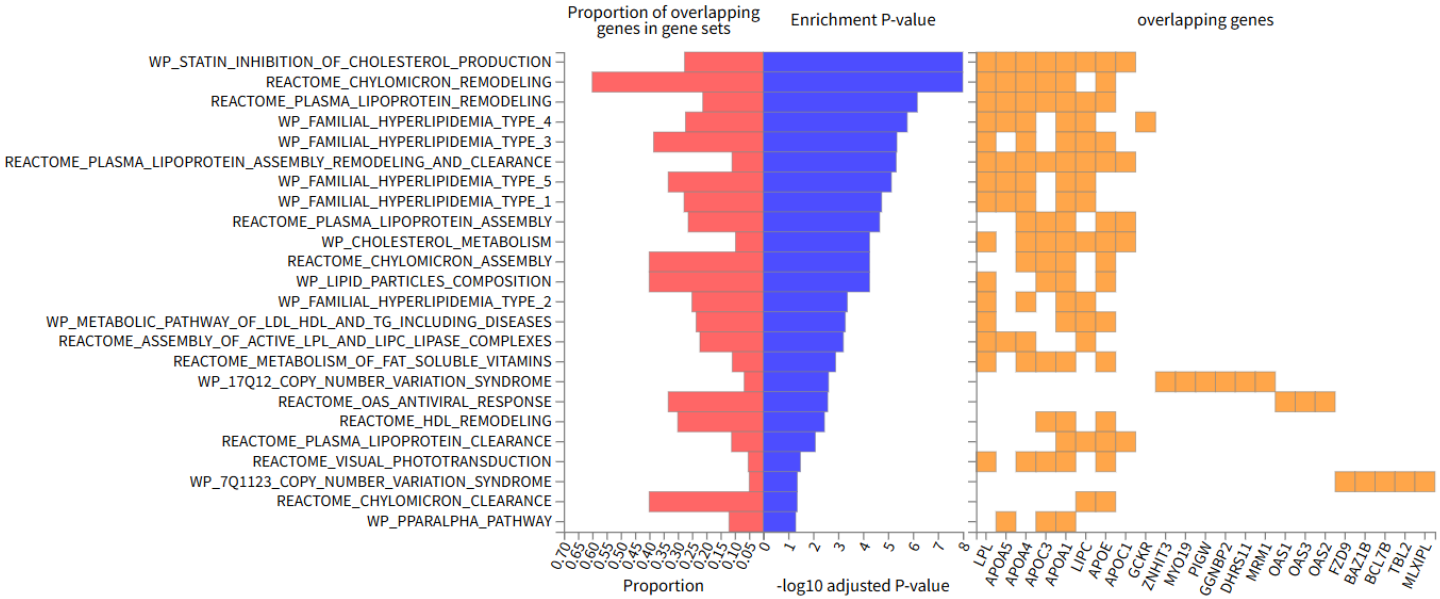


Figure 8. Enrichment analysis of a MASLD, regarding all canonical pathways processes (MASLD defined based on FLI ≥30)

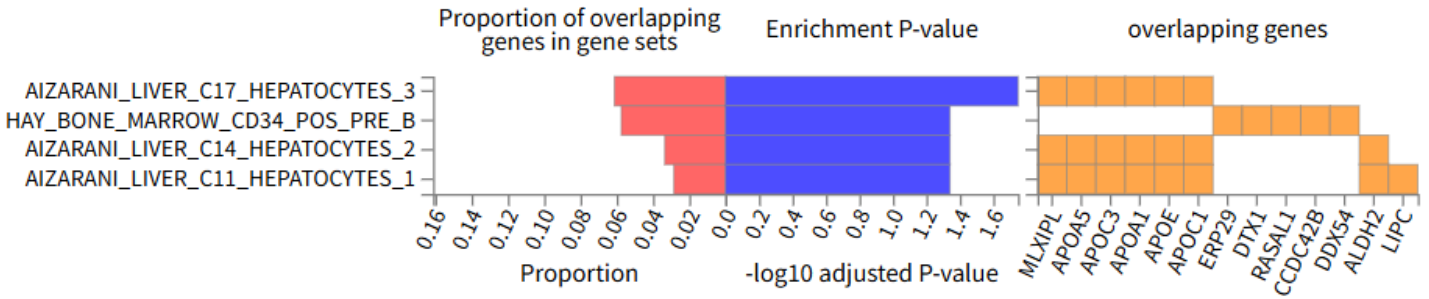


Figure 9. Enrichment analysis of a MASLD, regarding cell type signature (MASLD defined based on FLI ≥30)

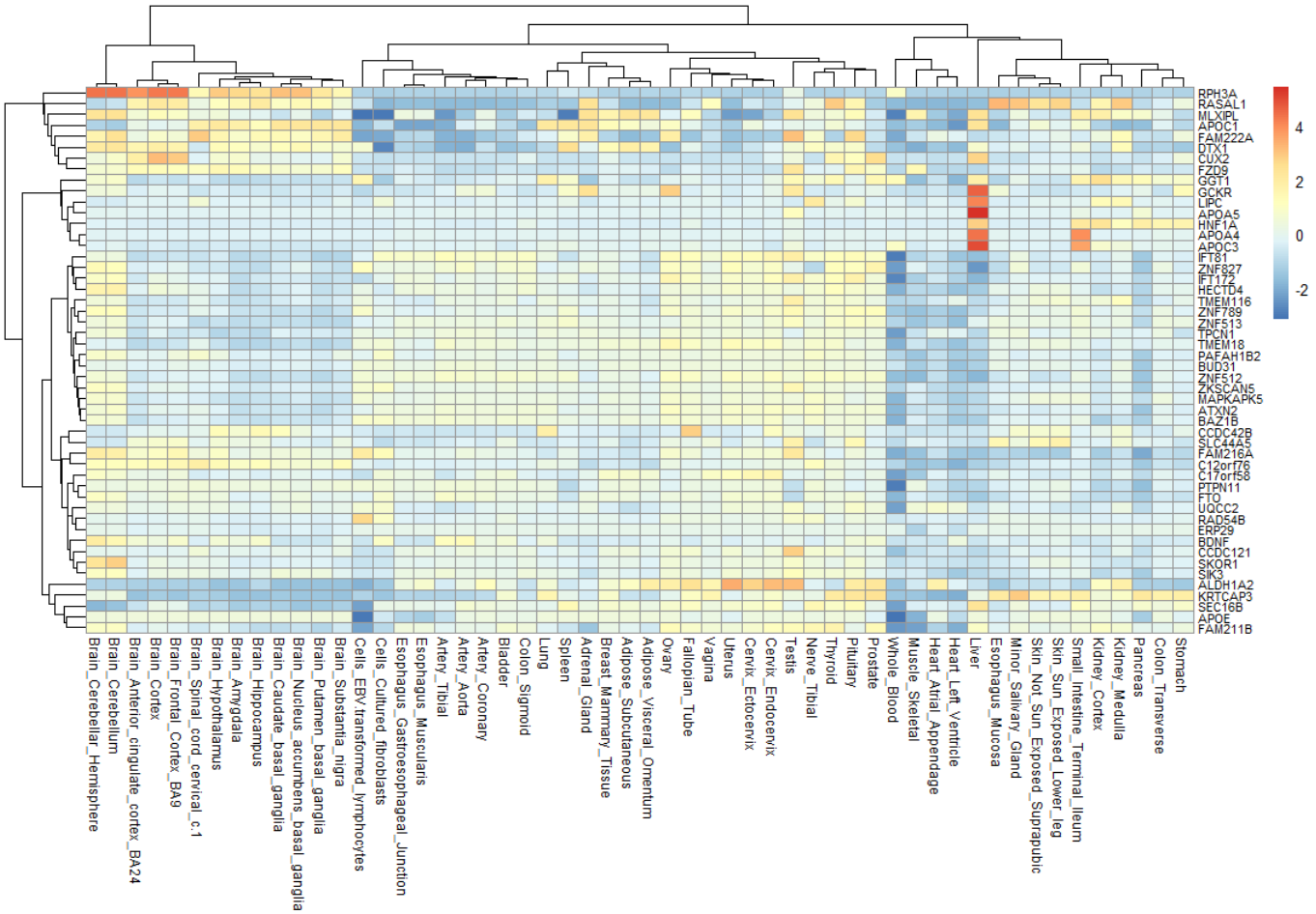


Figure 10. Heatmap of MASLD gene expression in different tissues (MASLD defined based on FLI ≥30)

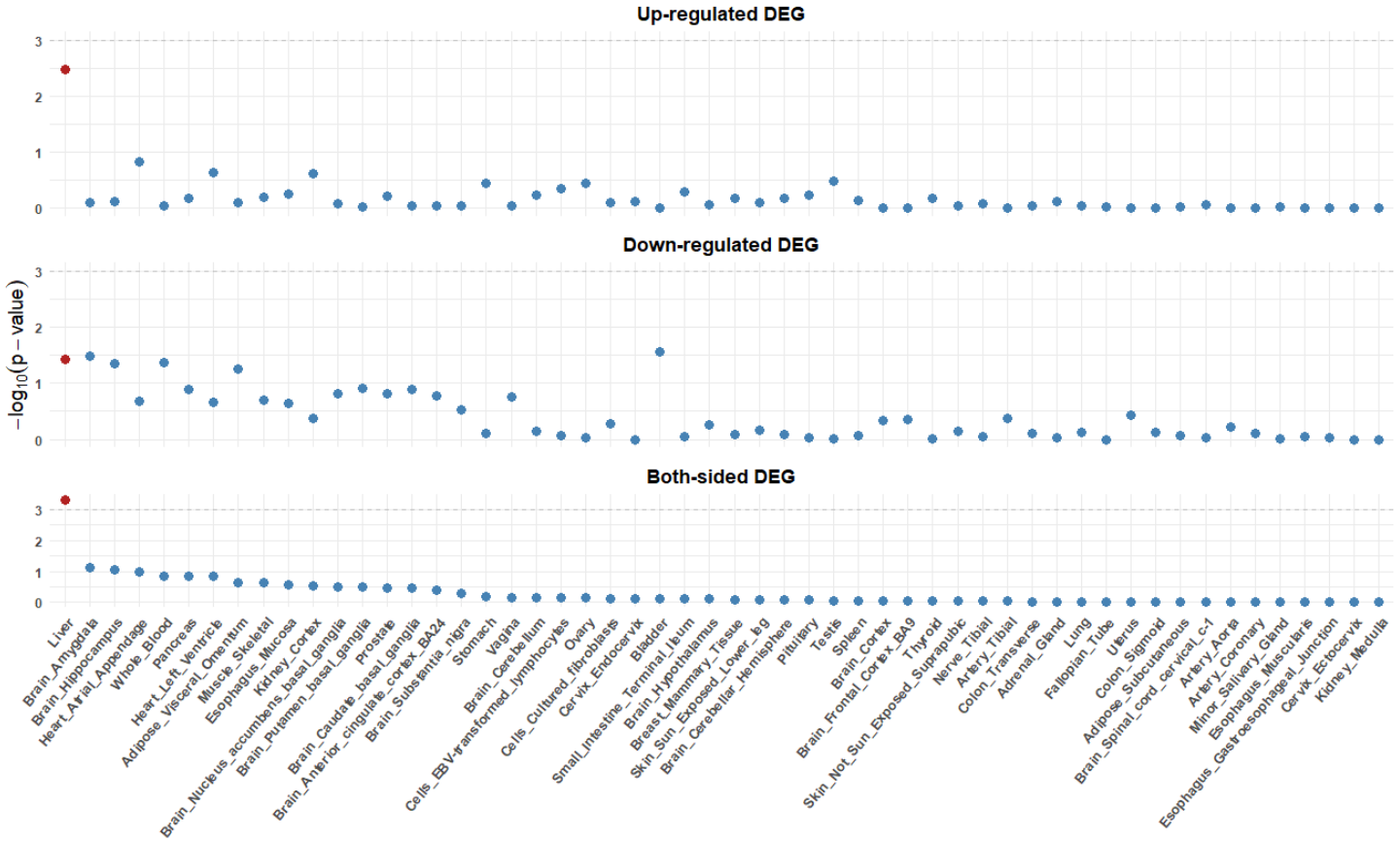


Figure 11. Tissue specificity expression of MASLD related genes in different tissue (MASLD defined based on FLI ≥30)

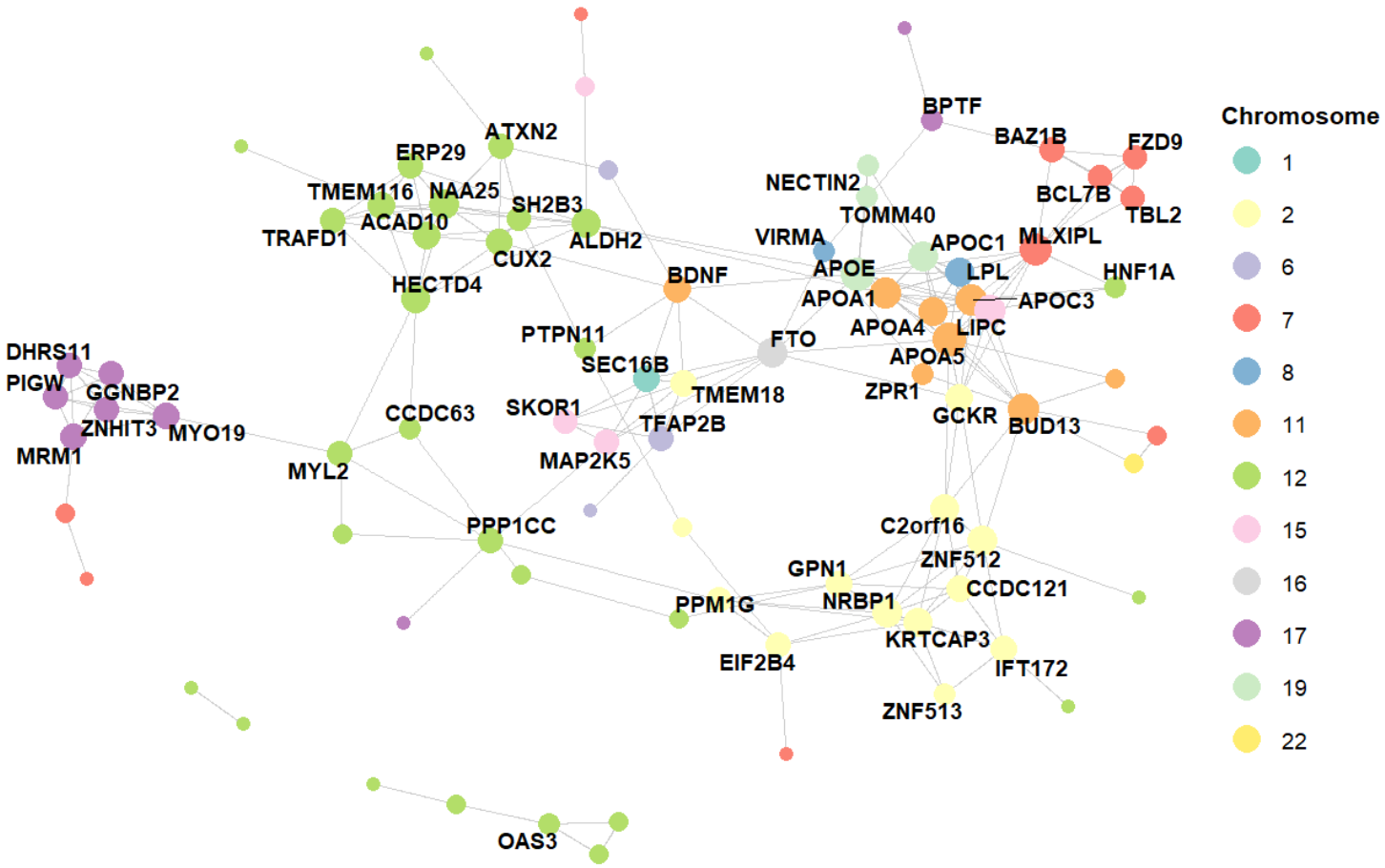


Figure 12. STRING-based protein–protein interaction network of MASLD-associated genes

4. Gene-environment interaction

Variance component analyses were conducted using GENIE to decompose the phenotypic variance of MASLD into genetic, GxE, and noise-environment interaction (NxE) components (Figure 13). For MASLD, the total heritability was estimated to be 38.6%, with genetic factors accounting for 6.5%, and GxE contributing an additional 5.3%. Notably, the NxE component explained 26.7% of the variance, reflecting the substantial role of environmental exposure independent of genetic background as well as random noise or unexplained variability.

GxE analysis was conducted for the 81 candidate SNPs. Based on nominal p-values ($p < 0.05$), several SNPs exhibited statistically significant interactions. Most of the essential variants showed positive GxE values, indicating increased genetic effects on MASLD in the presence of smoking (Table 22). Notably, rs671 (ALDH2) showed a strong genetic main effect (LRT = 55.528) and significant interaction with smoking (LRT = 18.756). The joint 2-degree-of-freedom test further confirmed this association (LRT = 74.284). Similar interaction patterns were observed for variants of PHACTR1 (rs11065933), UBE2L3, and PLCE1 (Table 22). In contrast, variants such as SLC9A1 and SLC6A2 showed negative β_GxE values, indicating attenuated genetic effects among smokers. These results suggest that smoking may modify genetic susceptibility to MASLD in a variant-specific manner (Figure 14).

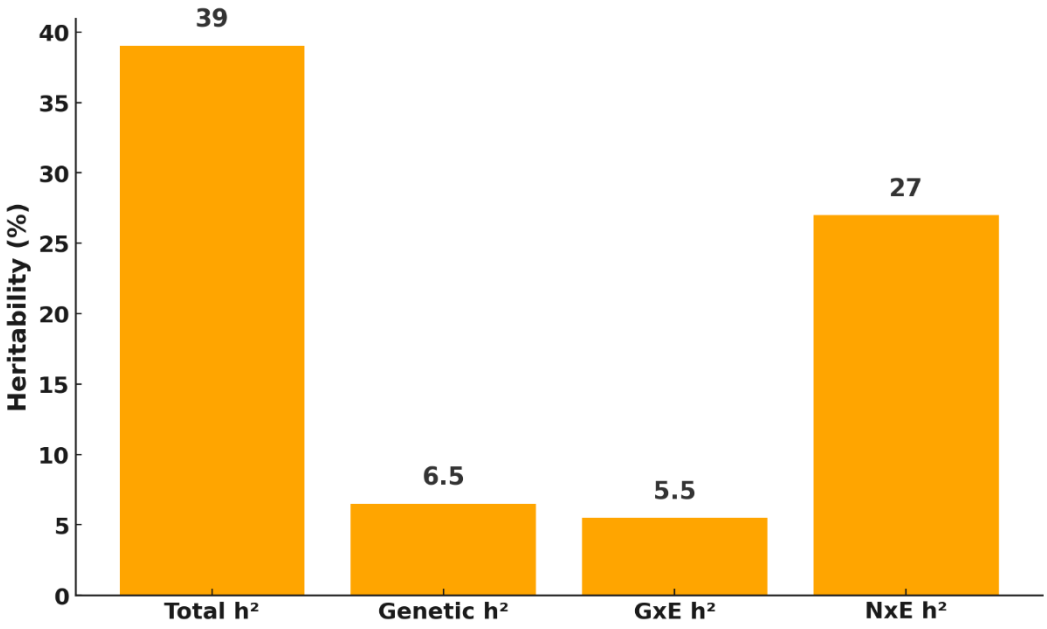


Figure 13. Contribution of genetic and smoking-environment interactions to MASLD heritability

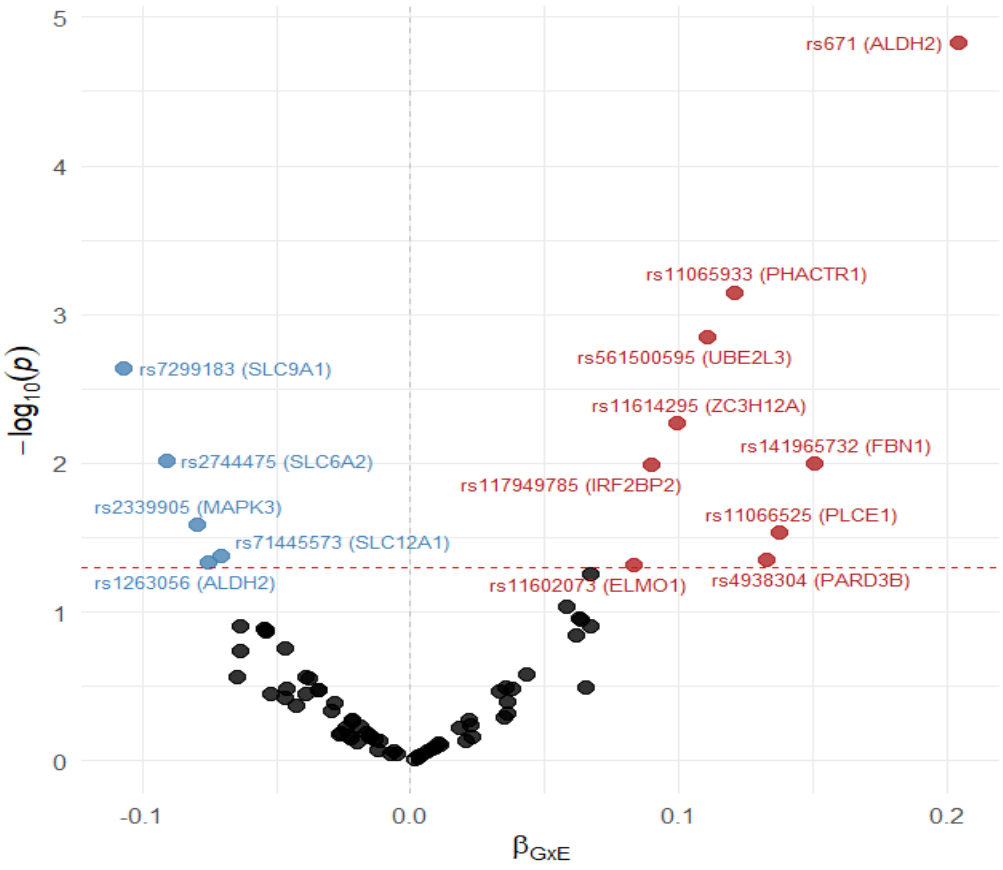


Figure 14. Gene–Smoking interaction effects on MASLD

Table 22. Genetic associations and interactions with smoking on MASLD risk

No	SNP	Gene	Betadg	Lrtdg	Betagxe	Lrtgxe	Lrt2df	Betacase	Lrtcase	Betactrl	Lrtctrl
1	rs671	ALDH2	-0.165	55.528	0.204	18.756	74.284	-0.023	11.173	-0.030	68.170
2	rs11065933	PHACTR1	-0.085	25.113	0.120	11.458	36.572	-0.011	4.751	-0.015	28.059
3	rs561500595	UBE2L3	-0.044	7.121	0.110	10.195	17.315	-0.007	1.712	-0.014	26.048
4	rs7299183	SLC9A1	0.048	8.368	-0.107	9.314	17.682	0.000	0.001	0.013	21.126
5	rs11614295	ZC3H12A	-0.069	16.401	0.099	7.752	24.153	-0.008	2.601	-0.011	14.658
6	rs2744475	SLC6A2	0.078	21.785	-0.091	6.722	28.507	-0.003	0.433	0.007	5.726
7	rs141965732	FBN1	-0.129	22.162	0.150	6.620	28.782	-0.012	2.191	-0.020	20.284
8	rs117949785	IRF2BP2	-0.059	12.359	0.090	6.614	18.973	-0.007	1.980	-0.009	10.580
9	rs2339905	MAPK3	0.076	19.491	-0.079	4.952	24.444	0.006	1.548	0.008	7.372
10	rs11066525	PLCE1	-0.107	12.768	0.137	4.758	17.526	-0.006	0.456	-0.009	3.310
11	rs71445573	SLC12A1	0.062	13.696	-0.071	4.134	17.830	0.010	3.861	0.010	13.751
12	rs4938304	PARD3B	-0.088	7.854	0.133	4.053	11.906	-0.001	0.013	-0.013	6.210
13	rs1263056	ALDH2	-0.081	20.480	-0.075	3.973	24.454	-0.013	5.319	0.000	0.000
14	rs11602073	ELMO1	-0.110	30.388	0.083	3.913	34.301	-0.004	0.353	-0.014	17.592

SNP, Single nucleotide polymorphism; Betadg, Effect size of the genetic main effect (G) on MASLD; Lrtdg, Likelihood ratio test statistic for genetic main effect; Betagxe, Effect size of the gene–smoking interaction (G×E) on MASLD; Lrtgxe, Likelihood ratio test statistic for the G×E interaction; lrt2df, Joint likelihood ratio test statistic for both G and G×E (2 degrees of freedom); betacase, Genetic effect estimate among smokers; lrtcase, LRT statistic for smokers only; betactrl, Genetic effect estimate among non-smokers; lrtctrl, LRT statistic for non-smokers only;

IV. DISCUSSION

4.1. Main findings

This study integrated observational and genetic epidemiological approaches to comprehensively evaluate the association and potential causality between MASLD and CVD. First, a survival analysis using a large Korean cohort revealed that individuals with MASLD had a significantly higher risk of developing CVD. Second, both the one-sample and two-sample MR analyses supported the potential causal effect of MASLD on CAD and IS, with consistent findings across datasets from the KCPS-II, UKB, and BBJ. Third, functional characterization of GWAS-identified genes using differential expression heatmaps, tissue-specific enrichment, and gene network analysis highlighted several MASLD-related genes, including APOA5, APOC3, LPL, HNF1A, GCKR, ALDH2, and RPH3A. These genes were found to be either functionally enriched or organized into distinct co-expression modules. Notably, lipid metabolism-related genes formed a tightly connected cluster, and the liver tissue exhibited statistically significant enrichment in two-sided differential expression analysis, underscoring its central role in MASLD pathophysiology. Fourth, GxE interaction analysis demonstrated that the genetic effects on MASLD varied with smoking exposure, with ALDH2 showing a notably stronger effect among smokers.

These findings provide robust evidence supporting the complex association between MASLD and CVD and emphasize the need to consider both genetic and environmental factors to understand metabolic liver disease and its cardiovascular consequences.

4.2. Previous observational and emerging evidence on systemic effect of MASLD

Substantial epidemiological evidence from extensive cohort studies indicates that MASLD is an independent risk factor for CVD morbidity and mortality. In a Swedish cohort study of over 10,000 individuals with biopsy-confirmed MASLD and matched controls, the presence of MASLD was linked to an elevated risk of CVD events over a median follow-up of 13.6 years, independent of conventional cardiometabolic risk factors. The adjusted HRs increased with liver disease severity: 1.67 (95% CI, 1.47–1.89) for non-cirrhotic fibrosis and 2.15 (95% CI, 1.77–2.61) for cirrhosis.[54]

Supporting these findings, a meta-analysis including 36 studies showed that MASLD—defined by imaging, diagnostic codes, or biopsy—was associated with a pooled HR of 1.45 (95% CI, 1.31–1.61) for fatal and nonfatal CVD outcomes. The risk was even higher among individuals with more advanced disease, especially those with MASH and higher fibrosis stages (random-effects HR 2.50, 95% CI 1.68–3.72).[55]

While the above studies relied primarily on histological or clinical definitions, similar associations were observed in large-scale population data. A study based on a nationwide health screening cohort of 8.8 million South Korean adults reported that MASLD, defined as an FLI ≥ 30, was independently associated with an increased risk of incident CVD events, including myocardial infarction, ischemic stroke, heart failure, and cardiovascular death, with an adjusted HR of 1.39 (95% CI 1.38–1.40).[18]

While most observational studies have focused on the association between MASLD and CVD outcomes via hepatic metabolic dysfunction, emerging evidence suggests a potential link between MASLD and neuropsychiatric conditions, such as cognitive decline, depression, and sleep disorders. The first studies evaluating brain health in patients with MASLD were conducted by Newton and colleagues,[56,57] and nearly half of the included patients with MASLD had fatigue and cognitive impairment. Additionally, MASLD is associated with an increased risk of depression and anxiety. Depression has been shown to affect executive function, learning, and memory.[58] MASLD is associated with obstructive sleep apnea. One study showed that obstructive sleep apnea in patients with MASLD was associated with reductions in the volumes of cortical and subcortical brain structures as well as cognitive impairment.[59]

These findings are consistent with the results of the present study, in which both MASLD and SLD were significantly associated with increased cardiovascular risk, including myocardial infarction and stroke, after adjusting for potential confounders. Neuropsychiatric outcomes, such as cognitive decline or depression, were not directly assessed in the present study; however, the identification of MASLD-associated genetic variants expressed in brain tissues suggests potential shared mechanisms contributing to brain health outcomes. These findings are consistent with a growing body of evidence indicating that MASLD may exert systemic effects beyond the liver and cardiovascular systems, warranting further investigation.

4.3. Genetic findings and functional interpretation

Genetic variants associated with increased liver fat content influence hepatic lipid metabolism through various mechanisms, including altered lipid droplet turnover, lipoprotein secretion, and de novo lipogenesis. The PNPLA3 allele, which increases liver fat, impairs lipid droplet remodeling and turnover, leading to triglyceride accumulation in hepatocytes. It is the best-characterized genetic variant to date and is most strongly associated with the development of steatohepatitis, advanced fibrosis, cirrhosis, hepatic decompensation, and liver-related mortality in patients with MASLD.[60-62] Also, the frequency and impact of this allele vary substantially across ethnic groups, with particularly high prevalence and stronger disease associations observed in Hispanic populations, moderate in East Asians and Europeans, and relatively low in individuals of African ancestry.[63] As a gene that encodes a major structural component of VLDL and chylomicrons, APOE may influence hepatic lipid metabolism; liver fat-promoting alleles at this locus can reduce VLDL secretion, thereby promoting triglyceride accumulation in the liver.[64] TM6SF2 contributes to hepatic lipid accumulation by reducing lipid export, thereby facilitating intracellular fat storage and predisposing the liver to conditions such as NASH and progressive fibrosis.[65] Increased glucokinase activity driven by the GCKR liver fat-associated allele promotes de novo lipogenesis, contributing to hepatic triglyceride buildup. [66] TRIB1 influences VLDL metabolism, lowering circulating TG levels.[67] In the present study, significant associations were observed between APOE and GCKR variants,

consistent with previous findings. However, other well-established variants, such as PNPLA3, TM6SF2, and TRIB1, did not reach statistical significance, possibly reflecting population-specific genetic architecture or differences in the study design.

In addition to liver-specific pathways, the present study also identified brain-expressing genes, such as RPH3A. RPH3A encodes a synaptic vesicle-associated protein that regulates exocytosis at presynaptic terminals and is critical for synaptic stabilization.[68,69] A recent study revealed that missense mutations in RPH3A cause an ultra-rare neurodevelopmental disorder, with varied expression levels associated with learning difficulties, intellectual disability, autism spectrum disorder, and epilepsy.[70] This finding may suggest a potential neurobiological component in the pathophysiology of MASLD.

This study identified evidence of GxE, particularly concerning smoking status. Notably, *rs671 (ALDH2)* exhibited a markedly amplified genetic effect on the risk of developing MASLD in smokers. Additionally, PHACTR1, UBE2L3, and PLCE1 exhibit potential GxE effects. These genes are involved in key biological pathways, including vascular endothelial function (PHACTR1)[71], immune and inflammatory regulation (UBE2L3)[72], and intracellular signaling and metabolic regulation (PLCE1)[73], all of which may contribute to the pathophysiology of MASLD. Conversely, SLC9A1 and SLC6A2 exhibited more substantial genetic effects in nonsmokers. SLC9A1 encodes a sodium/hydrogen exchanger that regulates intracellular pH and metabolic homeostasis, whereas SLC6A2 encodes a norepinephrine transporter involved in sympathetic nervous system activity.[74] These findings suggest that variants of these genes may modulate

MASLD risk through systemic or neuroregulatory pathways, especially in the absence of environmental stressors such as smoking.

In addition to GxE findings, this study employed a series of functional follow-up analyses, including tissue enrichment, differential expression heatmaps, and gene network analyses, to better understand the biological mechanisms underlying MASLD. Several genes, including APOA5, APOC3, LPL, LIPC, HNF1A, GCKR, and ALDH2 were consistently prioritized across multiple analyses, highlighting their potential relevance to lipid metabolism and systemic regulation. Notably, these genes were not only enriched in liver-specific pathways but also exhibited coordinated expression patterns, forming tightly connected clusters in the network analysis. The one-sample MR analysis using a less stringent LD clumping threshold (e.g., $p < 5 \times 10^{-5}$) included a greater number of SNPs and yielded statistically significant associations with CVD outcomes. Although relaxing the threshold increases the risk of including weak instruments, it may also capture biologically important variants that do not reach conventional significance levels. This underscores the need for integrative multiomic approaches to validate and functionally annotate these variants.

Lastly, although rs671 (ALDH2) was excluded as an outlier in the MR analysis, it demonstrated a notable gene-environment interaction effect with smoking in the GxE analysis and remained consistently prioritized in other studies. These findings suggest that variants, such as rs671, may have context-dependent biological roles, further supporting the importance of evaluating genetic effects in diverse analytical frameworks.

Evidence from MR studies on the causal role of MASLD in CVD is controversial. Genetically defined MASLD is associated with a higher risk of arterial stiffness and heart failure but not with coronary artery disease, stroke, ischemic stroke, or its subtypes.[75] MR studies of the association between MASLD and cancers are scarce but suggest no association between genetically predicted MASLD and the risk of intrahepatic cholangiocarcinoma.[76] Some studies aiming to assess the causal impact of MASLD have relied on circulating levels of liver enzymes as proxies for the condition. However, liver enzymes, such as alanine aminotransferase, are imperfect predictors of MASLD.[77] This study defined MASLD using the FLI and found evidence supporting a potential causal association with CVD.

4.4. Risk factors and mechanisms of MASLD

Obesity, insulin resistance, T2DM, hypertension, and dyslipidemia, including elevated triglyceride levels and reduced HDL-C, are widely recognized as key metabolic risk factors for MASLD.[78] A meta-analysis indicated that increased waist circumference (pooled OR 2.34, 95% CI: 1.83–3.00) and higher body mass index (pooled OR 2.85, 95% CI, 1.60–5.08) are significantly associated with a greater likelihood of MASLD.[79] Globally, MASLD is estimated to affect 55.5% (95% CI, 47.3–63.7) of individuals with T2DM, while approximately 37.3% (95% CI, 24.7–50.0) of these individuals are affected by MASH.[80] In addition, a meta-analysis of 11 studies reported that hypertension is associated with a significantly increased risk of incident MASLD (HR 1.63, 95% CI, 1.41–1.88).[81] Dyslipidemia is also commonly observed in individuals with MASLD, and approximately 50% of patients with dyslipidemia attending lipid clinics were found to have MASLD.[82] Although MASLD predominantly affects obese adults, cases among lean individuals are increasingly being recognized, often referred to as "lean liver disease".[83] This trend likely reflects the complex and heterogeneous pathophysiology of MASLD. These observations suggest that a combination of various exogenous factors and endogenous conditions may contribute to liver injury.[84] For instance, hepatic steatosis is promoted by a combination of external stressors, such as diet, lifestyle, medication use, and hepatotoxic exposures, along with internal factors, including genetic predisposition, insulin resistance, increased fat synthesis, toxic lipid accumulation, and gut microbiome imbalance.[85] Overnutrition is

widely recognized as the predominant external factor contributing to MASLD, leading to adipose tissue expansion and ectopic fat accumulation. This process alters tissue metabolism and induces the dysregulation of hepatocytes, such as insulin resistance.[86] In addition, hepatocellular stress, metabolic disturbances, and microbial translocation form a compromised gut barrier via the portal vein, thereby promoting immune dysregulation. This creates a proinflammatory environment in the liver and systemically accelerates disease progression.[87] Therefore, MASLD is closely linked to metabolic conditions, such as T2DM and obesity, with overlapping pathophysiological mechanisms. In addition to lipid accumulation and lipotoxicity, its progression involves a complex network of metabolic processes, gut microbiome, and immune responses. Multiple interacting factors, including diet, microbial components, and inflammatory mediators, cause this disease. Alterations in the gut microbiome and adipose tissue inflammation play key roles in creating a vicious cycle of metabolic dysregulation and immune activation, which drives disease progression. Various therapeutic options that affect metabolism, lipotoxicity, inflammation, appetite, and the gut microbiome, all of which are involved in disease pathophysiology, are currently being tested for the treatment of MASLD and MASH.[88]

Therefore, treatment strategies should consider the multifactorial nature of MASLD. Although individual therapeutic targets, such as lipid metabolism, inflammation, and insulin resistance, are important, a comprehensive approach may offer greater benefits. Pharmacological agents such as GLP-1 receptor agonists, which simultaneously address weight reduction, appetite regulation, and systemic inflammation, represent a promising

therapeutic option.[89] Combining lifestyle interventions with such agents may pave the way for effective and sustainable management of metabolic liver disease in the future.

4.5. Strengths and limitations

This study has several strengths. First, a large sample size of more than 111,637 participants ensured adequate statistical power. A comprehensive range of cardiovascular and cerebrovascular outcomes, including overall CVD, IHD, MI, total stroke, thrombotic stroke, and hemorrhagic stroke, were evaluated. To enhance the accuracy of the estimates, the statistical models were sequentially adjusted for important confounders, including age, sex, smoking, alcohol use, physical activity, insurance status, and eGFR. A series of sensitivity analyses was conducted to confirm the robustness and consistency of the results. To our knowledge, this is the first study to report an association between the newly defined MASLD and the risk of CVD according to the FIB-4 score. Genetic annotation was performed to explore the potential biological mechanisms underlying genetic variants associated with MASLD, further reinforcing the biological plausibility of the observed associations. The use of one- and two-sample MR approaches strengthened the robustness of our findings regarding the causal association between MASLD score and CVD outcomes. One of the significant strengths of this study is an integrated analysis of both observational and genetic data, allowing for a more comprehensive interpretation and supportive evidence toward causality between MASLD and cardiovascular outcomes.

However, this study has several unresolved issues. First, because general characteristics were obtained through a self-reported questionnaire, recall bias may have occurred. The relatively young age of the study population may have led to differences in the multivariate

analysis outcomes compared to those observed in other studies. With additional follow-ups and an expanded age range, findings that are more consistent with those of previous studies will likely emerge. Because the data were collected from the general Korean population during health checkups, the findings may not be generalizable to other populations or ethnic groups. Also, because this study did not include information on dietary habits, further research incorporating these variables is recommended. Histological findings and imaging-based diagnostic data were not analyzed; instead, the FLI was used as an adjustment variable. Although the FLI is a non-invasive marker, it has been validated as a useful surrogate in various epidemiological studies. However, reliance on FLI, rather than direct imaging or biopsy-based diagnosis, may limit the precision of phenotypic definitions in genetic analyses. While the use of the FLI does not significantly undermine the reliability of the findings, it remains a potential limitation for both the observational and genetic components of this study. The use of both observational and genetic approaches, including MR, enhanced the robustness of the findings by addressing potential confounding factors. As in all nonrandomized studies, residual confounding due to unknown or unmeasured factors cannot be entirely ruled out. The exposure and outcome summary statistics used in the MR analysis were derived from different ancestral populations (East Asian and European). This trans-ancestry design may introduce bias due to differences in the linkage disequilibrium structure, allele frequencies, and SNP effect sizes, potentially affecting the accuracy and generalizability of the estimated causal effects. Therefore, the results should be interpreted with caution, and replication in ancestry-matched populations is warranted.

Ⅴ. Conclusion

In this large-scale prospective cohort study of approximately 111,637 Koreans, the MASLD was significantly associated with an increased risk of CVD. Individuals with MASLD showed a significantly increased risk of developing CVD compared with metabolically healthy controls; this association remained consistent after adjustment for various covariates. The one-sample MR analysis demonstrated a modest but significant association between genetically predicted MASLD and CVD risk. In contrast, the two-sample MR using BBJ summary statistics showed consistent CAD findings. Transcriptomic and functional enrichment analyses indicated that MASLD-related genes were involved in lipid metabolism, metabolic syndrome, and coronary artery disease pathways. G×E analysis further revealed that genetic effects on MASLD may vary depending on environmental exposure, with stronger associations observed among smokers for variants such as rs671 (ALDH2).

The study findings suggest a potential causal role for MASLD in the development of CVD, emphasizing the importance of metabolic liver health in CVD prevention. The results also highlight the potential value of integrating genetic and metabolic liver profiles into personalized strategies for predicting and preventing CVD. Overall, this study supports the role of metabolic liver dysfunction as an upstream determinant of cardiovascular risk.

References

1. Chan WK, Chuah KH, Rajaram RB, Lim LL, Ratnasingam J, Vethakkan SR. Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD): A State-of-the-Art Review. *J Obes Metab Syndr*. Sep 30 2023;32(3):197-213. doi:10.7570/jomes23052

2. Miao L, Targher G, Byrne CD, Cao Y-Y, Zheng M-H. Current status and future trends of the global burden of MASLD. *Trends in Endocrinology & Metabolism*. 2024/08/01/ 2024;35(8):697-707. doi:https://doi.org/10.1016/j.tem.2024.02.007

3. Paik JM, Henry L, Younossi Y, Ong J, Alqahtani S, Younossi ZM. The burden of nonalcoholic fatty liver disease (NAFLD) is rapidly growing in every region of the world from 1990 to 2019. *Hepatology communications*. 2023;7(10):e0251.

4. Lekakis V, Papatheodoridis GV. Natural history of metabolic dysfunction-associated steatotic liver disease. *European Journal of Internal Medicine*. 2024;122:3-10.

5. Eslam M, Sanyal AJ, George J, et al. MAFLD: a consensus-driven proposed nomenclature for metabolic associated fatty liver disease. *Gastroenterology*. 2020;158(7):1999-2014. e1.

6. Rinella ME, Lazarus JV, Ratziu V, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology*. 2023;78(6):1966-1986. doi:10.1097/hep.0000000000000520

7. Rinella ME, Lazarus JV, Ratziu V, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology*. 2023;78(6):1966-1986.

8. Adams LA, Angulo P, Lindor KD. Nonalcoholic fatty liver disease. *Cmaj*. 2005;172(7):899-905.

9. Idilman IS, Aniktar H, Idilman R, et al. Hepatic steatosis: quantification by proton density fat fraction with MR imaging versus liver biopsy. *Radiology*. 2013;267(3):767-775.

10. Park HJ, Lee S, Lee JS. Differences in the prevalence of NAFLD, MAFLD, and MASLD according to changes in the nomenclature in a health check-up using MRI-derived proton density fat fraction. *Abdominal Radiology*. 2024:1-9.

11. Zelber-Sagi S, Webb M, Assy N, et al. Comparison of fatty liver index with noninvasive methods for steatosis detection and quantification. *World J Gastroenterol*. Jan 7 2013;19(1):57-64. doi:10.3748/wjg.v19.i1.57
12. Zhou Y-J, Zhou Y-F, Zheng J-N, et al. NAFL screening score: A basic score identifying ultrasound-diagnosed non-alcoholic fatty liver. *Clinica Chimica Acta*. 2017/12/01/ 2017;475:44-50. doi:https://doi.org/10.1016/j.cca.2017.09.020
13. Eslam M, Sarin SK, Wong VW-S, et al. The Asian Pacific Association for the Study of the Liver clinical practice guidelines for the diagnosis and management of metabolic associated fatty liver disease. *Hepatology international*. 2020;14:889-919.
14. Bedogni G, Bellentani S, Miglioli L, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC gastroenterology*. 2006;6:1-7.
15. Nagai A, Hirata M, Kamatani Y, et al. Overview of the BioBank Japan Project: Study design and profile. *Journal of epidemiology*. 2017;27(Supplement_III):S2-S8.
16. Golabi P, Paik JM, AlQahtani S, Younossi Y, Tuncer G, Younossi ZM. Burden of non-alcoholic fatty liver disease in Asia, the Middle East and North Africa: Data from Global Burden of Disease 2009-2019. *Journal of hepatology*. 2021;75(4):795-809.
17. Rosato V, Masarone M, Dallio M, Federico A, Aglitti A, Persico M. NAFLD and extra-hepatic comorbidities: current evidence on a multi-organ metabolic syndrome. *International journal of environmental research and public health*. 2019;16(18):3415.
18. Lee H-H, Lee HA, Kim E-J, et al. Metabolic dysfunction-associated steatotic liver disease and risk of cardiovascular disease. *Gut*. 2024;73(3):533-540.
19. Moon JH, Jeong S, Jang H, Koo BK, Kim W. Metabolic dysfunction-associated steatotic liver disease increases the risk of incident cardiovascular disease: a nationwide cohort study. *EClinicalMedicine*. 2023;65
20. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. *American journal of preventive medicine*. 2007;33(2):155-161.
21. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*. 2019;20(8):467-484.

22. Sookoian S, Rotman Y, Valenti L. Genetics of Metabolic Dysfunction-associated Steatotic Liver Disease: The State of the Art Update. *Clinical Gastroenterology and Hepatology*. 2024/11/01/ 2024;22(11):2177-2187.e3. doi:https://doi.org/10.1016/j.cgh.2024.05.052

23. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nature communications*. 2017;8(1):1826.

24. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome medicine*. 2020;12(1):44.

25. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nature protocols*. 2020;15(9):2759-2772.

26. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *Jama*. 2017;318(19):1925-1926.

27. Targher G, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: a meta-analysis. *Journal of hepatology*. 2016;65(3):589-600.

28. Wong VWS, Wong GLH, Yeung JCL, et al. Long-term clinical outcomes after fatty liver screening in patients undergoing coronary angiogram: a prospective cohort study. *Hepatology*. 2016;63(3):754-763.

29. Manuck SB, McCaffery JM. Gene-environment interaction. *Annual review of psychology*. 2014;65(1):41-70.

30. Dunn AR, O'Connell KM, Kaczorowski CC. Gene-by-environment interactions in Alzheimer's disease and Parkinson's disease. *Neuroscience & Biobehavioral Reviews*. 2019;103:73-80.

31. Costenbader K, Karlson E. Cigarette smoking and autoimmune disease: what can we learn from epidemiology? *Lupus*. 2006;15(11):737-745.

32. Jee YH, Emberson J, Jung KJ, et al. Cohort Profile: The Korean Cancer Prevention Study-II (KCPS-II) Biobank. *International Journal of Epidemiology*. 2017;47(2):385-386f. doi:10.1093/ije/dyx226

33. Eslam M, Newsome PN, Sarin SK, et al. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. *Journal of hepatology*. 2020;73(1):202-209.
34. European Association for the Study of the Liver EAftSoD, European Association for the Study of Obesity. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *Obesity facts*. 2016;9(2):65-90.
35. Lee Y-h, Bang H, Park YM, et al. Non–laboratory-based self-assessment screening score for non-alcoholic fatty liver disease: development, validation and comparison with other scores. *PloS one*. 2014;9(9):e107584.
36. Guerrero-Romero F, Simental-Mendía LE, Gonzalez-Ortiz M, et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. *The Journal of Clinical Endocrinology & Metabolism*. 2010;95(7):3347-3351.
37. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clinical gastroenterology and hepatology*. 2009;7(10):1104-1112.
38. Yi SW, Mok Y, Ohrr H, et al. Low Systolic Blood Pressure and Vascular Mortality Among More Than 1 Million Korean Adults. *Circulation*. Jun 14 2016;133(24):2381-90. doi:10.1161/circulationaha.115.020752
39. Koyama S, Ito K, Terao C, et al. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. *Nature genetics*. 2020;52(11):1169-1177.
40. Nakamura Y. The BioBank Japan Project. *Clin Adv Hematol Oncol*. 2007;5:696-697.
41. Allen N, Sudlow C, Downey P, et al. UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology*. 2012;1(3):123-126.
42. Jung KJ, Baek JW, Shin SY, Jee SH. Genetically determined alcohol consumption and cancer risk in Korea. *Epidemiology and Health*. 2023;45:e2023077.
43. Targher G, Byrne CD, Tilg H. MASLD: a systemic metabolic disorder with cardiovascular and malignant complications. *Gut*. 2024;73(4):691-702.

44. Yanai H, Adachi H, Hakoshima M, Iida S, Katsuyama H. Metabolic-Dysfunction-Associated Steatotic Liver Disease—Its Pathophysiology, Association with Atherosclerosis and Cardiovascular Disease, and Treatments. *International journal of molecular sciences*. 2023;24(20):15473.
45. Che R, Motsinger-Reif AA. Evaluation of genetic risk score models in the presence of interaction and linkage disequilibrium. *Frontiers in genetics*. 2013;4:138.
46. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human molecular genetics*. 2014;23(R1):R89-R98.
47. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology*. 2015;44(2):512-525.
48. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic epidemiology*. 2016;40(4):304-314.
49. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology*. 2017;46(6):1985-1998.
50. Bowden J, Spiller W, Del Greco M F, et al. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *International journal of epidemiology*. 2018;47(4):1264-1278.
51. Pazokitoroudi A, Liu Z, Dahl A, Zaitlen N, Rosset S, Sankararaman S. A scalable and robust variance components method reveals insights into the architecture of gene-environment interactions underlying complex traits. *Am J Hum Genet*. Jul 11 2024;111(7):1462-1480. doi:10.1016/j.ajhg.2024.05.015
52. Murcray CE, Lewinger JP, Gauderman WJ. Gene-environment interaction in genome-wide association studies. *American journal of epidemiology*. 2009;169(2):219-226.

53. Tall AR, Thomas DG, Gonzalez-Cabodevilla AG, Goldberg IJ. Addressing dyslipidemic risk beyond LDL-cholesterol. *The Journal of clinical investigation*. 2022;132(1)

54. Simon TG, Roelstraete B, Hagström H, Sundström J, Ludvigsson JF. Non-alcoholic fatty liver disease and incident major adverse cardiovascular events: results from a nationwide histology cohort. *Gut*. 2022;71(9):1867-1875.

55. Mantovani A, Csermely A, Petracca G, et al. Non-alcoholic fatty liver disease and risk of fatal and non-fatal cardiovascular events: an updated systematic review and meta-analysis. *The lancet Gastroenterology & hepatology*. 2021;6(11):903-913.

56. Newton JL, Jones DE, Henderson E, et al. Fatigue in non-alcoholic fatty liver disease (NAFLD) is significant and associates with inactivity and excessive daytime sleepiness but not with liver disease severity or insulin resistance. *Gut*. 2008;57(6):807-813.

57. Newton JL. Systemic symptoms in non-alcoholic fatty liver disease. *Digestive diseases*. 2010;28(1):214-219.

58. Pan Z, Park C, Brietzke E, et al. Cognitive impairment in major depressive disorder. *CNS spectrums*. 2019;24(1):22-29.

59. Filipovic B, Đuric V, Filipovic N, et al. Anatomical brain changes and cognitive abilities in patients with obstructive sleep apnea syndrome and nonalcoholic fatty liver disease. *Canadian Journal of Gastroenterology and Hepatology*. 2021;2021(1):8873652.

60. Chen L-Z, Xin Y-N, Geng N, Jiang M, Zhang D-D, Xuan S-Y. PNPLA3 I148M variant in nonalcoholic fatty liver disease: demographic and ethnic characteristics and the role of the variant in nonalcoholic fatty liver fibrosis. *World Journal of Gastroenterology: WJG*. 2015;21(3):794.

61. Petta S, Armandi A, Bugianesi E. Impact of PNPLA3 I148M on clinical outcomes in patients with MASLD. *Liver International*. 2025;45(3):e16133.

62. Seko Y, Lin H, Wong VWS, Okanoue T. Impact of PNPLA3 in lean individuals and in cryptogenic steatotic liver disease. *Liver International*. 2025;45(4):e16164.

63. Kozlitina J, Sookoian S. Global epidemiological impact of PNPLA3 I148M on liver disease. *Liver International*. 2025;45(3):e16123.

64. Zhao N, Liu C-C, Van Ingelgom AJ, et al. Apolipoprotein E4 impairs neuronal insulin signaling by trapping insulin receptor in the endosomes. *Neuron*. 2017;96(1):115-129. e5.

65. Zhang X, Liu S, Dong Q, Xin Y, Xuan S. The genetics of clinical liver diseases: insight into the TM6SF2 E167K variant. *Journal of Clinical and Translational Hepatology*. 2018;6(3):326.

66. Kitamoto A, Kitamoto T, Nakamura T, et al. Association of polymorphisms in GCKR and TRIB1 with nonalcoholic fatty liver disease and metabolic syndrome traits. *Endocrine journal*. 2014;61(7):683-689.

67. Iwamoto S, Boonvisut S, Makishima S, Ishizuka Y, Watanabe K, Nakayama K. The role of TRIB1 in lipid metabolism; from genetics to pathways. *Biochemical Society transactions*. 2015;43(5):1063-1068.

68. Franchini L, Stanic J, Ponzoni L, et al. Linking NMDA receptor synaptic retention to synaptic plasticity and cognition. *Iscience*. 2019;19:927-939.

69. Stanic J, Mellone M, Napolitano F, et al. Rabphilin 3A: A novel target for the treatment of levodopa-induced dyskinesias. *Neurobiology of Disease*. 2017;108:54-64.

70. Pavinato L, Stanic J, Barzasi M, et al. Missense variants in RPH3A cause defects in excitatory synaptic function and are associated with a clinically variable neurodevelopmental disorder. *Genetics in Medicine*. 2023;25(11):100922.

71. Debette S, Kamatani Y, Metso TM, et al. Common variation in PHACTR1 is associated with susceptibility to cervical artery dissection. *Nature genetics*. 2015;47(1):78-83.

72. Zhang X, Huo C, Liu Y, Su R, Zhao Y, Li Y. Mechanism and disease association with a ubiquitin conjugating E2 enzyme: UBE2L3. *Frontiers in immunology*. 2022;13:793610.

73. Hinkes B, Wiggins RC, Gbadegesin R, et al. Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. *Nature genetics*. 2006;38(12):1397-1405.

74. Pizzagalli MD, Bensimon A, Superti-Furga G. A guide to plasma membrane solute carrier proteins. *The FEBS journal*. 2021;288(9):2784-2835.

75. Ahmed A, Cule M, Bell JD, Sattar N, Yaghootkar H. Differing genetic variants associated with liver fat and their contrasting relationships with cardiovascular diseases and cancer. *Journal of Hepatology*. 2024;81(6):921-929. doi:10.1016/j.jhep.2024.06.030

76. Qin S-s, Pan G-q, Meng Q-b, Liu J-b, Tian Z-y, Luan S-j. The causal relationship between metabolic factors, drinking, smoking and intrahepatic cholangiocarcinoma: a Mendelian randomization study. *Frontiers in oncology*. 2023;13:1203685.

77. Ren Z, Simons PI, Wesselius A, Stehouwer CD, Brouwers MC. Relationship between NAFLD and coronary artery disease: A Mendelian randomization study. *Hepatology*. 2023;77(1):230-238.

78. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. *Hepatology*. 2018;67(1):328-357.

79. Pang Q, Zhang J-Y, Song S-D, et al. Central obesity and nonalcoholic fatty liver disease risk after adjusting for body mass index. *World Journal of Gastroenterology: WJG*. 2015;21(5):1650.

80. Younossi ZM, Golabi P, de Avila L, et al. The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: a systematic review and meta-analysis. *Journal of hepatology*. 2019;71(4):793-801.

81. Li G, Peng Y, Chen Z, Li H, Liu D, Ye X. Bidirectional Association between Hypertension and NAFLD: A Systematic Review and Meta-Analysis of Observational Studies. *International Journal of Endocrinology*. 2022;2022(1):8463640.

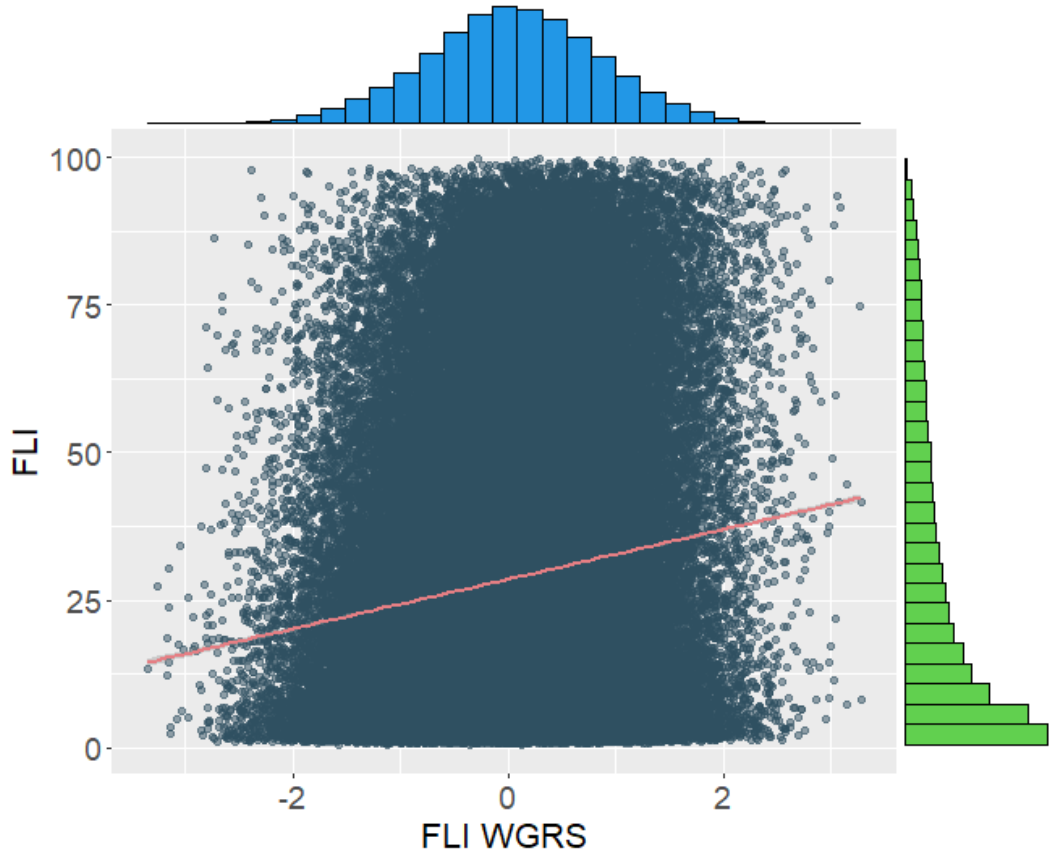
82. Assy N, Kaita K, Mymin D, Levy C, Rosser B, Minuk G. Fatty infiltration of liver in hyperlipidemic patients. *Digestive diseases and sciences*. 2000;45:1929-1934.

83. JAISWAL A, JAIN K, BABU N. Metabolic Profile of Lean/Non Obese NAFLD (Non Alcoholic Fatty Liver Disease) Subjects. *Journal of Clinical & Diagnostic Research*. 2021;15(11)

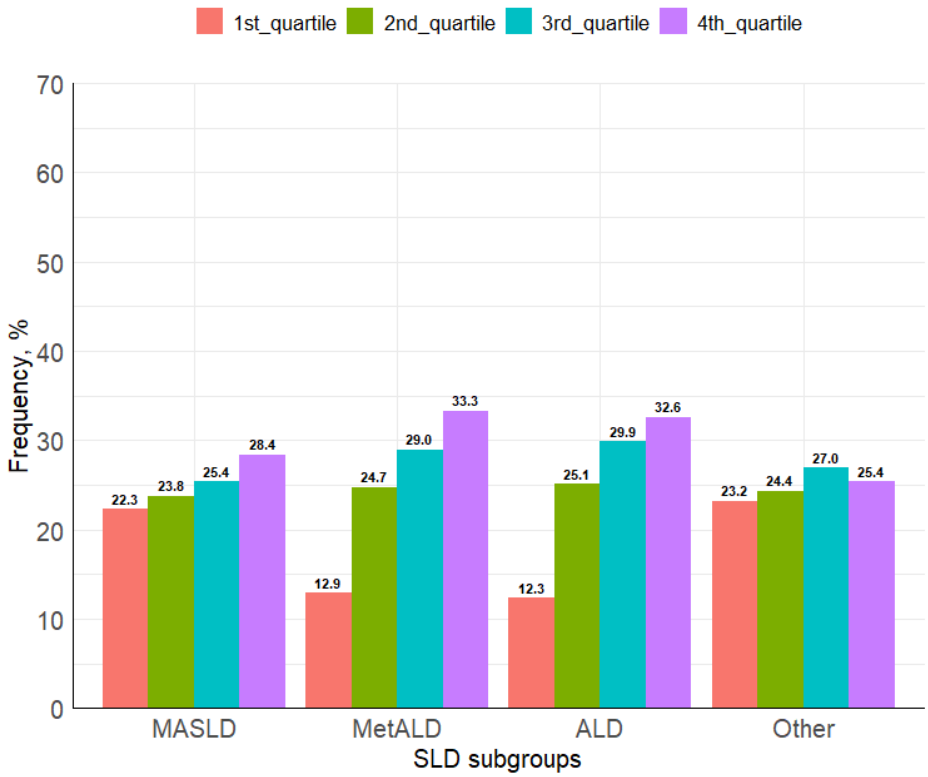
84. Tilg H, Adolph TE, Moschen AR. Multiple parallel hits hypothesis in nonalcoholic fatty liver disease: revisited after a decade. *Hepatology*. 2021;73(2):833-842.

85. Tilg H, Adolph TE, Dudek M, Knolle P. Non-alcoholic fatty liver disease: the interplay between metabolism, microbes and immunity. *Nature metabolism*. 2021;3(12):1596-1607.
86. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. *Nature medicine*. 2018;24(7):908-922.
87. Chu H, Duan Y, Yang L, Schnabl B. Small metabolites, possible big changes: a microbiota-centered view of non-alcoholic fatty liver disease. *Gut*. 2019;68(2):359-370.
88. Vuppalanchi R, Noureddin M, Alkhouri N, Sanyal AJ. Therapeutic pipeline in nonalcoholic steatohepatitis. *Nature Reviews Gastroenterology & Hepatology*. 2021;18(6):373-392.
89. Schwärzler J, Grabherr F, Grander C, Adolph TE, Tilg H. The pathophysiology of MASLD: an immunometabolic perspective. *Expert review of clinical immunology*. 2024;20(4):375-386.

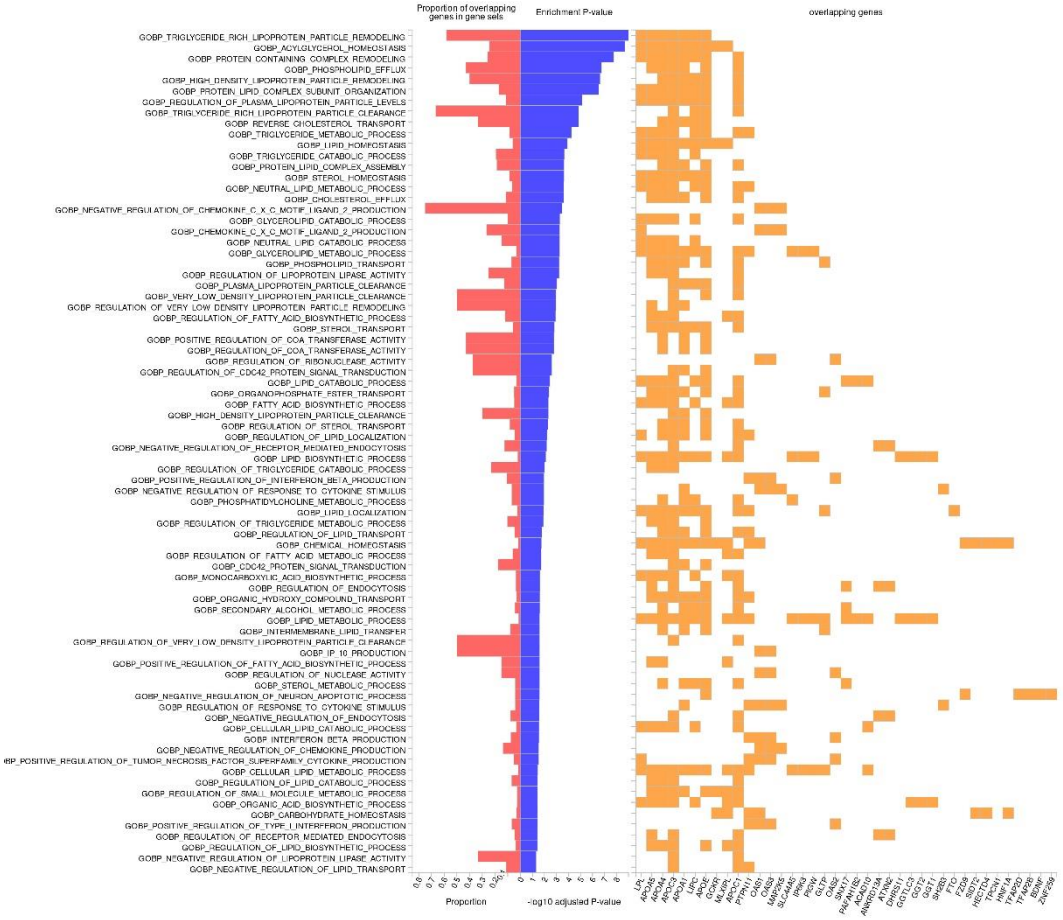
APPENDICES



Appendix 1. Scatter plot of WGRS for FLI vs. measured FLI



Appendix 2. FLI WGRS quartile distribution by SLD subtype



Appendix 3. Enrichment analysis of a MASLD, regarding GO biological processes (MASLD defined based on FLI ≥30)

Korean Abstract

대사기능 이상 관련 지방간질환과 심혈관계질환 위험의 연관성

연세대학교 대학원 보건학과

백 지 우

연구배경: 대사기능 이상 관련 지방간질환 (MASLD)은 비만과 대사증후군과 관련된 간질환으로, 심혈관계질환 (CVD) 위험 증가와의 연관성이 보고되어 왔다. 그러나, 기존 관찰연구의 한계로 인해 이 연관성이 실제 인과관계인지는 명확하지 않다. UK Biobank (UKB)와 같은 대형 바이오뱅크를 활용한 전장유전체분석(GWAS) 및 Mendelian Randomization (MR) 연구가 수행되었으나, 결과는 일관되지 않았다. 이에 본 연구는 한국과 영국의 개별 수준 데이터 및 일본의 요약 수준 데이터를 활용하여 MASLD 와 CVD 간의 연관성을 관찰연구 및 유전연구 방법을 통해 종합적으로 평가하고자 한다.

연구방법: 본 연구는 한국인암예방연구-II (Korean Cancer Prevention Study-II, KCPS-II) 코호트 자료를 기반으로 다단계 분석 전략을 수행하였다. 지방간 (SLD)은 Fatty Liver Index (FLI) 30 이상으로 정의하였다.

첫째, MASLD 와 CVD 및 그 하위 질환의 발생 간 연관성을 평가하기 위해 Cox 비례위험 회귀모형을 활용한 전향적 관찰연구를 수행하였다. 둘째, KCPS-II, UKB, Biobank Japan (BBJ)의 대규모 자료를 활용하여 MASLD 가 CVD 에 미치는 인과적 영향을 평가하기 위한 1 표본 및 양방향 2 표본 MR 분석을 수행하였다. 셋째, MASLD 관련 유전변이를 규명하기 위해 GWAS 를 수행하고, 유전자 기반 분석 및 조직 특이적 발현 분석을 실시하였다. 넷째, MASLD 에 대한 유전 효과가 흡연 노출에 따라 어떻게 달라지는지를 평가하고자 유전-흡연 상호작용 분석을 실시하였다.

연구결과: 총 111,637 명이 분석에 포함되었으며, 연령 중앙값은 39 세, 여성 비율은 35.7%였다. 기저 시점에서 32,018 명 (28.7%)이 MASLD 분류 기준을 충족하였으며, 중앙값 10 년의 추적 기간 동안 3,926 건의 CVD 사건이 발생하였다. 다변량 보정 후 MASLD 는 CVD 발생 위험을 유의하게 증가시켰으며, 위험비는 1.69 (95% CI: 1.57-1.82)였다. KCPS-II 개별 수준 데이터를 활용한 1 표본 MR 분석에서는 유전적으로 예측된 MASLD 가 전체

CVD 위험 증가와 유의한 연관성을 보였으며, 공변량 보정한 모형에서 HR 은 1.05 (95% CI: 1.03–1.08)이었다.

인과성을 추가로 검증하기 위해 BBJ 및 UKB 의 유전 데이터를 outcome 으로 활용한 2 표본 MR 분석을 수행한 결과, MASLD 가 CAD 위험을 증가시킨다는 IVW 분석 결과가 도출되었다. BBJ 에서는 OR 은 1.08 (95% CI: 1.05–1.13, $p = 1.83 \times 10^{-5}$), UKB 에서 OR 은 1.04 (95% CI: 1.01–1.08, $p = 6.03 \times 10^{-3}$)였다. 반대 방향의 분석에서는, UKB 에서 정의된 MASLD 와 KCPS-II 의 CVD 위험 증가와 유의한 관련성을 보여 OR 은 1.15 (95% CI: 1.08–1.24, $p = 4.39 \times 10^{-5}$)였다.

KCPS-II 데이터를 활용한 GWAS 분석에서는 MASLD 와 연관된 여러 유전자좌가 확인되었고, 그중 와 FTO 와 CUX2 가 주요 유전자였다. 발현 분석 결과 APOA1, APOC3, GCKR, and HNF1A 는 간 조직에서, RPH3A 는 뇌 조직에서 특이적으로 발현되는 유전자였다. 기능적 경로 분석 결과, 지방대사 및 관상동맥질환 관련 경로와 유의미한 관련성이 확인되었다.

MASLD 의 총 유전적 설명력은 38.6%로, 이 중 6.5%는 유전 요인, 5.3%는 유전-흡연 상호작용, 26.7%는 환경 잡음 성분에 의해 설명되었다. 또한, rs671 (ALDH2)을 포함한 일부 유전자가 흡연에 노출 시 더 강한 유전효과를 나타냈다.

결론: 본 연구는 MASLD 가 CVD 위험 증가와 유의한 관련이 있음을 관찰연구와 유전연구 접근을 통해 종합적으로 분석하였다. 관찰연구에서는 MASLD 가 CVD 발생과 유의한 연관성을 보였으며, 유전분석에서는 MASLD 가 CVD 발생에 기여할 수 있는 인과적 역할이 확인되었다. MASLD 관련 유전변이와 유전–환경 상호작용 분석은 이 질환의 복합적인 유전적 구조와 흡연과 같은 환경요인의 조절 효과를 부각시켰다. 이러한 결과는 대사이상 관련 간 건강이 심혈관질환 예방에 있어 중요성을 강조하며, MASLD 를 심혈관 위험 감소를 위한 조절 가능한 위험요인으로 고려할 수 있음을 시사한다.

키워드: 전향적 코호트, 유전학적 연구, 멘델리안무작위화, 대사기능이상 연관 지방간질환, 심혈관계질환