BMJ Open Gastroenterology

Metabolomic profiles of incident gallstone disease

Eugenia Uche-Anya,[1,2] Jane Ha ,[1,2] Raji Balasubramanian,[3] Kathryn M Rexrode,[4] Andrew T Chan[1,2,5,6]

To cite: Uche-Anya E, Ha J, Balasubramanian R, *et al*. Metabolomic profiles of incident gallstone disease. *BMJ Open Gastroenterol* 2024;**11**:e001417. doi:10.1136/bmjgast-2024-001417

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjgast-2024-001417).

Received 22 March 2024
Accepted 26 July 2024

For numbered affiliations see end of article.

Correspondence to
Dr Eugenia Uche-Anya; eucheanya@mgh.harvard.edu

ABSTRACT

Background and aims Gallstone disease affects ≥40 million people in the USA and accounts for health costs of ≥$4 billion a year. Risk factors such as obesity and metabolic syndrome are well established. However, data are limited on relevant metabolomic alterations that could offer mechanistic and predictive insights into gallstone disease. This study prospectively identifies and externally validates circulating prediagnostic metabolites associated with incident gallstone disease.

Methods Female participants in Nurses' Health Study (NHS) and Nurses' Health Study II (NHS II) who were free of known gallstones (N=9960) were prospectively followed up after baseline metabolomic profiling with liquid chromatography–tandem mass spectrometry. Multivariable logistic regression and enrichment analysis were used to identify metabolites and metabolite groups associated with incident gallstone disease at P_{FDR}<0.05. Findings were validated in 1866 female participants in the Women's Health Initiative and a comparative analysis was performed with 2178 male participants in the Health Professionals Follow-up Study.

Results After multivariate adjustment for lifestyle and putative risk factors, we identified and externally validated 17 metabolites associated with incident gallstone disease in women—nine triacylglycerols (TAGs) and diacylglycerols (DAGs) were positively associated, while eight plasmalogens and cholesterol ester (CE) were negatively associated. Enrichment analysis in male and female cohorts revealed positive class associations with DAGs, TAGs (≤56 carbon atoms and ≤3 double bonds) and de novo TAG biosynthesis pathways, as well as inverse associations with CEs.

Conclusions This study highlights several metabolites (TAGs, DAGs, plasmalogens and CE) that could be implicated in the aetiopathogenesis of gallstone disease and serve as clinically relevant markers.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Gallstone disease is prevalent and costly, yet remains understudied.

⇒ Obesity and metabolic syndrome are established risk factors, but data are limited on relevant prediagnostic metabolomic alterations that could offer mechanistic and predictive insights.

WHAT THIS STUDY ADDS

⇒ 17 validated metabolites were associated with incident gallstone disease in women.

⇒ Metabolite set associations were seen in male and female cohorts with diacylglycerols, triacylglycerols (≤56 carbon atoms and ≤3 double bonds), de novo triacylglycerol biosynthesis pathways and cholesterol esters.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study highlights several metabolites, metabolite classes and pathways that could be instrumental for future mechanistic studies that advance the understanding of gallstone disease and also leveraged as risk-prediction biomarkers, as well as preventive and therapeutic targets.

INTRODUCTION

Gallstone disease is one of the most common digestive disorders in the USA and a leading driver of health costs, accounting for over $4 billion annually.[1] This substantial burden to the health system is expected to grow as the prevalence of gallstone disease and its complications has steadily increased over the last few decades.[2–4] Yet, despite affecting >10% of US adults,[2 5] gallstone disease is comparatively understudied and significant research gaps remain in clinical management.[6]

Established risk factors for cholesterol gallstone formation include obesity and metabolic syndrome,[7–9] which interplay with lithogenic gene variants and environmental influences, to alter cholesterol homeostasis and promote biliary cholesterol supersaturation.[10] However, mounting recent evidence points towards a more complex multidirectional relationship with shared but incompletely elucidated mechanistic pathways.[11 12] Further characterisation of these underlying mechanisms could prove to be mutually beneficial in developing better preventive, diagnostic and therapeutic strategies.

Metabolomics can provide insight into physiological processes and pathophysiological alterations that contribute to disease states, which could be instrumental in identifying prognostic biomarkers and therapeutic

targets of clinical value.[13] Obesity and metabolic syndrome have been associated with variations in plasma metabolites, some of which have shown improved predictive and risk-stratification performance over anthropometric measures such as body mass index (BMI).[14–16] Data on the plasma metabolome in relation to gallstone disease are limited but could be of comparable utility. This study aims to identify prediagnostic circulating metabolites that are associated with incident gallstone disease in two prospective female cohorts, validate findings in an external independent female cohort and perform a comparative analysis with a male cohort.

MATERIAL AND METHODS

Study population

The primary analysis was performed in two large US prospective cohorts: the Nurses' Health Study (NHS) and the Nurses' Health Study II (NHS II). NHS enrolled 121 700 female registered nurses aged 30–55 years at its inception in 1976, while NHS II enrolled 116 429 female registered nurses aged 25–42 years at its inception in 1989.[17] Participants reported lifestyle and health-related data at baseline and biennially on follow-up questionnaires. Blood samples were provided by 32 826 NHS participants in 1989–1990 and 29 611 NHS II participants in 1996–1999. Blood samples were drawn in sodium heparin tubes and shipped with an ice pack by overnight courier to the laboratory, processed (centrifuged and aliquoted into plasma, white blood cell and red blood cell components) and subsequently stored in liquid nitrogen freezers at −130°C.[17 18] 73.6% of participants were fasting at the time of blood draw. Participants with baseline metabolomic profiling from 13 nested case-control studies in NHS and NHS II were included (see online supplemental table 1). Exclusion criteria included history of gallbladder disease prior to blood sample collection and loss-to follow-up afterwards. After exclusions, 9960 participants with prediagnostic specimens and follow-up through 2021 were included in the primary analysis. The NHS and NHS II study protocols were approved by the Institutional Review Boards (IRBs) at Brigham and Women's Hospital and Harvard T. H. Chan School of Public Health. Voluntary return of questionnaires and blood samples was regarded as informed consent according to the IRBs.

Validation analysis was performed with female participants from a nested case-control study of coronary artery disease[19] in the Women's Health Initiative (WHI) cohort. The WHI cohort is a large US prospective study which follows 161 808 postmenopausal women aged 50–79 since enrolment in 1993–1999.[20] A total of 93 676 participants were enrolled in an observational study (WHI-OS), while 68 132 participants were enrolled in four randomised controlled trials, two of which were hormone-replacement therapy (WHI-HT) randomised placebo-controlled trials involving 16 608 participants.[21] A total of 2306 participants had baseline prediagnostic metabolomic profiling with follow-up through 2005, including 1362 and 944 matched cases and controls from WHI-HT (active and placebo arms) and WHI-OS, respectively.[19] 99.7% of participants were fasting at the time of blood draw. Samples were either processed and stored at −70°C within 2 hours of collection or shipped with dry ice at −20°C for up to 2 days prior to processing and storage at −70°C.[19] After excluding participants with history of gallbladder disease at baseline, 1866 participants were included in the validation analysis. The WHI study protocols were approved by the IRB at Brigham and Women's Hospital. All participants provided written informed consent prior to participating in the study.

A comparative analysis was performed with male participants from six nested case-control studies in the Health Professionals Follow-up Study (HPFS) cohort. HPFS is a large US prospective cohort which enrolled 51 529 male health professionals aged 40–75 years at its launch in 1986.[22] Blood samples were collected from 18 018 HPFS participants in 1993–1995.[23] 59.0% of participants were fasting at time of blood draw. The study protocol in HPFS was similar to that in NHS and NHS II and was approved by the IRBs at Brigham and Women's Hospital and Harvard T. H. Chan School of Public Health. After exclusions, 2178 HPFS participants with prediagnostic specimens and no history of gallstone disease were included in the analysis with follow-up through 2022.

Metabolomics

Metabolomic profiling in NHS, NHS II, HPFS and WHI was performed using a liquid chromatography–tandem mass spectrometry (LC-MS) platform at the Broad Institute of Massachusetts Institute of Technology and Harvard University.[18 19] Water-soluble metabolites were analysed using hydrophilic interaction liquid chromatography (HILIC) in positive-ion mode analyses (HILIC-positive) and negative-ion mode analyses (HILIC-negative). Plasma lipids were profiled using C8-positive analyses, while free fatty acids and bile acids were profiled using C18-negative analyses. The methods for HILIC-positive, HILIC-negative, C8-positive and C18-negative analyses are described in detail (see online supplemental table 2). Of note, although all WHI participants had metabolites analysed via all four methods, only a reduced proportion of NHS and NHS II participants had HILIC-negative and C18-negative metabolites. Metabolite quantification was conducted by the integration of LC-MS peak areas, which are proportional to metabolite concentration values. Pooled plasma reference samples were analysed every 20 samples for standardisation of each result by multiplying the ratio of the sample measurement to that of the nearest pooled reference measurement by the median of all reference values for the metabolite. Only annotated metabolites were included in the analysis.

Outcome

NHS, NHS II and HPFS participants were asked on biennial questionnaires to disclose whether they were

diagnosed with gallstones or underwent cholecystectomy in the preceding 2 years. Prior validation with medical record review in a random sample of participants demonstrated ≥99% accuracy with self-report of gallstones and/or cholecystectomy.[24 25] Case definition for incident gallstone disease was limited to reports of cholecystectomy as it was unknown whether gallstones in the absence of cholecystectomy were symptomatic. In WHI, case definition for incident gallstone disease was ascertained by self-report of gallbladder disease based on the available information provided on follow-up questionnaires.

Covariates

In NHS, NHS II, HPFS and WHI, participants were asked to report baseline demographic, lifestyle and health-related information including age, height, weight, BMI, physical activity, tobacco use, alcohol use, dietary information quantified with the Alternative Healthy Eating Index-2010 (AHEI) score (excluding alcohol component),[26 27] thiazide use and fasting status at time of sample collection. Female participants also reported data on parity, oral contraceptive use (NHS II only), menopausal status and menopausal hormonal therapy use.

Statistical analysis

As only a smaller subset of NHS, NHS II and HPFS participants had HILIC-negative and C18-negative analyses performed, the primary analysis was restricted to HILIC-positive and C8-positive metabolites. Additionally, metabolites affected by delayed sample processing[18] or metabolites with an intraclass correlation coefficient >0.3 were excluded from the analysis. Metabolites were log-transformed and standardised with further exclusion of metabolites with ≥25% missingness across participants and subsequent imputation of the remaining missing values using random forest imputation as is typically recommended for metabolomics data analyses.[28] After exclusions, a total of 238 metabolites were included in the analysis.

Multivariable-adjusted logistic regression models were used to investigate associations of individual metabolites with incident gallstone disease (see online supplemental table 3). The model was adjusted for age (years), fasting status (≥8 hours) at time of sample collection, case-control status, BMI (kg/m^2), rapid weight change (±15 lbs in the preceding 2 years), tobacco use (never smoker, past smoker or current smoker), alcohol use (g/day), physical activity in metabolic-equivalent hours (MET-hours/week), AHEI dietary score (excluding alcohol component), thiazide use (yes or no), parity (yes or no), oral contraceptive use (never or ever) and menopausal hormonal therapy use (never or ever). The p value threshold was set <0.05 and adjusted for multiple testing by controlling the false discovery rate (FDR) using the Benjamini-Hochberg approach. In the HPFS and WHI replication analyses, the same multivariable adjusted model was applied to metabolites that were included the primary analysis (all HILIC-positive and C8-positive). A secondary analysis was performed in WHI with metabolites not analysed in NHS and NHS II, which included HILIC-negative and C18-negative metabolites.

To identify metabolite classes associated with incident gallstone disease, metabolite set enrichment analysis (MSEA)[29] was performed based on the metabolite regression coefficients from the multivariable-adjusted logistic regression model. MSEAs were done for metabolite sets that were grouped by chemical structure as well as for metabolite sets grouped by involvement in similar biological pathways. Metabolite sets included in the analysis are detailed (see online supplemental table 4). Statistical analyses were conducted in SAS V.9.4 and R V.4.2.

RESULTS

Baseline characteristics of 9960 NHS and NHS II participants in the primary analysis as well as 1866 WHI participants in the validation analysis and 2178 HPFS participants in the comparative analysis are described in detail (see table 1). WHI and HPFS participants tended to be older (mean age 66.9 and 63.5 years, respectively) compared with NHS participants (mean age 56.8 years), while NHS II participants were younger (mean age 44.6 years). Less than 4% of the NHS, NHS II and HPFS cohorts identified as non-white compared with 21% of the WHI cohort. NHS II participants had lower tobacco and alcohol use, HPFS participants had higher alcohol use and physical activity, while WHI participants had lower levels of physical activity. Gallstone incidence was higher in NHS (17.7%) and NHS II (12.2%) compared with WHI (9.1%) and HPFS (7.8%).

In the primary analysis conducted in the pooled NHS and NHS II cohorts, 1585 (15.9%) incident cases of gallstone disease were identified among participants who had prediagnostic metabolomic profiling over a median follow-up of 24.1 years. After adjusting for confounders and lifestyle risk factors in the multivariable model, 86 metabolites (50 positive, 36 negative) were significantly associated with incident gallstone disease at $P_{FDR}<0.05$. When stratified by cohort, 62 metabolites and 9 metabolites were significantly associated with incident gallstone disease at $P_{FDR}<0.05$ in NHS (N=6792, cases=1200) and NHS II (N=3168, cases=385), respectively. Results from the primary discovery analysis are described in detail (see online supplemental table 5).

In the replication analyses, 170 (7.8%) and 169 (9.14%) incident cases of gallstone disease were identified over a median follow-up of 20.5 and 7.9 years among HPFS and WHI participants, respectively. Using the same multivariable model for metabolites that met $P_{FDR}<0.05$ in the primary analysis, we validated the significant associations seen with incident gallstone disease for 17 metabolites in the independent female WHI cohort. All the validated metabolites that were positively associated were either triacylglycerols (TAGs: 46:1, 46:2, 48:1, 48:2, 48:3, 50:2 and 50:3) or diacylglycerols (DAGs: 32:1 and 34:2), while

Table 1 Baseline characteristics of participants in NHS, NHS II, HPFS and WHI cohorts

Characteristic	Discovery cohorts		Replication cohorts	
	NHS (N=6792)	NHS II (N=3168)	HPFS (N=2178)	WHI (N=1866)
Age, years	56.8 (6.8)	44.6 (4.5)	63.5 (8.3)	66.9 (6.9)
White	96.1%	97.1%	98.3%	78.8%
Black	3.3%	1.2%	0.2%	13.6%
Hispanic	0.9%	1.1%	–	2.6%
Asian or Pacific Islander	0.5%	1.3%	0.2%	1.9%
Native American	0.1%	0.2%	–	0.6%
Other	0.2%	0.3%	1.2%	2.5%
BMI, kg/m^2	25.5 (4.6)	25.6 (5.4)	25.6 (3.0)	28.3 (5.9)
Never smoker, %	46.6%	67.1%	49.6%	49.4%
Alcohol, g/day	5.9 (10.4)	3.5 (6.4)	11.3 (14.8)	4.9 (11.4)
AHEI dietary score	47.4 (9.9)	45.8 (10.1)	49.1 (11.0)	47.0 (9.7)
Physical activity, MET-hours/week	16.0 (22.5)	18.1 (22.0)	32.4 (29.1)	11.4 (13.0)
Thiazide use, %	13.4%	3.8%	4.0%	6.0%
Menopausal hormone therapy use, %	31.8%	11.0%	–	47.8%
Premenopausal	19.5%	75.1%	–	0.0%
Parity, %	92.9%	82.6%	–	88.4%
Fasting at sample collection, %	74.6%	71.3%	60.0%	99.6%
Incident gallstone cases, N	1200 (17.7%)	385 (12.2%)	170 (7.8%)	169 (9.1%)
Time to event, years	10.5 (7.9)	9.1 (5.8)	10.6 (6.9)	4.7 (2.5)
Follow-up time, years	25.4 (7.0)	22.2 (3.5)	19.6 (7.8)	7.4 (2.0)

AHEI, Alternative Healthy Eating Index-2010 score; BMI, body mass index; HPFS, Health Professionals Follow-up Study; NHS, Nurses' Health Study; WHI, Women's Health Initiative.

all those that were inversely associated were either phosphatidylcholine (PC) or phosphatidylethanolamine (PE) plasmalogens (34:3 PE, 36:3 PC, 36:4 PE, 36:5 PE, 38:6 PE, 38:7 PE and 40:7 PE plasmalogens) or cholesterol ester (CE: 22:6). Of note, no statistically significant associations at $P_{FDR}<0.05$ were seen for individual metabolites in the comparative analysis done in the male HPFS cohort. Results for validated metabolites in the WHI cohort are illustrated (see table 2). Stratifying the analysis for all female participants by cohort, C50:3 TAG was significantly associated with incident gallstone disease in the NHS, NHS II and WHI cohorts. Five amino acids—alanine, histidine, threonine, proline and lysine—were positively associated in the pooled discovery analysis, as well as the younger NHS II cohort but mostly non-significant in the largely postmenopausal NHS and WHI cohorts (see table 2 and online supplemental table 5). In a secondary analysis where the multivariable model was applied in WHI for metabolites not analysed in the primary analysis (particularly but not limited to HILIC-negative and C18-negative metabolites), glycolithocholate and C20:2 CE were found to be positively and negatively associated with incident gallstone disease, respectively (see online supplemental figure 1).

Metabolite classes associated with incident gallstone disease in enrichment analyses are shown (see figure 1). TAGs with ≤56 C and ≤3 double bonds and DAGs were positively associated with incident gallstone disease while plasmalogens and CEs were inversely associated with incident gallstone disease in the NHS and NHS II cohorts. The positive relationship seen with TAGs with ≤56 C and ≤3 double bonds and DAGs was replicated in both HPFS and WHI cohorts. The inverse relationship with plasmalogens was reproduced in WHI but not HPFS. Carnitines were also inversely associated with incident gallstone disease in the HPFS and WHI cohorts.

Enrichment analyses by biological pathway in NHS and NHS II revealed that metabolites involved in de novo TAG biosynthesis, glycerolipid metabolism, aminoacyl t-RNA biosynthesis and PE/PC biosynthesis were positively associated with incident gallstone disease. The signal for aminoacyl t-RNA biosynthesis appeared to be driven by the younger NHS II cohort. The positive association with de novo TAG biosynthesis and glycerolipid metabolism was re-demonstrated in

Table 2 Risk estimates for validated metabolite associations with incident gallstone disease

HMDB ID	Metabolite	NHS and NHS II (OR (95% CI, P_{FDR}))	WHI (OR (95% CI, P_{FDR}))	HPFS (OR (95% CI, P_{FDR}))
Positive				
HMDB0005433	C50:3 TAG	1.16 (1.09–1.23, 0.0001)*	1.33 (1.11–1.60, 0.045)*	1.19 (1.01–1.41, 0.198)
HMDB0005377	C50:2 TAG	1.12 (1.06–1.19, 0.001)*	1.37 (1.14–1.64, 0.045)*	1.20 (1.03–1.41, 0.178)
HMDB0005432	C48:3 TAG	1.12 (1.06–1.19, 0.0007)*	1.34 (1.11–1.60, 0.045)*	1.23 (1.06–1.43, 0.146)
HMDB0005376	C48:2 TAG	1.11 (1.05–1.18, 0.002)*	1.34 (1.12–1.61, 0.045)*	1.19 (1.02–1.40, 0.195)
HMDB0007103	C34:2 DAG	1.11 (1.05–1.17, 0.003)*	1.30 (1.09–1.55, 0.045)*	1.21 (1.05–1.40, 0.146)
HMDB0010419	C46:2 TAG	1.10 (1.04–1.16, 0.006)*	1.31 (1.10–1.57, 0.045)*	1.19 (1.03–1.38, 0.178)
HMDB0005359	C48:1 TAG	1.09 (1.03–1.15, 0.012)*	1.32 (1.10–1.58, 0.045)*	1.19 (1.02–1.39, 0.178)
HMDB0007099	C32:1 DAG	1.09 (1.03–1.15, 0.011)*	1.37 (1.14–1.65, 0.045)*	1.19 (1.03–1.37, 0.178)
HMDB0010412	C46:1 TAG	1.09 (1.03–1.15, 0.012)*	1.31 (1.09–1.57, 0.045)*	1.17 (1.01–1.36, 0.225)
Negative				
HMDB0011394	C40:7 PE plasmalogen	0.87 (0.82–0.92, 0.0001)*	0.71 (0.60–0.85, 0.010)*	1.08 (0.92–1.26, 0.753)
HMDB0006733	C22:6 CE	0.87 (0.82–0.93, 0.0003)*	0.76 (0.64–0.90, 0.045)*	1.00 (0.84–1.19, 0.995)
HMDB0011420	C38:7 PE plasmalogen	0.89 (0.83–0.94, 0.001)*	0.77 (0.65–0.91, 0.045)*	1.05 (0.89–1.24, 0.932)
HMDB0011244	C36:3 PC plasmalogen	0.91 (0.85–0.96, 0.008)*	0.77 (0.65–0.91, 0.045)*	0.87 (0.73–1.04, 0.445)
HMDB0011442	C36:4 PE plasmalogen	0.91 (0.86–0.96, 0.008)*	0.79 (0.68–0.93, 0.045)*	1.02 (0.87–1.20, 0.990)
HMDB0011387	C38:6 PE plasmalogen	0.91 (0.86–0.96, 0.008)*	0.77 (0.65–0.91, 0.045)*	0.99 (0.84–1.16, 0.995)
HMDB0011343	C34:3 PE plasmalogen	0.91 (0.86–0.96, 0.006)*	0.74 (0.63–0.86, 0.010)*	1.08 (0.92–1.26, 0.730)
HMDB0011410	C36:5 PE plasmalogen	0.92 (0.87–0.98, 0.023)*	0.77 (0.65–0.91, 0.045)*	1.01 (0.86–1.18, 0.995)

*Estimates for covariates adjusted for in the multivariable model are detailed in online supplemental table 3.
CE, cholesterol ester; DAG, diacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; TAG, triacylglycerol.

the HPFS and WHI cohorts. PE/PC biosynthesis was replicated in WHI only, while glycerolipid metabolism was replicated in HPFS only.

DISCUSSION

Although the causal mechanism has not been fully elucidated, altered lipid metabolism is widely considered to be a driver of gallstone disease.[30–32] Large-scale genetic studies of gallstone disease have associated several loci involved in lipid homeostasis with increased gallstone disease risk.[33–35] Also, there are data demonstrating that gut microbial factors promote gallstone formation by influencing lipid metabolism.[36] Prospective data on metabolomic correlates for gallstone disease risk are limited but could provide further insights. We identified and validated 17 prediagnostic metabolites associated with incident gallstone disease in women, after adjusting for lifestyle and putative risk factors. These included seven TAGs and two DAGs that were positively associated with incident gallstone disease, while one CE and seven PE and PC plasmalogens were inversely associated (see table 2). Although no individual metabolite associations were seen in the male HPFS cohort, DAGs, TAGs with ≤56 carbon atoms and ≤3 double bonds, de novo TAG biosynthesis pathways and glycerolipid metabolism pathways were enriched in both male and female cohorts. When the pooled primary analysis was stratified by cohort, C50:3 TAG was consistently positively associated with incident gallstone disease across NHS, NHS II and WHI cohorts. Of note, studies in the literature involving murine models and human liver tissue have identified all nine validated TAGs (46:1, 46:2, 48:1, 48:2, 48:3, 50:2 and 50:3) and validated DAGs (32:1 and 34:2) as biomarkers of hepatic steatosis.[37–39] This offers biological plausibility for our findings given increasing evidence on the bidirectional relationship between metabolic dysfunction-associated steatotic liver disease and gallstone disease.[40 41]

Plasmalogens—a class of PC or PE glycerophospholipids found on cell membranes that contain a vinyl ester bond in the sn-1 position—were inversely associated with incident gallstone risk in this analysis. Plasmalogens have been inversely associated with other disorders including neurodegenerative disorders, peroxisome diseases and coronary artery disease, although the exact mechanisms of action have only recently become an area of active research.[42] Murine model studies of atherosclerosis suggest that PE plasmalogens influences lipid metabolism by converting excess cholesterol to bile acids and excretion into faeces via increased cholesterol 7-alpha hydroxylase expression and suppressed farnesoid X receptor expression[43 44]—factors that have been implicated

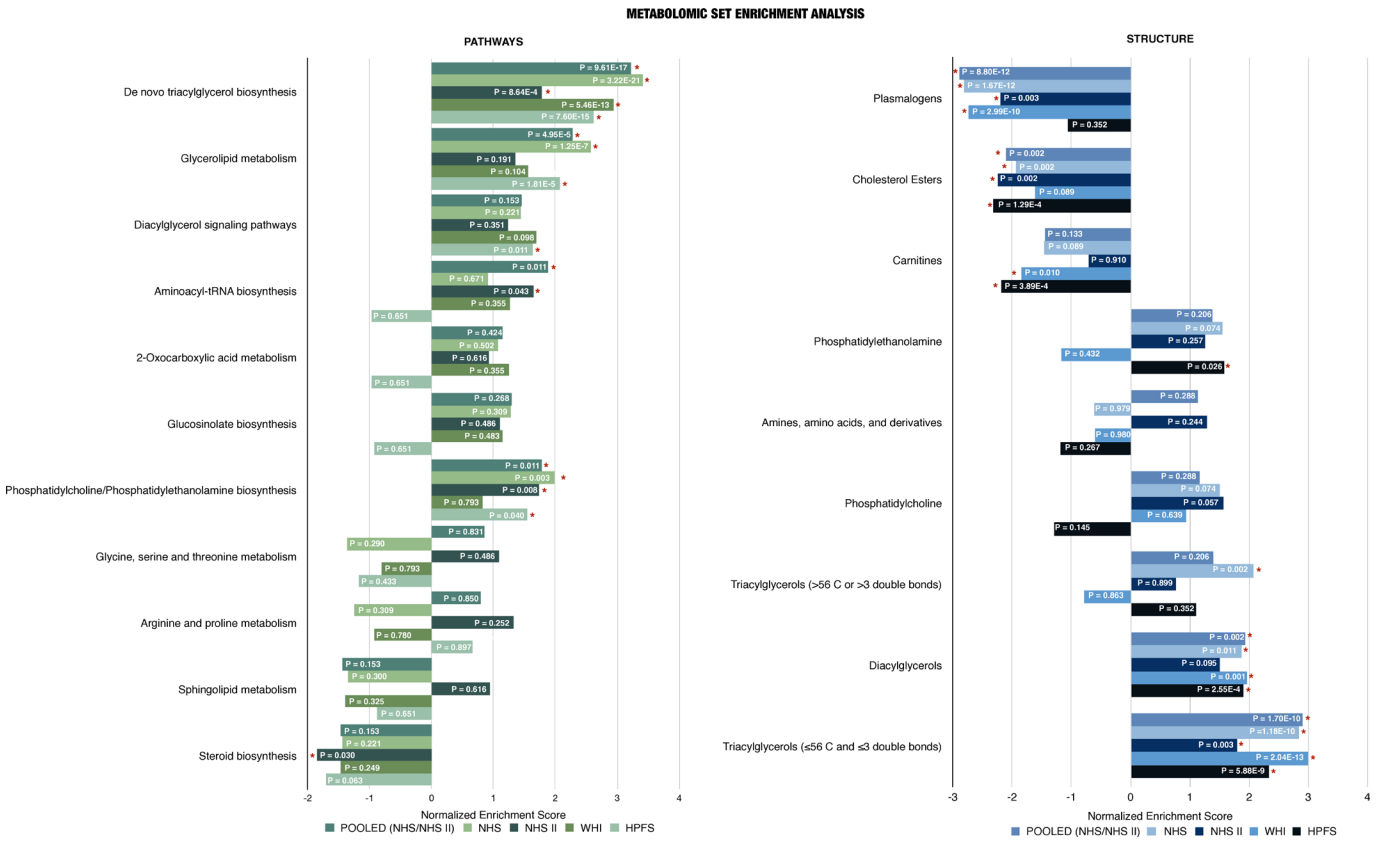

Figure 1 Metabolite set enrichment analysis (structure and pathway). HPFS, Health Professionals Follow-up Study; NHS, Nurses' Health Study; WHI, Women's Health Initiative.

in gallstone formation.[32] In the secondary analysis of HILIC negative and C18 negative metabolites performed in WHI, the only bile acid found to be associated with incident gallstone disease was glycolithocholate. Cross-sectional studies have reported associations with circulating plasma bile acids such as ursodeoxycholic acid, taurochenodeoxycholic acid, taurohyodeoxycholic acid, taurocholic acid, glycocholic acid and glycochenodeoxycholic acid[45 46]; however, these associations were not demonstrated in our prospective analysis.

CEs were also inversely associated with incident gallstone disease in this study. Validated CE, C22:6, along with C20:2 in the secondary analysis and C20:5 in NHS II cohort, are CEs of polyunsaturated fatty acids (PUFAs), docosahexaenoic acid (DHA), eicosadienoic acid and eicosapentaenoic acid (EPA), respectively. Epidemiological studies that have linked PUFA intake with reduced risk of gallstone disease[47 48] in humans and animal studies have demonstrated that administration of PUFAs, particularly DHA and EHA, can contribute to gallstone dissolution by decreasing cholesterol saturation.[49–51] Further mechanistic studies investigating the role of PUFA-CEs in humans are needed.

A prospective paediatric study and a Mendelian randomisation study in adults identified sphingolipids as metabolites linked with gallstone disease.[52 53] Sphingolipids could be potentially linked to gallstone disease through alkaline sphingomyelinase, an enzyme that acts in a bile acid–dependent manner to hydrolyse and release sphingomyelins from the canalicular membrane, which in turn can influence cholesterol crystallisation in bile and promote gallstone formation.[54 55] However, in our prospective study, the P_{FDR} value for sphingolipid metabolism was non-significant by MSEA across all four cohorts. The variation in our findings suggests a need for continued investigation in this area.

Amino acids alanine, histidine, threonine, proline and lysine were adversely associated with incident gallstone disease in the pooled primary analysis. This finding appeared to be driven by the younger NHS II cohort which had a higher proportion of premenopausal women, as these associations were not seen in the largely postmenopausal NHS and WHI cohorts. A similar finding was seen by enrichment analysis with aminoacyl tRNA biosynthesis, which was similarly significant in the pooled primary analysis and NHS II cohort but not in NHS and WHI. This may suggest that amino acids play a role in age-specific or hormone-related risk differences in women; however, further research is warranted to determine this.

Of note, gallstone incidence varied across cohorts. The lowest incidence was observed in the male HPFS cohort (7.8%), which is expected as female

sex increases the risk of gallstones by at least twofold.[56 57] This is likely related to hormonal differences as oestrogen promotes biliary cholesterol supersaturation via the oestrogen receptor-alpha (ER-α) and the G-coupled oestrogen receptor (GPER) pathways.[58–60] Gallstone incidence was generally higher in female cohorts, although to a greater extent in NHS (17.7%) compared with NHS II (12.2%). This finding could be attributed to age, a known risk factor for gallstone disease,[57 61 62] as participants in NHS were older on average. Although WHI also enrolled older female participants, gallstone incidence was comparatively lower (9.1%), likely because length of follow-up was much shorter compared with NHS (7 years vs 25 years).

The strengths of this study include the large sample size, prospective profiling of prediagnostic metabolites, robust data on covariates and risk factors, long-term follow-up and external replication in a more diverse independent cohort with a shared, well-validated metabolomic platform. However, there are limitations. Residual confounding from unmeasured confounders cannot be completely ruled out due to the observational study design. Gallstone cases were not confirmed with ultrasound or pathology but were determined by self-report. However, it is reassuring that prior validation studies have shown very high accuracy with self-report in the discovery cohorts.[24 25] Cholecystectomy was used for case definition in NHS, NHS II and HPFS, but it could be performed for indications other than gallstones. However, the vast majority of cholecystectomies are performed for gallstone-related disease.[63] In WHI, cases were defined by report of gallbladder disease, which is suboptimal, as there was no information on symptoms or cholecystectomies. Additionally, robust analysis was limited to annotated metabolites that were mostly HILIC-positive or C8-positive and does not comprehensively evaluate all unannotated or unknown metabolites that may be associated with gallstone disease. Finally, metabolite profiling was assessed at a single point in time which may not completely reflect the metabolome's dynamic nature. Further studies with repeated metabolite profiling over time are needed.

CONCLUSION

This study identifies 17 metabolites (TAGs, DAGs, CEs and plasmalogens) associated with incident gallstone disease in women with validation in an independent external cohort. Metabolite class association patterns were seen on enrichment analysis across male and female cohorts: DAGs, TAGs with ≤56 carbon atoms and ≤3 double bonds and de novo TAG biosynthesis pathways were positively associated, while CEs and plasmalogens were inversely associated. These identified metabolites, metabolite classes and pathways could not only be potentially relevant for future mechanistic studies that advance the understanding of gallstone disease but also leveraged in clinical practice as biomarkers for risk prediction, prevention and treatment.

Author affiliations
[1]Division of Gastroenterology, Massachusetts General Hospital, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, USA
[2]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA
[3]Department of Biostatistics and Epidemiology, University of Massachusetts Amherst, Amherst, Massachusetts, USA
[4]Division of Women's Health, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[5]Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[6]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Acknowledgements The authors appreciate the participants and staff of the HPFS, NHS, NHS II and WHI studies as well as the Channing Division of Network Medicine at Brigham and Women's Hospital for their valuable contributions.

Contributors EU-A and ATC designed the research and are the guarantors of the manuscript. EU-A and JH conducted analysis and RB provided statistical expertise. RB, KMR and ATC participated in acquisition of data. EU-A led manuscript drafting. KMR and ATC obtained funding. EU-A, JH, RB, KMR and ATC contributed to results interpretation and critical manuscript revision for important intellectual content. All authors approved the final version of the manuscript.

Funding The HPFS, NHS and NHS II studies are supported by the National Institutes of Health Grants U01 CA167552, UM1 CA186107, U01 CA176726, R01 CA49449 and R01 CA67262. Metabolomic analysis in the WHI was funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contract HHSN268201300008C. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C and HHSN268201600004C. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests ATC served as a consultant for Pfizer Inc. and Boehringer Ingelheim and received grants from Pfizer Inc, Zoe Ltd and Freenome.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Institutional Review Boards (IRBs) at Brigham and Women's Hospital and Harvard T. H. Chan School of Public Health (1999P003389, 1999P011114). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; internally peer reviewed.

Data availability statement Data are available upon reasonable request. Data from Nurses' Health Study (NHS), Nurses' Health Study II (NHS II), Health Professionals Follow-Up Study (HPFS and Women's Health Initiative (WHI) cohorts are not publicly available due to patient privacy concerns. Data access applications will be reviewed after submission at https://www.nurseshealthstudy.org/researchers for NHS/NHS II, https://sites.sph.harvard.edu/hpfs/for-collaborators/ for HPFS or https://www.whi.org/propose-a-paper for WHI.

ORCID iD
Jane Ha http://orcid.org/0000-0002-6175-8793

REFERENCES

1 Peery AF, Crockett SD, Murphy CC, *et al*. Burden and Cost of Gastrointestinal, Liver, and Pancreatic Diseases in the United States: Update 2021. *Gastroenterology* 2022;162:621–44.
2 Unalp-Arida A, Ruhl CE. Increasing gallstone disease prevalence and associations with gallbladder and biliary tract mortality in the US. *Hepatology* 2023;77:1882–95.
3 Li S, Guizzetti L, Ma C, *et al*. Epidemiology and outcomes of choledocholithiasis and cholangitis in the United States: trends and urban-rural variations. *BMC Gastroenterol* 2023;23:254.
4 Wadhwa V, Jobanputra Y, Garg SK, *et al*. Nationwide trends of hospital admissions for acute cholecystitis in the United States. *Gastroenterol Rep (Oxf)* 2017;5:gow015:36–42:.
5 Kratzer W, Mason RA, Kächele V. Prevalence of gallstones in sonographic surveys worldwide. *J Clin Ultrasound* 1999;27:1–7.
6 van Dijk AH, de Reuver PR, Besselink MG, *et al*. Assessment of available evidence in the management of gallbladder and bile duct stones: a systematic review of international guidelines. *HPB (Oxf)* 2017;19:297–309.
7 Di Ciaula A, Wang D-H, Portincasa P. An update on the pathogenesis of cholesterol gallstone disease. *Curr Opin Gastroenterol* 2018;34:71–80.
8 Stender S, Nordestgaard BG, Tybjaerg-Hansen A. Elevated body mass index as a causal risk factor for symptomatic gallstone disease: a Mendelian randomization study. *Hepatology* 2013;58:2133–41.
9 Lammert F, Gurusamy K, Ko CW, *et al*. Gallstones. *Nat Rev Dis Primers* 2016;2:16024.
10 Portincasa P, Di Ciaula A, Bonfrate L, *et al*. Metabolic dysfunction-associated gallstone disease: expecting more from critical care manifestations. *Intern Emerg Med* 2023;18:1897–918.
11 Cortés VA, Barrera F, Nervi F. Pathophysiological connections between gallstone disease, insulin resistance, and obesity. *Obes Rev* 2020;21:e12983.
12 Grundy SM. Cholesterol gallstones: a fellow traveler with metabolic syndrome? *Am J Clin Nutr* 2004;80:1–2.
13 Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov* 2016;15:473–84.
14 Talmor-Barkan Y, Bar N, Shaul AA, *et al*. Metabolomic and microbiome profiling reveals personalized risk factors for coronary artery disease. *Nat Med* 2022;28:295–302.
15 Cirulli ET, Guo L, Leon Swisher C, *et al*. Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. *Cell Metab* 2019;29:488–500.
16 Rangel-Huerta OD, Pastor-Villaescusa B, Gil A. Are we close to defining a metabolomic signature of human obesity? A systematic review of metabolomics studies. *Metabolomics (Los Angel)* 2019;15:93.
17 Bao Y, Bertoia ML, Lenart EB, *et al*. Origin, Methods, and Evolution of the Three Nurses' Health Studies. *Am J Public Health* 2016;106:1573–81.
18 Townsend MK, Clish CB, Kraft P, *et al*. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. *Clin Chem* 2013;59:1657–67.
19 Circulation. Metabolic predictors of incident coronary heart disease in women. Available: https://www.ahajournals.org/doi/full/10.1161/CIRCULATIONAHA.117.029468 [Accessed 4 Dec 2023].
20 ScienceDirect. Design of the women's health initiative clinical trial and observational study. Available: https://www.sciencedirect.com/science/article/abs/pii/S0197245697000780 [Accessed 4 Dec 2024].
21 Writing Group for the Women's Health Initiative Investigators. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: principal Results From the Women's Health Initiative Randomized Controlled Trial. *JAMA* 2002;288:321–33.
22 Rimm EB, Giovannucci EL, Willett WC, *et al*. Prospective study of alcohol consumption and risk of coronary disease in men. *The Lancet* 1991;338:464–8.
23 Platz EA, Clinton SK, Giovannucci E. Association between plasma cholesterol and prostate cancer in the PSA era. *Int J Cancer* 2008;123:1693–8.
24 Leitzmann MF, Giovannucci EL, Rimm EB, *et al*. The relation of physical activity to risk for symptomatic gallstone disease in men. *Ann Intern Med* 1998;128:417–25.
25 Maclure KM, Hayes KC, Colditz GA, *et al*. Weight, diet, and the risk of symptomatic gallstones in middle-aged women. *N Engl J Med* 1989;321:563–9.
26 McCullough ML, Feskanich D, Stampfer MJ, *et al*. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am J Clin Nutr* 2002;76:1261–71.
27 Chiuve SE, Fung TT, Rimm EB, *et al*. Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr* 2012;142:1009–18.
28 Wei R, Wang J, Su M, *et al*. Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci Rep* 2018;8:663.
29 Subramanian A, Tamayo P, Mootha VK, *et al*. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
30 Chen L, Yang H, Li H, *et al*. Insights into modifiable risk factors of cholelithiasis: a Mendelian randomization study. *Hepatology* 2022;75:785–96.
31 Stender S, Frikke-Schmidt R, Benn M, *et al*. Low-density lipoprotein cholesterol and risk of gallstone disease: a Mendelian randomization study and meta-analyses. *J Hepatol* 2013;58:126–33.
32 Smelt AHM. Triglycerides and gallstone formation. *Clin Chim Acta* 2010;411:1625–31.
33 Ferkingstad E, Oddsson A, Gretarsdottir S, *et al*. Genome-wide association meta-analysis yields 20 loci associated with gallstone disease. *Nat Commun* 2018;9:5101.
34 Joshi AD, Andersson C, Buch S, *et al*. Four Susceptibility Loci for Gallstone Disease Identified in a Meta-analysis of Genome-Wide Association Studies. *Gastroenterology* 2016;151:351–63.
35 Rodriguez S, Gaunt TR, Guo Y, *et al*. Lipids, obesity and gallbladder disease in women: insights from genetic studies using the cardiovascular gene-centric 50K SNP array. *Eur J Hum Genet* 2016;24:106–12.
36 Hu H, Shao W, Liu Q, *et al*. Gut microbiota promotes cholesterol gallstone formation by modulating bile acid composition and biliary cholesterol secretion. *Nat Commun* 2022;13:252.
37 Yetukuri L, Katajamaa M, Medina-Gomez G, *et al*. Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. *BMC Syst Biol* 2007;1:12.
38 van Ginneken V, Verhey E, Poelmann R, *et al*. Metabolomics (liver and blood profiling) in a mouse model in response to fasting: a study of hepatic steatosis. *Biochim et Biophys Acta (BBA) - Mol Cell Biol Lipids* 2007;1771:1263–70.
39 Gorden DL, Ivanova PT, Myers DS, *et al*. Increased diacylglycerols characterize hepatic lipid changes in progression of human nonalcoholic fatty liver disease; comparison to a murine model. *PLoS One* 2011;6:e22775.
40 Chang Y, Noh Y-H, Suh B-S, *et al*. Bidirectional Association between Nonalcoholic Fatty Liver Disease and Gallstone Disease: a Cohort Study. *J Clin Med* 2018;7:458.
41 Arrese M, Cortés V, Barrera F, *et al*. Nonalcoholic fatty liver disease, cholesterol gallstones, and cholecystectomy: new insights on a complex relationship. *Curr Opin Gastroenterol* 2018;34:90–6.
42 Bozelli JC, Azher S, Epand RM. Plasmalogens and Chronic Inflammatory Diseases. *Front Physiol* 2021;12:730829.
43 Ding L, Zhang L, Shi H, *et al*. EPA-enriched ethanolamine plasmalogen alleviates atherosclerosis via mediating bile acids metabolism. *J Funct Foods* 2020;66:103824.
44 Ding L, Zhang L-Y, Shi H-H, *et al*. Eicosapentaenoic Acid-Enriched Phosphoethanolamine Plasmalogens Alleviated Atherosclerosis by Remodeling Gut Microbiota to Regulate Bile Acid Metabolism in LDLR$^{-/-}$ Mice. *J Agric Food Chem* 2020;68:5339–48.
45 Wu L, Wang Y, Zhu S, *et al*. Changes in plasma bile acids are associated with gallbladder stones and polyps. *BMC Gastroenterol* 2020;20:363.
46 Wei J, Chen T, Liu Y, *et al*. Targeted bile acids metabolomics in cholesterol gallbladder polyps and gallstones: from analytical method development towards application to clinical samples. *J Pharm Anal* 2023;13:1080–7.
47 Sun Q, Gao N, Xia W. Association between omega-3/6 fatty acids and cholelithiasis: a mendelian randomization study. *Front Nutr* 2022;9:964805.
48 Tsai C-J, Leitzmann MF, Willett WC, *et al*. The effect of long-term intake of cis unsaturated fats on the risk for gallstone disease in men: a prospective cohort study. *Ann Intern Med* 2004;141:514–22.
49 Kim JK, Cho SM, Kang SH, *et al*. N-3 polyunsaturated fatty acid attenuates cholesterol gallstones by suppressing mucin production with a high cholesterol diet in mice. *J Gastroenterol Hepatol* 2012;27:1745–51.

50 Jang SI, Fang S, Kim KP, *et al*. Combination treatment with n-3 polyunsaturated fatty acids and ursodeoxycholic acid dissolves cholesterol gallstones in mice. *Sci Rep* 2019;9:12740.
51 Cho S-M, Park J-A, Kim N-H, *et al*. Effect of eicosapentaenoic acid on cholesterol gallstone formation in C57BL/6J mice. *Mol Med Rep* 2015;11:362–6.
52 Mi J, Jiang L, Liu Z, *et al*. Identification of blood metabolites linked to the risk of cholelithiasis: a comprehensive Mendelian randomization study. *Hepatol Int* 2022;16:1484–93.
53 Zdanowicz K, Bobrus-Chcociej A, Pogodzinska K, *et al*. Analysis of Sphingolipids in Pediatric Patients with Cholelithiasis-A Preliminary Study. *J Clin Med* 2022;11:5613.
54 Duan R-D. Alkaline sphingomyelinase (NPP7) in hepatobiliary diseases: a field that needs to be closely studied. *World J Hepatol* 2018;10:246–53.
55 Duan R-D. Alkaline sphingomyelinase: an old enzyme with novel implications. *Biochim Biophys Acta* 2006;1761:281–91.
56 Völzke H, Baumeister SE, Alte D, *et al*. Independent risk factors for gallstone formation in a region with high cholelithiasis prevalence. *Digestion* 2005;71:97–105.
57 Maurer KR, Everhart JE, Ezzati TM, *et al*. Prevalence of gallstone disease in Hispanic populations in the United States. *Gastroenterology* 1989;96:487–92.
58 Wang HH, Liu M, Clegg DJ, *et al*. New insights into the molecular mechanisms underlying effects of estrogen on cholesterol gallstone formation. *Biochim Biophys Acta* 2009;1791:1037–47.
59 Wang HH, de Bari O, Arnatt CK, *et al*. Activation of Estrogen Receptor G Protein-Coupled Receptor 30 Enhances Cholesterol Cholelithogenesis in Female Mice. *Hepatology* 2020;72:2077–89.
60 de Bari O, Wang TY, Liu M, *et al*. Estrogen induces two distinct cholesterol crystallization pathways by activating ERα and GPR30 in female mice. *J Lipid Res* 2015;56:1691–700.
61 Aerts R, Penninckx F. The burden of gallstone disease in Europe. *Aliment Pharmacol Ther* 2003;18:49–53.
62 Einarsson K, Nilsell K, Leijd B, *et al*. Influence of age on secretion of cholesterol and synthesis of bile acids by the liver. *N Engl J Med* 1985;313:277–82.
63 Alli VV, Yang J, Xu J, *et al*. Nineteen-year trends in incidence and indications for laparoscopic cholecystectomy: the NY State experience. *Surg Endosc* 2017;31:1651–8.