Effect of bariatric surgery on non-alcoholic fatty liver disease: an exploratory metabolomics and validation study

Short title: Bariatric surgery, metabolomics, and NAFLD remission

Mengyi Li, MD, PhD[1]*, Xinyu Wang, BM[2]*, Na Zeng, MSc[2]*, Zhiyu Wu, MSc[2], Canqing Yu. PhD[2,3,4], Dianjianyi Sun, PhD[2,3,4], Yang Liu, MD, PhD[1], Di Cao, MD, PhD[5], Peng Zhang, MD, PhD[1], Ling Yang, PhD[6,7], Yiping Chen, DPhil[6,7], Zhengming Chen, DPhil[6,7], Liming Li, MD[2,3,4], Zhongtao Zhang, MD, PhD[1], Jun Lv, PhD[2,3,4], Yuanjie Pang, DPhil[2,3]

*Contributed equally

1. Department of General Surgery, Beijing Friendship Hospital, Capital Medical University; State Key Lab of Digestive Health; National Clinical Research Center for Digestive Diseases, Beijing 100050, China
2. Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing 100191, China
3. Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education
4. Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing 100191, China
5. Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China
6. Medical Research Council Population Health Research Unit at the University of Oxford, Oxford, United Kingdom
7. Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, United Kingdom

Address for correspondence:

Dr Yuanjie Pang	Prof Jun Lv	Prof Zhongtao Zhang
Department of Epidemiology & Biostatistics	Department of Epidemiology & Biostatistics	Department of General Surgery
School of Public Health	School of Public Health	Beijing Friendship Hospital
Peking University	Peking University	Capital Medical University
38 Xueyuan Road	38 Xueyuan Road	95 Yongan Road
Beijing, 100191, China	Beijing, 100191, China	Beijing, 100050, China
Tel: 86-010-82801528	Tel: 86-010-82801528	Tel: 86-010-63139361
Fax: 86-010-82801530	Fax: 86-010-82801530	Fax: 86-010-63139361
yuanjie_p@163.com	epi.lvjun@vip.163.com	zhangzht@ccmu.edu.cn

Word count: abstract 246, main text 3016
(3 tables, 4 figures and 8 supplementary tables/figures)

Abbreviations

NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; CVD, cardiovascular disease; CKB, China Kadoorie Biobank; PDFF, proton density fat fraction; SG, sleeve gastrostomy; RYGB, Roux-en-Y gastric bypass; OAGB, one anastomosis gastric bypass; LFF, liver fat fraction; VCTE, vibration-controlled transient elastography; FAP, fat attenuation parameter; LSM, liver stiffness measurement; UPLC-MS/MS, ultra-performance liquid chromatography coupled to tandem mass spectrometry; FDR, false discovery rate; SMPD, small molecule pathway database; KEGG, kyoto encyclopedia of genes and genomes; CDCA, chenodeoxycholic acid; TCA, taurocholic acid

Acknowledgment

The most important acknowledgment is to the participants in the CKB study and the members of the survey teams in each of the 10 regional centers, as well as to the project development and management teams based at Beijing, Oxford and the 10 regional centers. We also would like to thank all the researchers in the Base-NAFLD cohort, especially the involved study investigators, clinicians, nurses, and technicians for dedicating their time and skills to the completion of this study from the principal site.

Data sharing statement

Data described in the article, code book, and analytic code will be made available from the China Kadoorie Biobank upon request (https://www.ckbiobank.org/data-

access/data-access-procedures) pending application and approval. The data sets generated and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Funding

This work was supported by National Natural Science Foundation of China (82304223, 82192901, 82192904, 82192900, 82070580), the National Key R&D Program of China (2023YFC3606302), and Beijing Natural Science Foundation (7242040). Base-NAFLD cohort was supported by Capital's Funds for Health Improvement and Research (2020-1-2021). The CKB baseline survey and the first re-survey were supported by a grant from the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up is supported by grants from the UK Wellcome Trust (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), grants (2016YFC0900500) from the National Key R&D Program of China, and Chinese Ministry of Science and Technology (2011BAI09B01). The funders had no role in the study design, data collection, data analysis and interpretation, writing of the report, or the decision to submit the article for publication.

Author contributions

ML and YP had full access to the data. ML, NZ, XW and YP conducted data analysis and are responsible for accuracy of the results and the decision to submit for publication. All authors were involved in study design, conduct, long-term follow-up, review and coding of disease events, interpretation of the results, or writing the report. All authors have read and approved the final version of the manuscript.

Ethical statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The metabolomics study in the Base-NAFLD study was approved by the Human Ethics Committee of Beijing Friendship Hospital, Capital Medical University (approval number: 2019-P2-194-03). The CKB study was approved by the Institutional Review Board of Peking University Health Science Center (approval number: IRB00001052-20040). Informed consent was obtained from all participants.

Abstract

Background: Bariatric surgery presents a significant alleviation for non-alcoholic fatty liver disease (NAFLD), which relies in part on achieving substantial weight loss in post-surgical period.

Objectives: We aimed to understand the effect of bariatric surgery on NAFLD remission via metabolomics and to validate the results in a general population-based cohort.

Settings: Bariatric surgery cohort at a single university hospital; general population-based cohort.

Methods: In a pilot study, ten patients with NAFLD who underwent bariatric surgery were enrolled. The remission of hepatic steatosis was assessed by MRI-derived proton density fat fraction (PDFF) before and 3-month after surgery. Temporal associations of body mass index (BMI) reduction, alteration in metabolomic biomarkers, and NAFLD remission were quantified by using cross-lagged models, which were then validated in a general population-based cohort (n=1,258).

Results: At 3-month after surgery, BMI reduction of 6.9 (SD 1.9) kg/m^2 and MRI-PDFF reduction of 9.6% (5.4) (all $p<0.001$) were achieved. Of the 64 metabolomic biomarkers quantified, 19 biomarkers showed significant differences between pre- and post-surgery (false discovery rate-corrected $p<0.05$). Temporal associations were observed between BMI reduction and 5 metabolomic biomarkers, while 3 (chenodeoxycholic acid [CDCA], palmitoylcarnitine, and hippuric acid) were further validated in the general population-based cohort. CDCA was able to explain 18% of the association between BMI reduction and NAFLD remission ($p<0.05$). In the general population-based cohort, Mendelian randomization showed that genetically elevated CDCA level was associated with a higher risk of liver fibrosis.

Conclusions: CDCA is a potential mediator and may predict long-term surgical benefits in liver fibrosis regression.

Keywords: *bariatric surgery; non-alcoholic fatty liver disease; metabolomics; chenodeoxycholic acid*

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a public health issue worldwide. According to the latest estimates, the prevalence of NAFLD was 32.5% in Chinese adults[1]. A recent nationwide study reported that the prevalence of liver steatosis was 44.39%[2]. NAFLD is the leading cause of liver-related morbidity and mortality, such as non-alcoholic steatohepatitis (NASH), advanced fibrosis, cirrhosis, and liver cancer[3]. Obesity, as one of the three defining criteria of metabolic dysfunction-associated fatty liver disease (MAFLD), is a critical determinant of clinical outcomes[4]. Adults with obesity have 10-fold higher risk of developing NAFLD compared with non-obese adults[5]. Bariatric surgery, also known as metabolic and bariatric surgery, is a minimal invasive procedure to modify gastrointestinal anatomy and to achieve substantial weight loss, and presents a therapeutic effect in patients with obesity-related diseases. Previous studies have shown that bariatric surgery provides long-term resolution of NASH and regression of fibrosis, and is associated with lower risks of cardiovascular disease (CVD), cancer, and all-cause mortality[6-8].

The health benefits of bariatric surgery can be classified into weight-dependent and weight-independent effects, through inducing changes in glucose homeostasis, gut hormones, bile acid metabolism, gut microbiota, and intestinal nutrient sensing[9]. Although animal studies have suggested potential mechanisms to explain these effects[10,11], there is limited evidence from population-based studies. In the past decades, advancement in metabolomics has provided emerging opportunities to understand the mechanisms underlying the metabolic improvements following

bariatric surgery. However, the majority of studies either examined the associations between weight loss and metabolomic biomarkers or compared levels of metabolomic biomarkers in NAFLD or NASH patients with healthy controls[12,13]. Few studies thus far have explored the temporal associations of weight loss following bariatric surgery, plasma metabolomic biomarkers, and NAFLD remission.

Understanding the temporal associations is important to examine the mediation role of key metabolomic biomarkers on the effect of weight loss on NAFLD remission, which provides valuable insights into the mechanisms of bariatric surgery. However, the follow-up period of published studies on bariatric surgery is mainly during 1-5 years, limiting the feasibility to evaluate the effects of key metabolomic biomarkers on long-term liver-related clinical outcomes including advanced fibrosis. An innovative solution to this problem is to conduct a pilot study in the bariatric surgery cohort and to validate study findings in the general population-based cohorts with both metabolomics and hepatic steatosis assessments. Large-scale population-based biobank studies (e.g. the China Kadoorie Biobank [CKB], UK Biobank), with stored blood samples and follow-up for long-term health outcomes, provide unique opportunities for such cross-disciplinary collaborations[14].

Therefore, the objectives of this study were (1) to determine the pre- vs post-surgery changes in circulating metabolomic biomarkers in patients undergoing bariatric surgery; (2) to explore the temporal associations between weight loss and metabolomic biomarkers, and to further assess the extent to which key metabolomic biomarkers mediated the effect on NAFLD remission; and (3) to validate the pilot

findings in the general population of the CKB cohort.

This manuscript is written following STROBE checklist.

Methods

Bariatric surgery cohort

"Bariatric surgery for non-alcoholic fatty liver diseases in individuals with obesity" (Base-NAFLD) is a prospective multi-center cohort (NCT04366999)[15]. It aims to adequately assess the effect of bariatric surgery in NAFLD-confirmed patients by preoperatively MRI-derived proton density fat fraction (PDFF) or intraoperatively liver histology (both assessments were needed before enrollment). All patients were diagnosed with obesity according to the WHO criteria for obesity in Asian populations (BMI ≥27.5 kg/m^2) and schedule for a primary standardized bariatric surgical procedure (sleeve gastrectomy [SG], Roux-en-Y gastric bypass [RYGB], or one anastomosis gastric bypass [OAGB]) recommended by a multidisciplinary team. The Ethics Committee of Beijing Friendship Hospital, Capital Medical University, approved the sample collection from this cohort, which was conducted as per the Declaration of Helsinki of the World Medical Association (2019-P2-194-04). Written informed consent was obtained from all patients.

All patients were examined using a 3.0-T MR scanner (GE Healthcare, Revolution HD 750W) with a multi-echo chemical-shift-encoded sequence (IDEAL IQ: TR: 5.9 ms, TE min full, six different echoes ranging from 1.0 to 5.0 ms, matrix: 140 × 160; FOV:44 cm, slice section thickness: 8 mm, NEX: 0.69, flip angle: 3° to minimize T1

bias, bandwidth: 111.11, acquisition time: 21s). Liver fat fraction (LFF) at baseline and 3 months measured by MRI-PDFF was to assess hepatic steatosis control over a short time frame. Two back-to-back radiologists (DC and HJ) from the principal site, blinded to the clinical and biological data, centrally reviewed the images using a GE Advantage Workstation (ADW4.6) and performed the LFF measurements. The LFF values were obtained by manually identifying three regions of interests (ROIs) approximately 1-2 cm in diameter in each of the 8 Couinaud liver segments (I to VIII) of the liver. The mean LFF level of each segment was measured twice, recorded, and an average value for each patient was obtained. Large vessels, hepatic margins, and apparent image artifacts were avoided.

Ten patients from the principal site with available follow-up data in this cohort were enrolled for further metabolomics analysis (**Fig. 1**). Based on the Base-NAFLD cohort, such sample size had a power of >80% to detect pre- and post-surgery difference of 1 (mean/standard deviation) in clinical parameters. Fasting serum specimens from these patients were collected before and 3 months after surgery. All samples were stored and measured according to the standard operating protocols. Blood samples were centrifuged for serum collection, divided into aliquots, and delivered to the study laboratory on dry ice. All samples were stored at -80°C freezer until analysis.

General population-based cohort

The validation study was conducted in the general population-based cohort, involving 1 258 participants attending both the first and the third resurvey with

metabolomics measurements in the China Kadoorie Biobank (CKB) (**Fig. 1**)[16]. Details of the study design is described in **Supplementary Methods**.

Trained and certified doctors performed vibration-controlled transient elastography (VCTE) using FibroTouch FT-100 (Hisky, Wuxi, China) in accordance with the standard protocol. Fat attenuation parameter (FAP) and liver stiffness measurement (LSM) were derived to assess liver steatosis and fibrosis, respectively. For a TE examination to be considered eligible, a success rate of >60% should be achieved with 10 successful measurements. NAFLD was diagnosed based on the liver steatosis (FAP ≥240 dB/m) in the absence of excessive alcohol consumption or positive HBsAg[2,17]. The presence of liver fibrosis was defined as liver fibrosis grade ≥F1 (LSM ≥7.3 kPa)[2].

Assessment of metabolomics

64 metabolomic biomarkers were quantified by targeted metabolomics platforms (**Table S1**), using a ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS) system. These metabolomic biomarkers included acylcarnitines (n=13), amino acids (n=19), amino acids-related (n=6), bile acids (n=12), biogenic amines (n=2), carbohydrates and related (n=1), carboxylic acids (n=3), fatty acids (n=6), and indoles derivatives (n=2). All metabolomic biomarkers were log-transformed and then standardized to have an SD of 1. Participants with metabolomic biomarkers levels >3SD away from the mean were excluded. Details of the analytical procedures were reported in **Supplementary Methods**.

Statistical analyses

We described the characteristics of patients in the bariatric surgery and the general population cohort. Paired t-test and independent t-test were used in each population. We then compared differences in metabolomics between pre- and post-surgery (**Fig. 2A**). Statistical analysis was performed using R version 4.3.2. The false discovery rate (FDR) was used for multiple comparisons[18], and the threshold of significance was considered to be less than 0.05. For metabolomic biomarkers that showed significant differences between post- and pre-surgery (FDR-corrected *p*-value <0.05), pathway analysis was performed using the web-based MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/)[19]. In total, 28 serum metabolic pathways were screened, including 23 from both the Small Molecule Pathway Database (SMPD) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, and 5 from KEGG alone.

We assessed the temporal associations of body mass index (BMI) decrease, changes in metabolomic biomarkers, and NAFLD remission in three steps (**Fig. 2B-D**). In Step 1 (**Fig. 2B**), temporal associations between BMI decrease and changes in metabolomic biomarkers were performed using cross-lagged models in the surgical cohort[20,21]. Essentially, cross-lagged models evaluate reciprocal relationships between multiple variables at different time points to evaluate the temporal associations and elucidate the directional influences between variables over time. Of the metabolomic biomarkers that were caused by BMI decrease but not the other way around (i.e. metabolomic biomarkers causing BMI decrease), the temporal

associations were validated in the CKB using Mendelian randomization. Mendelian randomization utilizes the random assortment of genes from parents to offspring at conception, and uses gene variants associated with the exposure of interest as unconfounded markers, thus providing an approach for assessing causation[22]. Mendelian randomization can be used to validate the temporal associations between BMI and metabolomics by utilizing genetically determined BMI at birth and associating it with metabolomics levels quantified in adulthood. We used 851 SNPs for BMI previously developed for the Chinese population as instrumental variables in Mendelian randomization[23]. Details of cross-lagged models and Mendelian randomization were shown in **Supplementary Methods**.

In Step 2 (**Fig. 2C**), of the metabolomic biomarkers that were validated in Step 1, we performed mediation analysis to assess whether metabolomic biomarkers were the mediators of the effect of BMI decrease on NAFLD remission[20,21]. Mediation analysis based on cross-lagged models utilizes the temporal associations among variables to explore the mediated effects between exposure and outcome. Including BMI, metabolomic biomarkers, and NAFLD status at two time points, the model estimates parameters a, b, and c' (**Fig. 3**) simultaneously and derives the proportion of mediating effects. In the bariatric surgery cohort, the remission of hepatic steatosis postoperatively was defined as LFF <5% on abdominal MRI-PDFF. For the metabolomic biomarkers identified as potential mediators, we examined the associations between these metabolomic biomarkers and risk of incident NAFLD in CKB using multivariable logistic regression.

In Step 3 (**Fig. 2D**), of the possible mediator(s) we identified in Step 2, we assessed the associations of genetically determined levels of metabolomic biomarkers (determined at birth) in relation to long-term risk of liver fibrosis (assessed at mean age of 64.9) by Mendelian randomization. We used 83 SNPs for chenodeoxycholic acid (CDCA) identified from an in-house GWAS as instrumental variables in Mendelian randomization. Details of genotyping, GWAS, and Mendelian randomization were described in **Supplementary Methods**.

Results

Baseline characteristics of study participants

Of the 138 participants in the Base-NAFLD cohort, 10 adults were selected including paired assessments pre-surgery and 3-month post-surgery (mean age 35.6, 80% women). Mean decrease in BMI was 6.9 kg/m^2 3-month after surgery (**Table 1**), and LFF levels assessed by MRI-PDFF was alleviated in all patients. In the CKB cohort, 1 258 adults were included at study baseline (mean age 52.1, 68.9% women). 651 participants were diagnosed with NAFLD at the 3rd resurvey (median follow-up 12.4 years). Participants who developed NAFLD were more likely to be female, from urban region, and more likely to smoke, but were less likely to drink (**Table S2**). Levels of systolic blood pressure, adiposity measures, blood biochemistry traits were also higher among participants who developed NAFLD.

Changes in metabolomic biomarkers before and after bariatric surgery

When comparing levels of metabolomic biomarkers pre- and post-surgery (**Table 2**),

significant differences were observed for 31 traits, and 19 survived FDR correction, including elevated levels of 6 traits (e.g. palmitoylcarnitine and glycine) and decreased levels of 13 traits (e.g. tyrosine and taurocholic acid [TCA]) post-surgery.

Those who survived FDR correction were included in the pathway analysis. Five metabolic pathways were identified in the analysis, including (1) aminoacyl-tRNA biosynthesis, (2) citrate cycle, (3) alanine, aspartate, and glutamate metabolism, (4) glyoxylate and dicarboxylate metabolism, and (5) phenylalanine, tyrosine, and tryptophan biosynthesis (**Fig. S1**).

Temporal associations between BMI and metabolomic biomarkers

Cross-lagged analysis in the bariatric surgery cohort showed temporal associations between BMI and 5 metabolomic biomarkers (**Table 3**), including CDCA, palmitoylcarnitine (C16), hippuric acid, DHA, and proline. The paths of BMI at baseline (i.e. pre-surgery) to levels of 5 metabolomic biomarkers at follow-up (i.e. post-surgery) were significant, suggesting that BMI had an effect on these biomarkers. The reverse paths of BMI at follow-up and metabolomic biomarkers at baseline were non-significant, suggesting that these biomarkers did not have an effect on BMI.

In the CKB cohort, the temporal associations for 3 metabolomic biomarkers were validated using Mendelian randomization (**Table S3**), including CDCA, C16, and hippuric acid. Genetically determined BMI was associated with higher levels of CDCA and C16, and was associated with lower levels of hippuric acid.

Mediation analysis of BMI, metabolomic biomarkers, and NAFLD

Of the 5 metabolomic biomarkers showing temporal associations with BMI, CDCA and hippuric acid were associated with NAFLD risk in CKB, while the association between C16 and NAFLD was borderline significant (**Table S4**). According to the results above, CACD and hippuric acid were selected as potential mediators in the mediation analysis in the bariatric surgery cohort. BMI was associated with NAFLD with a significant direct effect, indirect effect, and total effect (all $p<0.05$, **Fig. 3**). The level of CDCA explained 18% ($p<0.05$) of the association between BMI decrease and NAFLD remission. In contrast, the mediation effect of hippuric acid was non-significant (**Fig. S3**).

In the CKB cohort, Mendelian randomization showed associations of genetically elevated BMI with CDCA and of genetically elevated BMI with NAFLD, but there was no evidence of genetic association of CDCA and NAFLD (**Fig. 4**). For liver fibrosis, an advanced stage of NAFLD, genetically elevated CDCA was associated with higher risk of fibrosis.

Discussion

In this pilot study of patients undergoing bariatric surgery, significant pre- and post-surgery differences were observed both for NAFLD remission and in plasma levels of metabolomic biomarkers. These differences were driven by 19 metabolomic biomarkers and involved several pathways, including aminoacyl-tRNA biosynthesis, citrate cycle, as well as metabolism in glyoxylate, dicarboxylate, and amino acids.

Cross-lagged analysis revealed temporal associations between BMI and 5 metabolomic biomarkers, and three were validated in CKB, suggesting that BMI change was causally associated with changes in CDCA, palmitoylcarnitine, and hippuric acid. Mediation analysis showed that CDCA mediated 18% of the effect of BMI decrease on NAFLD remission in the bariatric surgery cohort, while genetically elevated CDCA level was also associated with risk of liver fibrosis in the CKB cohort.

At least 20 previous studies have examined the associations between BMI and metabolomic biomarkers in the general population[24]. Although several small studies examined these associations in patients undergoing bariatric surgery, these studies were inherently limited by the cross-sectional study design and provided limited insights into temporality. Nonetheless, these study findings showed that bariatric surgery induced significant weight loss and was associated with substantial changes in metabolomic biomarkers, involving amino acid derivatives, lipid derivatives, bile acids, microbiota-derived metabolomic biomarkers, tricarboxylic acid cycle-related metabolomic biomarkers, endocannabinoids, etc[12,13]. Consistent with previous studies, our studies showed that pathways regulating aminoacyl-tRNA biosynthesis, citrate cycle, as well as metabolism in glyoxylate, dicarboxylate, and amino acids were perturbed by bariatric surgery. By taking advantage of cross-lagged analysis, we showed temporal associations of BMI and 5 metabolomic biomarkers, and further validated three biomarkers in the general population.

A key finding in our study was that CDCA mediated 18% of the effect of BMI decrease on NAFLD remission in the bariatric surgery cohort. This suggests that

CDCA might be a promising biomarker to understand the mechanisms of weight loss on NAFLD remission as well as for predicting liver-related benefits after bariatric surgery. Indeed, although mediation effect could not be validated in the CKB cohort because of study design (i.e. CDCA was measured only once at baseline), CDCA was associated with higher risk of NAFLD in the CKB cohort. Moreover, there was evidence of positive association between genetically determined levels of CDCA and liver fibrosis risk, suggesting a possible causal role of CDCA in the advanced liver outcome. At least 10 case-control or cross-sectional studies have shown elevated levels of CDCA in NAFLD or NASH patients compared with healthy control[25-28]. However, none has examined the role of CDCA in bariatric surgery, nor as a biomarker of interventions in the clinical context. CDCA (chenodeoxycholic acid), a type of primary bile acids, is synthesized in hepatocytes and transported into the intestine where it is metabolized by the microbiota into the secondary bile acids[27]. A recent study found that 3-sucCA, a secondary bile acid, alleviates metabolic dysfunction-associated steatohepatitis by promoting the growth of Akkermansia muciniphila, a specific type of gut symbiont. This suggests a possible mechanism through which CDCA could influence NAFLD progression[29]. CDCA is the drug target for gallstone disease[30]. A recent study showed that CDCA is also a promising drug target for pancreatic necrosis[31]. Taken together findings from the pilot study in the bariatric surgery cohort and validation in the CKB cohort, CDCA alongside bile acid metabolisms are promising targets to understand the effects of weight loss interventions and to predict long-term liver benefits of such interventions.

Strengths of this study included the exploratory stage in the pilot study of the bariatric surgery cohort, the validation stage in the general population-based cohort, detailed assessment of clinical parameters including NAFLD, quantitative assessment of metabolomic biomarkers, and advanced statistical techniques of causal inference (i.e. cross-lagged models, Mendelian randomization). These strengths supported the validation of study findings and increased the generalizability. However, there were several limitations in this study. First, although the follow-up period of the bariatric surgery was 3 months, rapid BMI decrease alongside NAFLD remission were observed. Ongoing follow-up of this cohort will show long-term effects of bariatric surgery on metabolomics. Nonetheless, findings of the bariatric surgery cohort were validated by Mendelian randomization, which mimicked the lifelong effects of elevated BMI and CDCA. Second, Mendelian randomization rested on three assumptions and the interpretation of results relied heavily on these assumptions. However, our previous studies showed validation of the first two assumptions (i.e. the relevance assumption and the independence assumption). We also conducted sensitivity analysis (i.e. MR-Egger, **Table S3**) and showed that the results remained generally consistent, suggesting that the third assumption (i.e. the restriction exclusion assumption) is plausible. Third, despite validation in the general population in CKB, this study was conducted in the Asian population where a higher percentage of body fat (especially visceral fat) at lower BMI levels reflects an increased risk of major chronic diseases. Future studies are warranted in different ethnicities to support the generalization of our findings.

In conclusion, this study showed that metabolomics is promising to understand the effect of bariatric surgery on NAFLD remission and identified CDCA as a metabolomic biomarker to understand the weight loss-dependent effect of bariatric surgery on rapid remission of NAFLD. There was evidence of association between genetically determined CDCA and liver fibrosis, suggesting the role of CDCA in predicting long-term surgical benefits in lowering risk of liver disease progression.

Disclosure statement

The authors have no relevant financial or non-financial interests to disclose.

References:

1. Riazi K, Azhari H, Charette JH, et al. The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 2022; 7(9): 851-861.
2. Man S, Deng Y, Ma Y, et al. Prevalence of liver steatosis and fibrosis in the general population and various high-risk populations: a nationwide study with 5.7 million adults in China. *Gastroenterology* 2023; 165(4): 1025-1040.
3. Wong VW, Ekstedt M, Wong GL, Hagström H. Changing epidemiology, global trends and implications for outcomes of NAFLD. *J Hepatol* 2023; 79(3): 842-852.
4. Eslam M, Newsome PN, Sarin SK, et al. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. *J Hepatol* 2020; 73(1): 202-209.
5. Pang Y, Kartsonaki C, Turnbull I, et al. Adiposity in relation to risks of fatty liver, cirrhosis and liver cancer: a prospective study of 0.5 million Chinese adults. *Sci Rep* 2019; 9(1): 785.
6. Lassailly G, Caiazzo R, Ntandja-Wandji LC, et al. Bariatric surgery provides long-term resolution of nonalcoholic steatohepatitis and regression of fibrosis. *Gastroenterology* 2020; 159(4): 1290-1301.
7. Rustgi VK, Li Y, Gupta K, et al. Bariatric surgery reduces cancer risk in adults with nonalcoholic fatty liver disease and severe obesity. *Gastroenterology* 2021; 161(1): 171-184.
8. Adams TD, Davidson LE, Litwin SE, et al. Weight and metabolic outcomes 12

years after gastric bypass. *N Engl J Med* 2017; 377(12): 1143-1155.

9. Rubino F, Panagiotopoulos S. Metabolic surgery: the cutting edge of diabetes care. *Nat Rev Gastroenterol Hepatol* 2017; 14(7): 389-390.
10. Dreyfuss JM, Yuchi Y, Dong X, et al. High-throughput mediation analysis of human proteome and metabolome identifies mediators of post-bariatric surgical diabetes control. *Nat Commun* 2021; 12(1): 6951.
11. Quante M, Iske J, Heinbokel T, et al. Restored TDCA and valine levels imitate the effects of bariatric surgery. *Elife* 2021; 10.
12. Ha J, Kwon Y, Park S. Metabolomics in bariatric surgery: towards identification of mechanisms and biomarkers of metabolic outcomes. *Obes Surg* 2021; 31(10): 4564-4574.
13. Vaz M, Pereira SS, Monteiro MP. Metabolomic signatures after bariatric surgery - a systematic review. *Rev Endocr Metab Disord* 2022; 23(3): 503-519.
14. Chen Z, Emberson J, Collins R. Strategic Need for Large Prospective Studies in Different Populations. *JAMA* 2020; 323(4): 309-310.
15. Wei L, Li M, Zeng N, et al. Bariatric surgery for non-alcoholic fatty liver disease in individuals with obesity (Base-NAFLD): protocol of a prospective multicenter observational follow-up study. *BMC Surg* 2021; 21(1): 298.
16. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011; 40(6): 1652-1666.
17. Fan JG, Wei L, Zhuang H. Guidelines of prevention and treatment of

nonalcoholic fatty liver disease (2018, China). *J Dig Dis* 2019; 20(4): 163-173.

18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* 1995; 57(1): 289-300.
19. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* 2021; 49. doi:10.1093/nar/gkab382
20. Kenny DA. Cross-lagged panel correlation: a test for spuriousness. *Psychol Bull* 1975; 82(6): 887-903.
21. Kivimäki M, Feldt T, Vahtera J, Nurmi J-E. Sense of coherence and health: evidence from two cross-lagged longitudinal samples. *Soc Sci & Med* 2000; 50(4): 583-597.
22. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; 362. doi:10.1136/bmj.k601
23. Fairhurst-Hunter Z, Lin K, Millwood IY, et al. Trans-ancestry meta-analysis improves performance of genetic scores for multiple adiposity-related traits in East Asian populations. medRxiv 2022. doi: https://doi.org/10.1101/2022.07.05.22277254
24. Pang Y, Kartsonaki C, Lv J, et al. Adiposity, metabolomic biomarkers, and risk of nonalcoholic fatty liver disease: a case-cohort study. *Am J Clin Nutr* 2022; 115(3): 799-810.

25. Grzych G, Chávez-Talavera O, Descat A, et al. NASH-related increases in plasma bile acid levels depend on insulin resistance. *JHEP Rep* 2021; 3(2). doi:10.1016/j.jhepr.2020.100222

26. Caussy C, Hsu C, Singh S, et al. Serum bile acid patterns are associated with the presence of NAFLD in twins, and dose-dependent changes with increase in fibrosis stage in patients with biopsy-proven NAFLD. *Aliment Pharmacol Ther* 2019; 49(2): 183-193.

27. Chávez-Talavera O, Tailleux A, Lefebvre P, Staels B. Bile acid control of metabolism and inflammation in obesity, type 2 diabetes, dyslipidemia, and nonalcoholic fatty liver disease. *Gastroenterology* 2017; 152(7): 1679-1694.

28. Chow MD, Lee YH, Guo GL. The role of bile acids in nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Mol Aspects Med* 2017; 56: 34-44.

29. Nie Q, Luo X, Wang K, et al. Gut symbionts alleviate MASH through a secondary bile acid biosynthetic pathway. Cell 2024. doi:10.1016/j.cell.2024.03.034

30. Maudgal DP, Northfield TC. A practical guide to the nonsurgical treatment of gallstones. *Drugs* 1991; 41(2): 185-192.

31. Zhu Q, Yuan C, Dong X, et al. Bile acid metabolomics identifies chenodeoxycholic acid as a therapeutic agent for pancreatic necrosis. *Cell Rep Med* 2023; 4(12). doi:10.1016/j.xcrm.2023.101304.

Figure legends

Figure 1. Flow diagram

*The proportion of individuals with NAFLD at CKB 1st resurvey.

Figure 2. Conceptual framework and key results of the current study

(A) Differential metabolomic biomarkers pre-surgery vs post-surgery. (B) Temporal associations between BMI and metabolomics. (C) Cross-lagged panel model of the mediation effect of CDCA between BMI and NAFLD remission. (D) Validation of the associations between BMI, metabolomic biomarkers, and NAFLD in CKB.

Figure 3. Cross-lagged panel model of the mediation effect of CDCA between BMI and NAFLD remission

The subscript *t* indicates traits at baseline (i.e. CKB 1st resurvey), and the subscript *t+1* indicates traits at follow-up (i.e. CKB 3rd resurvey). Model fit: SRMR=standardized root mean squared residual, CFI=comparative fit index. **p*-value <0.05; ***p*-value <0.01; ****p*-value<0.001

Figure 4. Validation of the associations between BMI, metabolomic biomarkers, and NAFLD in CKB

Dotted lines denote significant results (*p*-value <0.05) from observational analyses. Black and grey solid lines denote significant (*p*-value <0.05) and insignificant (*p*-value ≥0.05) results from genetic analyses. Effect estimates are shown as follows: (1) SD difference in log-transformed CDCA levels per 1-SD higher BMI. In the genetic analysis, 2 stage least square (2SLS) was used, both stages adjusted for

age, sex, 10 regions, smoking status, alcohol status, physical activity, 11 national principal components, and GWAS array type at 1st resurvey; in the observational analysis, linear regression was adjusted for age, sex, 10 regions, smoking status, alcohol status, and physical activity at 1st resurvey. (2) OR of NAFLD per 1-SD higher log-transformed CDCA. In the genetic analysis, 2 sample Mendelian randomization (2SMR) was used, the SNP-outcome association adjusted for the same covariates as in 2SLS, restricted to those whose BMI >28 kg/m^2 without excessive alcohol consumption or positive HBsAg; in the observational analysis, logistic regression was adjusted for the same covariates as in linear regression. (3) OR of NAFLD per 1-SD higher BMI. Convention of the genetic and observational analysis as in (1) and (2), respectively; (4) OR of fibrosis per 1-SD higher log-transformed CDCA. Convention of the genetic analysis as in (2).

Table 1. Characteristics of patients at baseline and 3 months after surgery

Variables	Baseline (n=10)	3 months (n=10)	*p*-value
Age (SD), y	35.6 (4.4)	35.6 (4.4)	-
Women (%)	80.0	80.0	-
Blood and anthropometry			
Systolic blood pressure (SD), mmHg	138.1 (14.1)	128.5 (14.0)	0.109
Diastolic blood pressure (SD), mmHg	84.6 (12.4)	75.8 (8.2)	0.100
Body weight (SD), kg	120.9 (22.1)	101.6 (21.9)	<0.001*
BMI (SD), kg/m^2	43.4 (6.5)	36.5 (7.0)	<0.001*
Waist circumference (SD), cm	123.2 (13.8)	107.7 (11.4)	<0.001*
Biochemistry			
ALT (SD), U/L	44.6 (28.7)	21.3 (7.4)	0.029*
AST (SD), U/L	30.8 (20.8)	21.1 (4.7)	0.132
GGT (SD), U/L	28.1 (7.5)	16.8 (5.2)	<0.001*
Total cholesterol (SD), mmol/L	4.9 (0.8)	4.9 (1.0)	0.807
HDL-C (SD), mmol/L	1.0 (0.1)	1.1 (0.2)	0.029*
LDL-C (SD), mmol/L	3.1 (0.6)	3.1 (0.7)	0.855
Triglycerides (SD), mmol/L	1.7 (1.0)	1.2 (0.4)	0.035*
Albumin (SD), g/L	40.8 (2.6)	43.3 (2.0)	0.012*
Creatinine (SD), umol/L	64.2 (14.9)	63.8 (14.7)	0.845
Fasting glucose (SD), mmol/L	5.7 (1.1)	4.9 (0.3)	0.084
HbA1c (SD), %	5.8 (0.5)	5.4 (0.3)	0.017*
LFF(SD), %	16.7 (7.5)	7.1 (4.5)	<0.001*
NAFLD activity score (SD)	5.2 (1.0)	-	-
Fibrosis stage (%)		-	-
F0	2 (20.0)		
F1	3 (30.0)		
F2	3 (30.0)		
F3	2 (20.0)		

P-value was obtained from paired-sample t-test.

Abbreviations: SD, standard deviation; BMI, body mass index; ALT, alanine transaminase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transpeptidase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; LFF, liver fat fraction; NAFLD, non-alcoholic fatty liver disease.

44

564 **Table 2. Differences in metabolomic biomarkers between post- and pre-surgery**

565

Metabolite	Mean		Mean difference (95% CI)	*p*-value	FDR-adjusted *p*-value	Metabolite	Mean		Mean difference (95% CI)	*p*-value	FDR-adjusted *p*-value
	Pre-surgery	Post-surgery					Pre-surgery	Post-surgery			
Acylcarnitines						***Amino acids***					
Palmitoylcarnitine	0.09	0.13	-0.04 (-0.05, -0.02)	1.09E-04	**0.002**	Tyrosine	69.16	45.81	23.35 (17.67, 29.03)	6.52E-06	**4.17E-04**
L-Carnitine	49.79	38.05	11.74 (6.54, 16.93)	0.001	**0.007**	Tryptophan	57.17	40.98	16.19 (11.22, 21.16)	3.23E-05	**0.001**
Propionylcarnitine	0.39	0.25	0.14 (0.06, 0.22)	0.003	**0.015**	Methionine	28.66	21.06	7.60 (4.54, 10.67)	3.31E-04	**0.005**
Acetylcarnitine	7.6	13.91	-6.31 (-10.18, -2.45)	0.005	**0.023**	Lysine	224.1	181.18	42.93 (23.32, 62.54)	0.001	**0.007**
Tetradecanoylcarnitine	0.01	0.02	0.00 (-0.01, 0.00)	0.013	**0.043**	Glycine	175.55	246.71	-71.16 (-108.02, -34.30)	0.002	**0.013**
Stearylcarnitine	0.03	0.04	-0.01 (-0.01, 0.00)	0.026	0.067	Arginine	136.13	113.12	23.00 (10.06, 35.95)	0.003	**0.015**
Decanoylcarnitine	0.16	0.31	-0.15 (-0.30, -0.01)	0.039	0.084	Leucine	125.96	99.4	26.56 (9.21, 43.92)	0.007	**0.028**
Glutarylcarnitine	0.03	0.02	0.00 (0.00, 0.01)	0.039	0.084	Alanine	402.68	311.1	91.58 (31.29, 151.87)	0.007	**0.028**
Dodecanoylcarnitine	0.04	0.06	-0.02 (-0.04, 0.00)	0.037	0.084	Proline	157.91	132.32	25.59 (8.91, 42.28)	0.007	**0.028**
Octanoylcarnitine	0.13	0.23	-0.10 (-0.21, 0.01)	0.062	0.116	Histidine	73.47	63.9	9.57 (2.13, 17.01)	0.017	0.054
Valerylcarnitine	0.003	0.002	0.001 (0.000, 0.002)	0.154	0.227	Aspartic acid	20.53	25.35	-4.82 (-8.61, -1.04)	0.018	0.054
Hexanoylcarnitine	0.05	0.07	-0.02 (-0.05, 0.01)	0.184	0.262	Serine	116.87	136.91	-20.04 (-36.95, -3.13)	0.025	0.067
Malonylcarnitine	0.06	0.05	0.01 (-0.01, 0.02)	0.277	0.362	Valine	196.24	161.3	34.93 (4.40, 65.46)	0.029	0.069
Bile acids						Isoleucine	68.44	54.46	13.99 (-2.27, 30.24)	0.083	0.141
TCA	0.06	0.03	0.03 (0.01, 0.05)	0.009	**0.034**	Asparagine	35.3	31.48	3.82 (-1.55, 9.19)	0.142	0.216
GCDCA	1.56	0.57	0.99 (0.20, 1.78)	0.019	0.054	Threonine	61.73	52.7	9.03 (-4.17, 22.22)	0.156	0.227
GDCA	0.24	0.06	0.18 (0.03, 0.33)	0.027	0.067	Phenylalanine	97.65	87.89	9.76 (-8.99, 28.51)	0.269	0.359
GCA	0.32	0.12	0.20 (0.01, 0.40)	0.043	0.089	Glutamine	484.04	468.89	15.15 (-57.13, 87.43)	0.647	0.690
TCDCA	0.19	0.10	0.09 (-0.01, 0.18)	0.061	0.116	Glutamic acid	34.97	36.41	-1.44 (-8.65, 5.78)	0.663	0.696
CDCA	0.49	0.17	0.32 (-0.03, 0.66)	0.066	0.120	***Amino acids related***					
CA	0.28	0.09	0.19 (-0.02, 0.41)	0.076	0.131	Citrulline	21.74	19.21	2.53 (-0.03, 5.08)	0.052	0.104
TDCA	0.03	0.01	0.02 (0.00, 0.04)	0.091	0.145	Aminoadipic acid	4.19	1.81	2.38 (-0.21, 4.96)	0.067	0.120
GLCA-3S	0.16	0.23	-0.08 (-0.21, 0.05)	0.204	0.278	4-Hydroxyproline	5.50	6.14	-0.64 (-2.05, 0.77)	0.330	0.398

GLCA	0.003	0.001	0.001 (-0.002, 0.004)	0.304	0.382	1-Methylhistidine	5.55	5.00	0.55 (-0.81, 1.90)	0.387	0.458
DCA	0.16	0.09	0.06 (-0.08, 0.20)	0.323	0.398	Sarcosine	1.91	1.55	0.36 (-0.56, 1.29)	0.398	0.463
GUDCA	0.41	0.64	-0.22 (-0.89, 0.44)	0.466	0.523	Ornithine	73.08	69.57	3.50 (-10.81, 17.82)	0.593	0.655

Table 2. Continued

Metabolite	Mean		Mean difference (95% CI)	*p*-value	FDR-adjusted *p*-value	Metabolite	Mean		Mean difference (95% CI)	*p*-value	FDR-adjusted *p*-value
	Pre-surgery	Post-surgery					Pre-surgery	Post-surgery			
Biogenic amines						***Fatty acids-Free/non-covalently bond***					
GABA	0.13	0.12	0.01 (-0.05, 0.07)	0.688	0.711	Arachidonic acid	4.31	7.49	-3.18 (-4.84, -1.52)	0.002	**0.013**
Beta-Alanine	3.55	3.49	0.06 (-0.93, 1.04)	0.900	0.900	DHA	5.55	6.69	-1.15 (-2.06, -0.24)	0.019	0.054
Carboxylic acids						EPA	0.25	0.39	-0.13 (-0.29, 0.03)	0.091	0.145
Cis-Aconitic acid	44.46	65.05	-20.59 (-31.91, -9.27)	0.003	**0.015**	Myristic acid	9.34	8.47	0.87 (-0.56, 2.29)	0.202	0.278
Succinic acid	2.15	2.79	-0.64 (-0.99, -0.28)	0.003	**0.015**	Oleic acid	310.74	329.66	-18.92 (-74.36, 36.53)	0.460	0.523
Hippuric acid	0.58	1.23	-0.65 (-1.48, 0.19)	0.112	0.175	Dodecanoic acid	1.25	1.28	-0.03 (-0.24, 0.18)	0.732	0.743
Carbohydrates and related											
Glucose	5585.09	4422.42	1162.67 (635.51, 1689.83)	0.001	**0.007**						

Abbreviations: FDR, false discovery rate; TCA, taurocholic acid; GCDCA, glycochenodeoxycholic acid; GDCA, glycodeoxycholic acid; GCA, glycocholic acid; TCDCA, taurochenodeoxycholic acid; CDCA, chenodeoxycholic acid; CA, cholic acid; TDCA, taurodeoxycholic acid; GLCA, glucuronic acid; DCA, deoxycholic acid; GUDC, glycoursodeoxycholic acid; GABA, gamma-aminobutyric acid; DHA, docosahexaenoic acid; EPA, eicosapentaenoic acid.

Table 3. Cross-lagged associations between BMI and metabolomic biomarkers

Metabolite	$BMI_{baseline} \rightarrow Trait_{follow\text{-}up}$			$Trait_{baseline} \rightarrow BMI_{follow\text{-}up}$			Model fit	
	β	SE	p-value	β	SE	p-value	SRMR	CFI
CDCA	0.663	0.202	0.001	-0.026	0.114	0.818	<0.001	1
Palmitoylcarnitine	0.516	0.211	0.014	0.003	0.092	0.978	<0.001	1
Hippuric acid	-0.544	0.263	0.038	-0.053	0.085	0.533	<0.001	1
DHA	0.516	0.255	0.043	-0.103	0.143	0.471	<0.001	1
Proline	0.531	0.267	0.046	-0.047	0.112	0.674	<0.001	1

Abbreviations: CDCA, chenodeoxycholic acid; DHA, docosahexaenoic acid; SE, standard error; SRMR, standardized root mean squared residual; CFI, comparative fit index.

Figure 1. Flow diagram

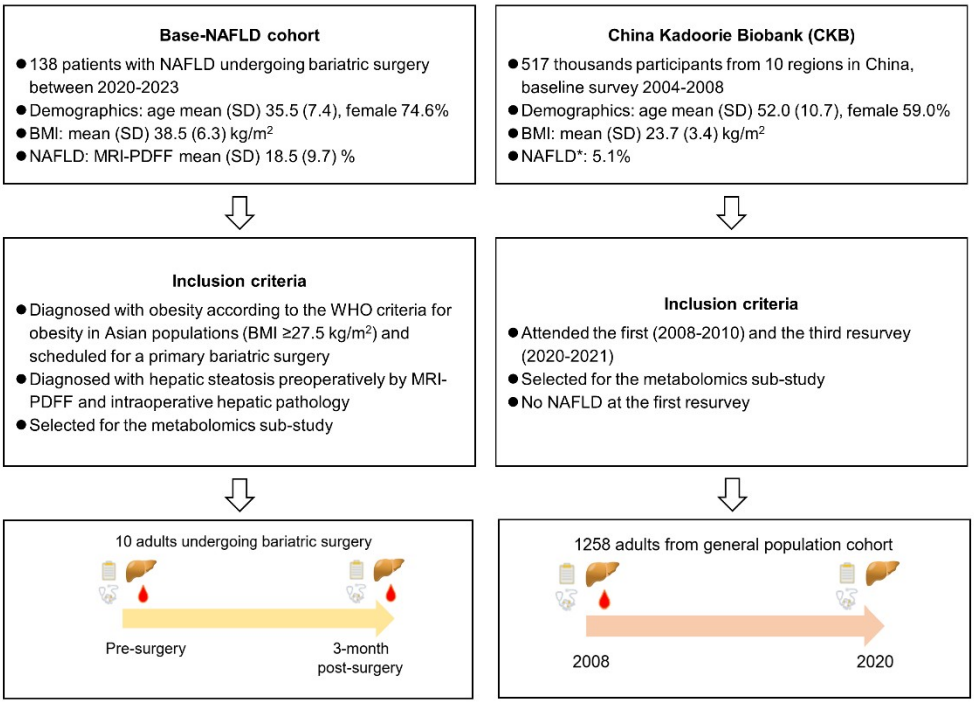

Figure 2. Conceptual framework and key results of the current study

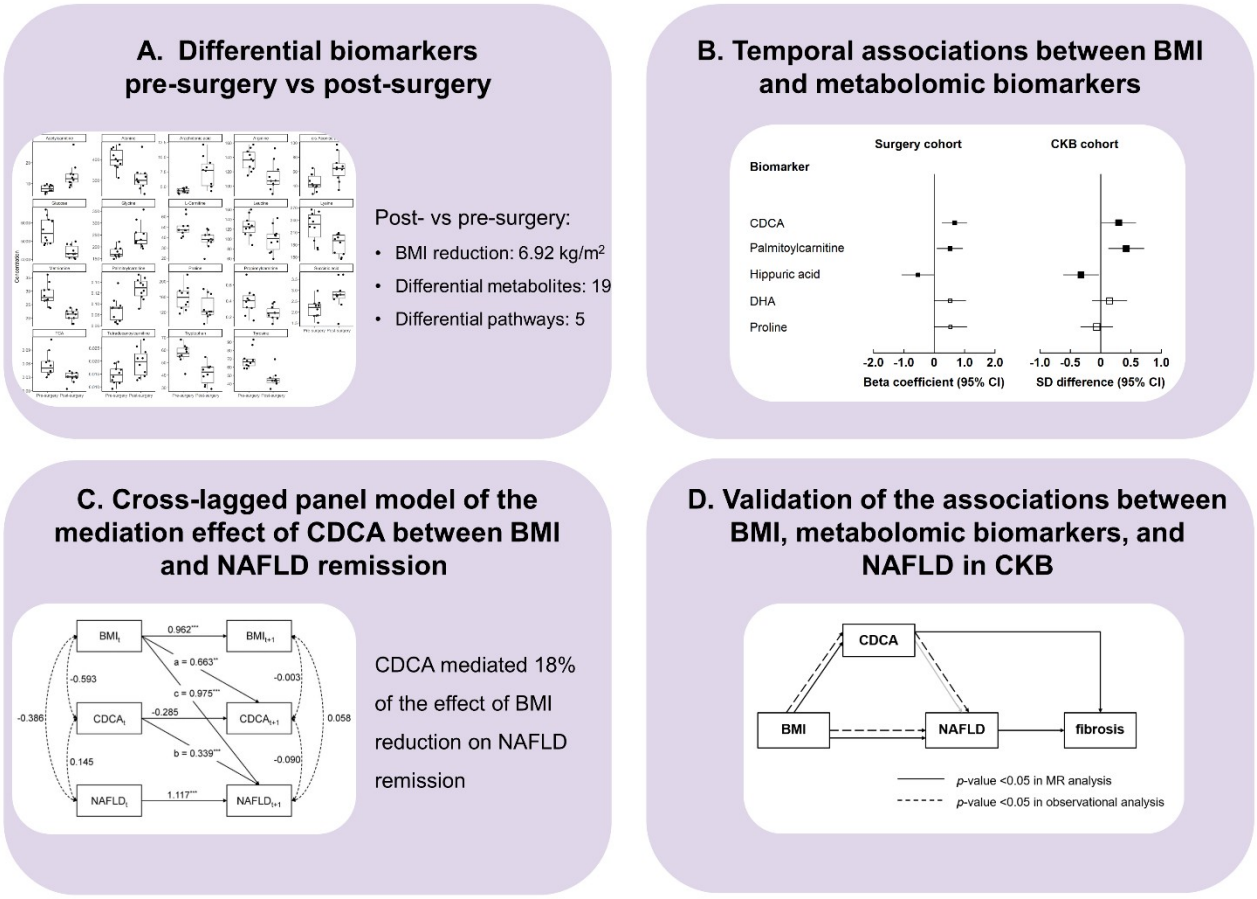

Figure 3. Cross-lagged panel model of the mediation effect of CDCA between BMI and NAFLD remission

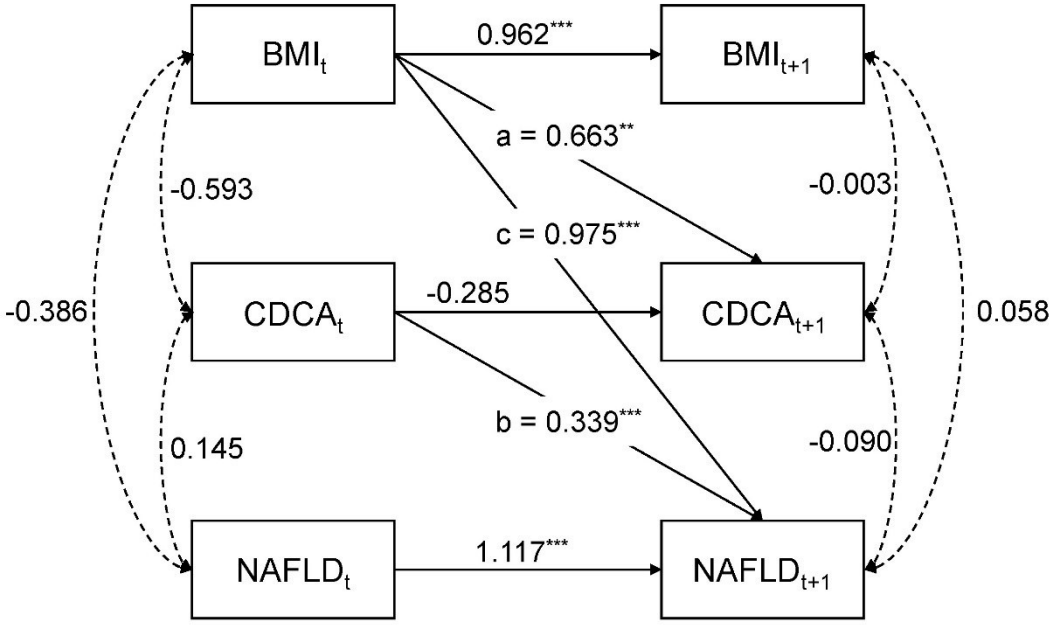

Figure 4. Validation of the associations between BMI, metabolomic biomarkers, and NAFLD in CKB