Supplementary Figure 1 Secondary Analysis–Volcano Plot of WHI Incident Gallstone Disease Risk Estimates for Metabolites(N=284) not Tested in the Primary Analysis

Supplementary Table 1 Gallstone-Free Participant Numbers from Nested Case Control Sub-Studies in NHS, NHS II & HPFS

STUDY	NHS	NHS II	HPFS
Amyotrophic Lateral Sclerosis	97	0	74
Breast Cancer	1557	1958	0
Colon Cancer	652	0	512
Diabetes (type II)	1218	0	0
Glaucoma (exfoliative)	273	0	61
Glaucoma (primary open angle)	594	138	284
Inflammatory Bowel Disease	268	114	0
Ovarian Cancer	434	105	0
Parkinson's Disease	493	0	610
Prostate Cancer	0	0	637
Racial Differences	169	0	0
Rheumatoid Arthritis	447	244	0
Stress	0	206	0
Stroke	590	110	0
Total	**6792**	**3168**	**2178**

Supplementary Table 2 Description of Metabolomic Profiling Methods

HILIC-POSITIVE

Hydrophilic interaction liquid chromatography (HILIC) analyses of water-soluble metabolites in the positive ionization mode (HILIC-positive) were performed with an LC–MS system comprised of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.; Marlborough, MA) coupled to a Q Exactive mass spectrometer (Thermo Fisher Scientific; Waltham, MA). Metabolites were extracted from plasma (10 µL) using 90 µL of acetonitrile/methanol/formic acid (74.9:24.9:0.2 v/v/v) containing stable isotope-labeled internal standards (valine-d8, Sigma-Aldrich; St. Louis, MO; and phenylalanine-d8, Cambridge Isotope Laboratories; Andover, MA). The samples were centrifuged (10 min, 9000 × *g*, 4 °C), and the supernatants were injected directly onto a 150 × 2 mm, 3 µm Atlantis HILIC column (Waters; Milford, MA). The column was eluted isocratically at a flow rate of 250 µL/min with 5% mobile phase A (10 mM ammonium formate and 0.1% formic acid in

water) for 0.5 min followed by a linear gradient to 40% mobile phase B (acetonitrile with 0.1% formic acid) over 10 min. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over 70–800 *m/z* at 70,000 resolution and 3 Hz data acquisition rate. Other MS settings were: sheath gas 40, sweep gas 2, spray voltage 3.5 kV, capillary temperature 350 °C, S-lens RF 40, heater temperature 300 °C, microscans 1, automatic gain control target 1e6, and maximum ion time 250 ms.

HILIC-NEGATIVE

HILIC analyses of water-soluble metabolites in the negative ionization mode (HILIC-negative) were performed with an LC-MS system composed of an AQUITY UPLC system (Waters and a 5500 QTRAP mass spectrometer [SCIEX]). Plasma samples (30 µL) were prepared via protein precipitation with the addition of 4 volumes of 80% methanol containing inosine-15N4, thymine-d4, and glycocholate-d4 internal standards (Cambridge Isotope Laboratories). The samples were centrifuged (10 minutes, 9000*g*, 4°C), and the supernatants were injected directly onto a 150×2.0 mm Luna NH2 column (Phenomenex). The column was eluted at a flow rate of 400 µL/min with initial conditions of 10% mobile phase A (20 mmol/L ammonium acetate and 20 mmol/L ammonium hydroxide in water) and 90% mobile phase B (10 mmol/L ammonium hydroxide in 75:25 v/v acetonitrile/methanol) followed by a 10-minute linear gradient to 100% mobile phase A. MS analyses were performed using electrospray ionization and selective multiple reaction monitoring scans in the negative ion mode. To create the method, declustering potentials and collision energies were optimized for each metabolite by infusion of reference standards. The ion spray voltage was –4.5 kV and the source temperature was 500°C.

C8-POSITIVE

Analyses of polar and non-polar plasma lipids in positive ion mode (C8-positive) was conducted using a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.; Marlborough, MA). Lipids were extracted from plasma (10 µL) using 190 µL of isopropanol containing 1,2- didodecanoyl-sn-glycero-3-phosphocholine as an internal standard (Avanti Polar Lipids; Alabaster, AL). After centrifugation, supernatants were injected directly onto a 100 × 2.1 mm, 1.7 µm ACQUITY BEH C8 column (Waters; Milford, MA). The column was eluted isocratically with 80% mobile phase A (95:5:0.1 vol/vol/vol 10 mM ammonium acetate/methanol/formic acid) for 1 min followed by a linear gradient to 80% mobile-phase B (99.9:0.1 vol/vol methanol/formic acid) over 2 min, a linear gradient to 100% mobile phase B over 7 min, then 3 min at 100% mobile-phase B. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over 200–1100 *m/z* at 70,000 resolution and 3 Hz data acquisition rate. Other MS settings were: sheath gas 50, in source CID 5 eV, sweep gas 5, spray voltage 3 kV, capillary temperature 300 °C, S-lens RF 60, heater temperature 300 °C, microscans 1, automatic gain control target 1e6, and maximum ion time 100 ms. Lipid identities were determined based on comparison with reference plasma extracts and were denoted by the total number of carbons in the lipid acyl chain(s) and total number of double bonds in the lipid acyl chain(s).

C18-NEGATIVE

Analyses of free fatty acids and bile acids in the negative ion mode (C18-negative) was conducted using an LC-MS system composed of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp) coupled to a Q Exactive hybrid quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific). Plasma samples (30 µL) were extracted using 90 µL of methanol containing PGE2-d4 (Cayman Chemical Co) and centrifuged (10 minutes, 9000*g*, 4°C). The samples were injected onto a 150×2 mm ACQUITY T3 column (Waters; Milford, MA). The column was eluted isocratically at a flow rate of 400 µL/min with 60% mobile phase A (0.1% formic acid in water) for 4 minutes followed by a linear gradient to 100% mobile phase B (acetonitrile with 0.1% formic acid) over 8 minutes. MS analyses were performed in the negative ion mode using electrospray ionization, full scan MS acquisition over 200 to 550 *m/z*, and a resolution setting of 70 000.

Metabolite identities were confirmed using authentic reference standards. Other MS settings were as follows: sheath gas 45, sweep gas 5, spray voltage –3.5 kV, capillary temperature 320°C, S-lens RF 60, heater temperature 300°C, microscans 1, automatic gain control target 1e6, and maximum ion time 250 ms.

Supplementary Table 3 Risk Estimates for Lifestyle & Demographic Covariates Adjusted for in the Multivariable Model

COVARIATE	NHS & NHS II [aOR (95% CI, P)]	WHI [aOR (95% CI, P_{FDR})]	HPFS [aOR (95% CI, P_{FDR})]
Age (years)	0.99 (0.98–1.00, 0.003)	0.98 (0.96–1.01, 0.197)	1.01 (0.99–1.03, 0.415)
BMI	1.09 (1.08–1.10, <0.001)	1.05 (1.02–1.08, <0.001)	1.10 (1.05–1.15, <0.001)
Activity	1.00 (1.00–1.00, 0.499)	0.99 (0.98–1.01, 0.329)	1.00 (0.99–1.01, 0.952)
AHEI Dietary Score	0.99 (0.99–1.00, 0.08)	1.01 (0.99–1.02, 0.605)	0.98 (0.97–1.00, 0.024)
Alcohol	0.99 (0.98–0.99, <0.001)	0.99 (0.97–1.01, 0.455)	0.99 (0.98–1.00, 0.135)
Current Tobacco Use	1.01 (0.54–1.80, 0.066)	1.37 (1.14–1.65, 0.965)	1.06 (0.90–1.25, 0.479)
Parity	1.25 (1.03–1.52, 0.026)	1.39 (0.80–2.64, 0.276)	–
Menopausal Hormone Therapy Use	1.41 (1.22–1.64, <0.001)	1.51 (1.05–2.19, 0.028)	–

Supplementary Table 4 Metabolite Sets

Metabolite Structure Sets	N (Metabolites)
Amines, amino acids, and derivatives	33

Carnitines	20
Purines and purine derivatives	6
Cholesterol Esters	9
Pyrimidine and pyrimidine derivatives	1
Nucleosides, nucleotides, and analogues	5
Sphingolipids	4
Ceramides	4
Triacylglycerols (≤56 C and ≤3 double bonds)	38
Triacylglycerols (>56 C or >3 double bonds)	21
Diacylglycerols	11
Phosphatidylcholines	20
Phosphatidylethanolamines	10
Plasmalogens	22
Lysophosphatidylcholines	3
Lysophophatidylethanolamines	5
Phosphatidylserine	1

Metabolite Pathway Sets	**N (Metabolites)**
beta-Alanine metabolism	3
Pyrimidine metabolism	4
Pantothenate and CoA biosynthesis	3
Glycine, serine and threonine metabolism	9
Methionine metabolism	2
Porphyrin metabolism	4
Biosynthesis of unsaturated fatty acids	2
Fatty acid metabolism	1
Steroid hormone biosynthesis	3
Arginine and proline metabolism	8
Primary bile acid biosynthesis	1
Glutathione metabolism	2
Alanine, aspartate and glutamate metabolism	3
Glutamate metabolism	4
Purine metabolism	3

Lysine degradation	6
Phosphonate and Phosphinate metabolism	1
Thiamine metabolism	3
Aminoacyl-tRNA biosynthesis	16
Glyoxylate and dicarboxylate metabolism	3
Tyrosine metabolism	2
Phenylalanine metabolism	4
Phenylalanine, tyrosine and tryptophan biosynthesis	3
Glucosinolate biosynthesis	8
2-Oxocarboxylic acid metabolism	10
Tropane, piperidine and pyridine alkaloid biosynthesis	6
Selenocompound metabolism	1
Cysteine and methionine metabolism	3
Taurine and hypotaurine metabolism	1
Valine, leucine and isoleucine biosynthesis	4
Cyanoamino acid metabolism	4
Valine, leucine and isoleucine degradation	3
Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid	4
Histidine metabolism	3
Biosynthesis of alkaloids derived from histidine and purine	3
Lysine biosynthesis	1
Biotin metabolism	1
Glycerophospholipid metabolism	4
Sphingolipid metabolism	7
Nucleotide metabolism	1
Amino sugar metabolism	1
Aspartate metabolism	1
Nicotinate and nicotinamide metabolism	3
Arginine biosynthesis	3
Biosynthesis of alkaloids derived from shikimate pathway	2
Methane metabolism	1
Tryptophan metabolism	1
Indole alkaloid biosynthesis	1

Biosynthesis of phenylpropanoids	1
Starch and sucrose metabolism	1
Caffeine metabolism	4
Bile secretion	1
Glycerolipid metabolism	31
De novo triacylglycerol biosynthesis	70
Steroid biosynthesis	5
Diacylglycerol signaling pathways	6
Phosphatidylinositol phosphate metabolism	1
Phosphatidylcholine/Phosphatidylethanolamine biosynthesis	30
Arachidonic acid metabolism	2
Linoleic acid metabolism	1
alpha-Linolenic acid metabolism	1

Supplementary Table 5 Risk Estimates for Individual Metabolite Associations with Incident Gallstone Disease in the Primary Analysis in NHS I & NHS II (external file attached)