SUPPLEMENTARY MATERIALS

Supplementary Methods

Biological reproducibility of bile acid concentrations

Prior to the present analyses on BAs and colon cancer risk, we evaluated the biological reproducibility of the BA metabolites using fasting plasma samples of participants of the EPIC-Heidelberg cohort (that was initially started in 1994 as part of the Europe-wide EPIC cohort). These participants were re-invited to the EPIC-Heidelberg study center (German Cancer Research Center, Heidelberg, Germany) for a sub-study on diet, physical activity, and body composition between 2010 and 2013 (1). They provided blood samples at two examinations, which were carried out one year apart. Out of the 799 participants of this sub-study, we randomly selected 50 individuals (50 % women, mean age 65.0 years, free of cancer and major cardiovascular diseases) for our reproducibility study. BAs were measured by the same validated method as the BAs in our main study (2) at the Health Research Institute Hospital La Fe (Valencia, Spain).

To test the biological reproducibility of BA concentrations over one year, we obtained Spearman's correlation coefficients (see Supplementary Table 1 below). The majority of BA metabolites (CA, CDCA, TCA, TCDCA, DCA, GDCA, TDCA, GUDCA, and GHCA) showed good biological reproducibility over one year, with Spearman's coefficients of $\rho>0.5$ to 0.78. Modest reproducibility ($\rho>0.3$-0.5) was observed for GCA, GLCA, TUDCA, HCA, and TaMCA. Reproducibility was poor ($\rho<0.3$) for UDCA, and THCA was not detectable in a majority of samples.

Cross-sectional associations between bile acids and covariates among the controls

Associations between bile acid concentrations and covariates were evaluated by Generalized Linear Models adjusted for age, sex, fasting status, smoking status, education level, physical activity level, BMI, height, and consumption of alcohol, fiber, red and processed meat, fat, and energy. Multivariable adjusted geometric means (95 % confidence intervals) of bile acid levels across covariate strata were calculated by the least-squares means method. P values for linear trend were obtained modelling covariate strata as continuous variables.
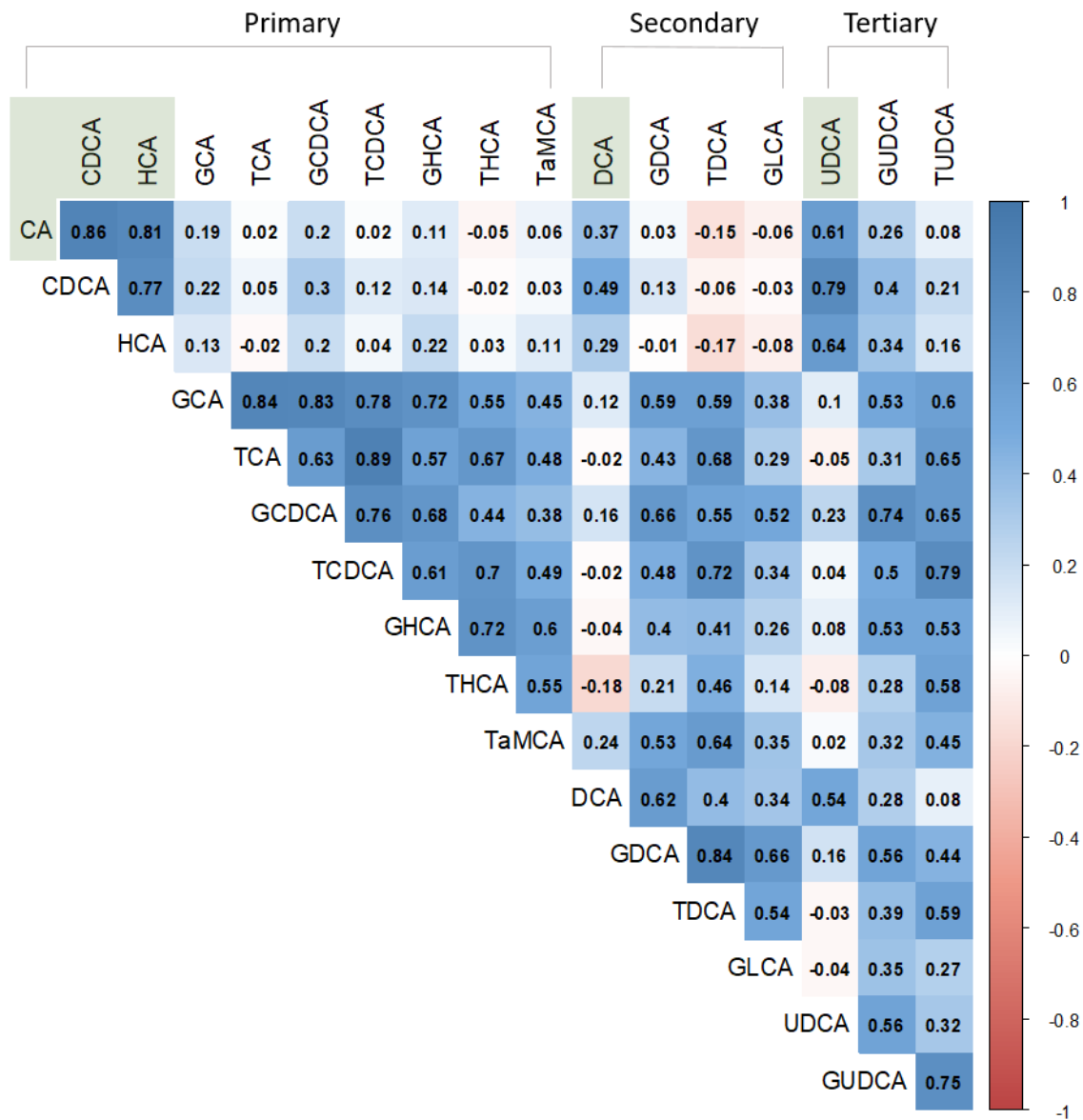
Correction of odds ratios from logistic regression analyses for measurement error

As described by Frost and Thompson (3), we used the data from the reproducibility study outlined above to correct odds ratios and 95 % confidence intervals from multivariable logistic regression analyses for regression dilution due to measurement error, using product-moment correlation coefficients. As we did not have information on the reproducibility of covariates, we assumed that these would be stable over time. Our correction for measurement error is further based on the assumption that the results of our one-year reproducibility study among 50 participants of the EPIC-Heidelberg cohort are representative and generalizable.

References

1. Kühn T, Sookthai D, Rolle-Kampczyk U, Otto W, von Bergen M, Kaaks R, et al. Mid- and long-term correlations of plasma metabolite concentrations measured by a targeted metabolomics approach. Metabolomics. 2016;12(12):184.

2. Garcia-Canaveras JC, Donato MT, Castell JV, Lahoz A. Targeted profiling of circulating and hepatic bile acids in human, mouse, and rat using a UPLC-MRM-MS-validated method. Journal of lipid research. 2012;53(10):2231-41.

3. Frost C, Thompson SG. Correcting for regression dilution bias: comparison of methods for a single predictor variable. JSTOR. 2000;163(2):173-89.

Supplementary Figure

Supplementary Figure 1. Inter-correlations between bile acid metabolites

Spearman's correlations between bile acid metabolites among the controls (n=569); bile acids marked in green are unconjugated metabolites, all others are conjugated metabolites. CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);

Supplementary Tables

Supplementary Table 1 One-year biological reproducibility of bile acid concentrations in a subsample of the EPIC-Heidelberg cohort (n=50)

Bile acid metabolite	N missing values	Spearman's ρ*
Primary bile acids, unconjugated		
Cholic Acid (CA)	0	0.59
Chenodeoxycholic Acid (CDCA)	0	0.66
Hyocholic Acid (HCA)	23	0.45
Primary bile acids, conjugated		
Glycocholic Acid (GCA)	0	0.49
Taurocholic Acid (TCA)	0	0.59
Glycochenodeoxycholic Acid (GCDCA)	0	0.55
Taurochenodeoxycholic Acid (TCDCA)	0	0.60
Glycohyocholic Acid (GHCA)	1	0.53
Taurohyocholic Acid (THCA)	40	-
Tauro-Alpha-Muricholic Acid (TaMCA)	11	0.44
Secondary bile acids, unconjugated		
Deoxycholic Acid (DCA)	1	0.70
Secondary bile acids, conjugated		
Glycodeoxycholic Acid (GDCA)	1	0.78
Taurodeoxycholic Acid (TDCA)	4	0.67
Glycolithocholic Acid (GLCA)	3	0.31
Tertiary bile acids, unconjugated		
Ursodeoxycholic Acid (UDCA)	15	0.01
Tertiary bile acids, conjugated		
Glycoursodeoxycholic Acid (GUDCA)	0	0.55
Tauroursodeoxycholic Acid (TUDCA)	19	0.30

*All correlations over one year were statistically significant at p-values <0.05, except for the correlations between plasma levels of UDCA;

Supplementary Table 2 Number of study participants with bile acid concentrations below the limit of detection

Bile acid metabolite	Classification	N (%) values below the limit of detection among cases / controls
Cholic Acid (CA)	Primary, unconjugated	3/6 (0.5%/1.1%)
Chenodeoxycholic Acid (CDCA)	Primary, unconjugated	1/4 (0.2%/0.7%)
Hyocholic Acid (HCA)	Primary, unconjugated	74/79 (13%/13.9%)
Glycocholic Acid (GCA)	Primary, conjugated	0/0
Taurocholic Acid (TCA)	Primary, conjugated	6/7 (1.1%/1.2%)
Glycochenodeoxycholic Acid (GCDCA)	Primary, conjugated	0/0
Taurochenodeoxycholic (TCDCA)	Primary, conjugated	2/1 (0.4%/0.2%)
Glycohyocholic Acid (GHCA)	Primary, conjugated	9/5 (1.6%/0.9%)
Taurohyocholic Acid (THCA)	Primary, conjugated	307/286 (54%/50.3%)
Tauro-Alpha-Muricholic Acid (TaMCA)	Primary, conjugated	104/80 (18.3%/14.1%)
Deoxycholic Acid (DCA)	Secondary, unconjugated	3/4 (0.5%/0.7%)
Glycodeoxycholic Acid (GDCA)	Secondary, conjugated	2/4 (0.4%/0.7%)
Taurodeoxycholic Acid (TDCA)	Secondary, conjugated	21/16 (3.7%/2.8%)
Glycolithocholic Acid (GLCA)	Secondary, conjugated	10/10 (1.8%/1.8%)
Ursodeoxycholic Acid (UDCA)	Tertiary, unconjugated	87/84 (15.3%/14.8%)
Glycoursodeoxycholic Acid (GUDCA)	Tertiary, conjugated	3/3 (0.5%/0.5%)
Tauroursodeoxycholic Acid (TUDCA)	Tertiary, conjugated	163/154 (28.6%/27.1%)

Supplementary Table 3 Geometric means (95% confidence intervals) of plasma bile acid concentrations, stratified by sex

Bile Acid*	Sex	Frequency	Geometric Mean (95% CI)	P†
CA	Male	214 (37.3)	97.2 (52.5,180.1)	0.007
	Female	360 (62.7)	54.4 (30.2,98.2)	
CDCA	Male	214 (37.3)	189.3 (108.9,329.2)	0.003
	Female	360 (62.7)	108.0 (63.6,183.3)	
GCA	Male	214 (37.3)	189.1 (126.7,282.2)	<0.001
	Female	360 (62.7)	109.8 (74.9,161.1)	
TCA	Male	214 (37.3)	19.8 (12.3,32.0)	0.08
	Female	360 (62.7)	14.8 (9.4,23.4)	
GCDCA	Male	214 (37.3)	760.9 (529.8,1092.8)	<0.001
	Female	360 (62.7)	498.9 (352.8,705.3)	
TCDCA	Male	214 (37.3)	74.7 (48.9,113.9)	0.15
	Female	360 (62.7)	60.4 (40.3,90.5)	
DCA	Male	214 (37.3)	172.9 (106.6,280.6)	0.32
	Female	360 (62.7)	146.7 (92.3,233.1)	
GDCA	Male	214 (37.3)	154.9 (99.0,242.4)	0.18
	Female	360 (62.7)	125.8 (81.9,193.0)	
TDCA	Male	214 (37.3)	19.8 (12.3,31.8)	0.85
	Female	360 (62.7)	20.4 (12.9,32.2)	
GLCA	Male	214 (37.3)	17.8 (11.8,26.9)	0.40
	Female	360 (62.7)	15.9 (10.7,23.5)	
UDCA	Male	214 (37.3)	19.1 (11.0,33.3)	0.11
	Female	360 (62.7)	14.1 (8.3,24.0)	
GUDCA	Male	214 (37.3)	85.1 (57.0,127.0)	0.03
	Female	360 (62.7)	62.9 (42.9,92.2)	
TUDCA	Male	214 (37.3)	3.7 (2.7,5.1)	0.46
	Female	360 (62.7)	3.4 (2.5,4.7)	
HCA	Male	214 (37.3)	8.1 (5.1,12.9)	0.002
	Female	360 (62.7)	5.0 (3.2,7.8)	
GHCA	Male	214 (37.3)	13.3 (10.2,17.3)	<0.001
	Female	360 (62.7)	9.6 (7.5,12.4)	
THCA	Male	214 (37.3)	2.6 (2.1,3.3)	0.01
	Female	360 (62.7)	2.1 (1.7,2.7)	
TaMCA	Male	214 (37.3)	4.9 (3.4,7.0)	0.73
	Female	360 (62.7)	5.1 (3.6,7.2)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, fasting status, smoking status, education level, physical activity level, BMI, height, and consumption of alcohol, fiber, red and processed meat, fat, and energy;

*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);

†Two-sided P value for difference from Generalized Linear Models;

Supplementary Table 4 Geometric means (95% confidence intervals) of plasma bile acid concentrations across strata of fasting duration

Bile Acid*	Fasting Duration	Frequency	Geometric Mean (95% CI)	P†
CA	>6 hours	108 (18.8)	66.3 (35.7,123.1)	0.70
	3-6 hours	91 (15.9)	75.1 (40.0,141.2)	
	<3 hours	375 (65.3)	77.2 (43.6,136.7)	
CDCA	>6 hours	108 (18.8)	149.8 (86.0,260.9)	0.86
	3-6 hours	91 (15.9)	142.4 (80.8,250.9)	
	<3 hours	375 (65.3)	137.1 (82.1,228.8)	
GCA	>6 hours	108 (18.8)	189.1 (126.5,282.6)	<0.001
	3-6 hours	91 (15.9)	194.2 (128.9,292.7)	
	<3 hours	375 (65.3)	81.5 (56.2,118.1)	
TCA	>6 hours	108 (18.8)	22.4 (13.8,36.1)	<0.001
	3-6 hours	91 (15.9)	23.6 (14.4,38.5)	
	<3 hours	375 (65.3)	9.5 (6.1,14.9)	
GCDCA	>6 hours	108 (18.8)	863.5 (600.6,1241.6)	<0.001
	3-6 hours	91 (15.9)	772.6 (533.4,1119.3)	
	<3 hours	375 (65.3)	350.5 (250.7,490.2)	
TCDCA	>6 hours	108 (18.8)	93.1 (61.0,142.3)	<0.001
	3-6 hours	91 (15.9)	87.0 (56.4,134.1)	
	<3 hours	375 (65.3)	37.4 (25.3,55.3)	
DCA	>6 hours	108 (18.8)	195.0 (120.0,316.9)	0.01
	3-6 hours	91 (15.9)	159.2 (97.0,261.3)	
	<3 hours	375 (65.3)	130.1 (83.1,203.7)	
GDCA	>6 hours	108 (18.8)	212.9 (135.9,333.5)	<0.001
	3-6 hours	91 (15.9)	178.3 (112.7,281.9)	
	<3 hours	375 (65.3)	71.7 (47.3,108.5)	
TDCA	>6 hours	108 (18.8)	31.5 (19.5,50.8)	<0.001
	3-6 hours	91 (15.9)	25.8 (15.9,42.0)	
	<3 hours	375 (65.3)	10.0 (6.4,15.5)	
GLCA	>6 hours	108 (18.8)	27.9 (18.5,42.0)	<0.001
	3-6 hours	91 (15.9)	19.2 (12.6,29.2)	
	<3 hours	375 (65.3)	8.9 (6.1,13.0)	
UDCA	>6 hours	108 (18.8)	15.3 (8.8,26.7)	0.79
	3-6 hours	91 (15.9)	17.2 (9.8,30.4)	
	<3 hours	375 (65.3)	16.9 (10.1,28.2)	
GUDCA	>6 hours	108 (18.8)	87.9 (58.8,131.3)	<0.001
	3-6 hours	91 (15.9)	89.3 (59.3,134.6)	
	<3 hours	375 (65.3)	49.9 (34.4,72.2)	
TUDCA	>6 hours	108 (18.8)	4.2 (3.0,5.8)	<0.001
	3-6 hours	91 (15.9)	4.2 (3.0,5.8)	
	<3 hours	375 (65.3)	2.6 (1.9,3.5)	
HCA	>6 hours	108 (18.8)	6.7 (4.2,10.7)	0.04
	3-6 hours	91 (15.9)	5.2 (3.3,8.4)	
	<3 hours	375 (65.3)	7.4 (4.8,11.4)	

Supplementary Table 4 (continued) Geometric means (95% confidence intervals) of plasma bile acid concentrations across strata of fasting duration

Bile Acid*	Fasting Duration	Frequency	Geometric Mean (95% CI)	P†
GHCA	>6 hours	108 (18.8)	13.3 (10.2,17.3)	<0.001
	3-6 hours	91 (15.9)	12.5 (9.5,16.4)	
	<3 hours	375 (65.3)	8.7 (6.8,11.1)	
THCA	>6 hours	108 (18.8)	2.8 (2.2,3.6)	<0.001
	3-6 hours	91 (15.9)	2.4 (1.9,3.1)	
	<3 hours	375 (65.3)	1.9 (1.5,2.4)	
TaMCA	>6 hours	108 (18.8)	6.3 (4.4,9.0)	<0.001
	3-6 hours	91 (15.9)	5.8 (4.0,8.3)	
	<3 hours	375 (65.3)	3.4 (2.5,4.8)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, sex, smoking status, education level, physical activity level, BMI, height, and consumption of alcohol, fiber, red and processed meat, fat, and energy;
*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);
†Two-sided P value for linear trend from Generalized Linear Models;

Supplementary Table 5 Geometric means (95% confidence intervals) of plasma bile acid concentrations across strata of BMI

Bile Acid*	BMI	Frequency	Geometric Mean (95% CI)	P†
CA	Normal Weight (< 25)	200 (34.8)	58.3 (31.8,106.9)	0.02
	Overweight (25-29.9)	281 (49)	79.9 (44.9,142.1)	
	Obese (≥30)	93 (16.2)	99.8 (52.0,191.3)	
CDCA	Normal Weight (< 25)	200 (34.8)	118.0 (68.6,203.2)	0.003
	Overweight (25-29.9)	281 (49)	152.9 (91.2,256.3)	
	Obese (≥30)	93 (16.2)	216.0 (120.5,387.2)	
GCA	Normal Weight (< 25)	200 (34.8)	145.0 (98.0,214.4)	0.85
	Overweight (25-29.9)	281 (49)	143.7 (99.1,208.5)	
	Obese (≥30)	93 (16.2)	153.6 (100.9,233.9)	
TCA	Normal Weight (< 25)	200 (34.8)	17.3 (10.8,27.5)	0.84
	Overweight (25-29.9)	281 (49)	17.3 (11.1,27.0)	
	Obese (≥30)	93 (16.2)	15.9 (9.6,26.3)	
GCDCA	Normal Weight (< 25)	200 (34.8)	632.8 (444.6,900.6)	0.27
	Overweight (25-29.9)	281 (49)	601.6 (430.1,841.5)	
	Obese (≥30)	93 (16.2)	715.1 (489.4,1044.8)	
TCDCA	Normal Weight (< 25)	200 (34.8)	70.9 (46.9,107.1)	0.73
	Overweight (25-29.9)	281 (49)	65.7 (44.4,97.3)	
	Obese (≥30)	93 (16.2)	68.6 (44.0,106.9)	
DCA	Normal Weight (< 25)	200 (34.8)	131.3 (81.7,211.0)	0.02
	Overweight (25-29.9)	281 (49)	174.6 (111.2,274.0)	
	Obese (≥30)	93 (16.2)	181.7 (109.2,302.5)	
GDCA	Normal Weight (< 25)	200 (34.8)	132.8 (85.7,205.8)	0.53
	Overweight (25-29.9)	281 (49)	142.6 (94.0,216.3)	
	Obese (≥30)	93 (16.2)	156.1 (97.5,249.9)	
TDCA	Normal Weight (< 25)	200 (34.8)	19.4 (12.2,30.9)	0.88
	Overweight (25-29.9)	281 (49)	20.5 (13.2,31.9)	
	Obese (≥30)	93 (16.2)	20.1 (12.2,33.1)	
GLCA	Normal Weight (< 25)	200 (34.8)	15.8 (10.6,23.6)	0.57
	Overweight (25-29.9)	281 (49)	17.3 (11.8,25.3)	
	Obese (≥30)	93 (16.2)	15.9 (10.3,24.4)	
UDCA	Normal Weight (< 25)	200 (34.8)	13.9 (8.1,24.0)	0.11
	Overweight (25-29.9)	281 (49)	17.7 (10.6,29.7)	
	Obese (≥30)	93 (16.2)	18.9 (10.5,33.8)	
GUDCA	Normal Weight (< 25)	200 (34.8)	77.8 (52.7,115.0)	0.33
	Overweight (25-29.9)	281 (49)	70.7 (48.7,102.4)	
	Obese (≥30)	93 (16.2)	82.3 (54.1,125.1)	
TUDCA	Normal Weight (< 25)	200 (34.8)	3.8 (2.7,5.2)	0.47
	Overweight (25-29.9)	281 (49)	3.5 (2.5,4.7)	
	Obese (≥30)	93 (16.2)	3.7 (2.6,5.2)	
HCA	Normal Weight (< 25)	200 (34.8)	5.9 (3.8,9.3)	0.51
	Overweight (25-29.9)	281 (49)	6.7 (4.3,10.3)	
	Obese (≥30)	93 (16.2)	6.6 (4.0,10.7)	
GHCA	Normal Weight (< 25)	200 (34.8)	12.3 (9.5,15.9)	0.09
	Overweight (25-29.9)	281 (49)	10.9 (8.5,13.9)	
	Obese (≥30)	93 (16.2)	10.5 (7.9,13.8)	

Table continued on next page;

Supplementary Table 5 (continued) Geometric means (95% confidence intervals) of plasma bile acid concentrations across strata of BMI

Bile Acid*	BMI	Frequency	Geometric Mean (95% CI)	P†
THCA	Normal Weight (< 25)	200 (34.8)	2.5 (2.0,3.2)	0.05
	Overweight (25-29.9)	281 (49)	2.3 (1.8,2.9)	
	Obese (≥30)	93 (16.2)	2.1 (1.6,2.7)	
TaMCA	Normal Weight (< 25)	200 (34.8)	5.1 (3.6,7.2)	0.68
	Overweight (25-29.9)	281 (49)	5.0 (3.6,7.0)	
	Obese (≥30)	93 (16.2)	4.6 (3.1,6.7)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, sex, fasting status, smoking status, education level, physical activity level, height, and consumption of alcohol, fiber, red and processed meat, fat, and energy;
*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);
†Two-sided P value for linear trend from Generalized Linear Models;

Supplementary Table 6 Geometric means (95% confidence intervals) of plasma bile acid concentrations across quartiles of red and processed meat consumption

Bile Acid*	Red and processed meat consumption	Frequency	Geometric Mean (95% CI)	P†
CA	Quartile 1	143 (24.9)	76.0 (40.8,141.5)	0.90
	Quartile 4	143 (24.9)	74.1 (40.2,136.9)	
CDCA	Quartile 1	143 (24.9)	140.5 (80.4,245.4)	0.99
	Quartile 4	143 (24.9)	148.0 (85.4,256.6)	
GCA	Quartile 1	143 (24.9)	141.6 (94.6,211.8)	0.67
	Quartile 4	143 (24.9)	152.2 (102.3,226.5)	
TCA	Quartile 1	143 (24.9)	16.3 (10.1,26.4)	0.67
	Quartile 4	143 (24.9)	18.1 (11.3,29.1)	
GCDCA	Quartile 1	143 (24.9)	591.1 (410.8,850.5)	0.56
	Quartile 4	143 (24.9)	641.8 (448.2,918.9)	
TCDCA	Quartile 1	143 (24.9)	63.5 (41.6,97.0)	0.56
	Quartile 4	143 (24.9)	69.3 (45.7,105.3)	
DCA	Quartile 1	143 (24.9)	171.5 (105.6,278.7)	0.79
	Quartile 4	143 (24.9)	159.1 (98.5,256.8)	
GDCA	Quartile 1	143 (24.9)	138.0 (88.0,216.4)	0.88
	Quartile 4	143 (24.9)	150.5 (96.5,234.5)	
TDCA	Quartile 1	143 (24.9)	19.7 (12.2,31.7)	0.85
	Quartile 4	143 (24.9)	21.7 (13.6,34.8)	
GLCA	Quartile 1	143 (24.9)	17.3 (11.4,26.1)	0.41
	Quartile 4	143 (24.9)	15.0 (10.0,22.5)	
UDCA	Quartile 1	143 (24.9)	19.1 (11.0,33.4)	0.61
	Quartile 4	143 (24.9)	15.6 (9.0,27.1)	
GUDCA	Quartile 1	143 (24.9)	76.8 (51.4,114.9)	0.82
	Quartile 4	143 (24.9)	75.7 (50.9,112.6)	
TUDCA	Quartile 1	143 (24.9)	3.6 (2.6,5.0)	0.92
	Quartile 4	143 (24.9)	3.6 (2.6,5.0)	
HCA	Quartile 1	143 (24.9)	5.9 (3.7,9.4)	0.76
	Quartile 4	143 (24.9)	6.6 (4.1,10.4)	
GHCA	Quartile 1	143 (24.9)	10.6 (8.1,13.8)	0.64
	Quartile 4	143 (24.9)	11.7 (9.0,15.1)	
THCA	Quartile 1	143 (24.9)	2.2 (1.8,2.8)	0.53
	Quartile 4	143 (24.9)	2.3 (1.8,3.0)	
TaMCA	Quartile 1	143 (24.9)	4.3 (3.0,6.1)	0.08
	Quartile 4	143 (24.9)	5.5 (3.9,7.9)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, sex, fasting status, smoking status, education level, physical activity level, BMI, height, and consumption of alcohol, fiber, fat, and energy;

*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);

†Two-sided P value for linear trend from Generalized Linear Models;

Supplementary Table 7 Geometric means (95% confidence intervals) of plasma bile acid concentrations across quartiles of fiber consumption

Bile Acid*	Fiber consumption	Frequency	Geometric Mean (95% CI)	P†
CA	Quartile 1	143 (24.9)	64.0 (34.0,120.2)	0.73
	Quartile 4	143 (24.9)	73.0 (39.9,133.4)	
CDCA	Quartile 1	143 (24.9)	124.6 (70.8,219.3)	0.56
	Quartile 4	143 (24.9)	146.1 (85.1,250.8)	
GCA	Quartile 1	143 (24.9)	135.3 (90.0,203.6)	0.06
	Quartile 4	143 (24.9)	175.3 (118.6,259.0)	
TCA	Quartile 1	143 (24.9)	16.1 (9.9,26.2)	0.09
	Quartile 4	143 (24.9)	20.1 (12.6,32.1)	
GCDCA	Quartile 1	143 (24.9)	594.1 (410.8,859.2)	0.09
	Quartile 4	143 (24.9)	716.0 (503.2,1018.8)	
TCDCA	Quartile 1	143 (24.9)	65.8 (42.8,101.1)	0.09
	Quartile 4	143 (24.9)	75.5 (50.0,114.0)	
DCA	Quartile 1	143 (24.9)	146.6 (89.3,240.5)	0.67
	Quartile 4	143 (24.9)	154.6 (96.3,248.1)	
GDCA	Quartile 1	143 (24.9)	137.6 (87.1,217.5)	0.68
	Quartile 4	143 (24.9)	149.8 (96.7,232.0)	
TDCA	Quartile 1	143 (24.9)	20.0 (12.3,32.4)	0.63
	Quartile 4	143 (24.9)	20.9 (13.1,33.3)	
GLCA	Quartile 1	143 (24.9)	18.7 (12.3,28.3)	0.21
	Quartile 4	143 (24.9)	16.8 (11.3,25.1)	
UDCA	Quartile 1	143 (24.9)	14.6 (8.3,25.7)	0.54
	Quartile 4	143 (24.9)	15.7 (9.1,27.0)	
GUDCA	Quartile 1	143 (24.9)	70.2 (46.6,105.7)	0.96
	Quartile 4	143 (24.9)	75.9 (51.3,112.3)	
TUDCA	Quartile 1	143 (24.9)	3.4 (2.4,4.7)	0.60
	Quartile 4	143 (24.9)	3.7 (2.7,5.1)	
HCA	Quartile 1	143 (24.9)	5.9 (3.7,9.5)	0.73
	Quartile 4	143 (24.9)	6.5 (4.1,10.2)	
GHCA	Quartile 1	143 (24.9)	10.3 (7.9,13.5)	0.03
	Quartile 4	143 (24.9)	13.2 (10.2,17.1)	
THCA	Quartile 1	143 (24.9)	2.3 (1.8,2.9)	0.24
	Quartile 4	143 (24.9)	2.5 (2.0,3.1)	
TaMCA	Quartile 1	143 (24.9)	5.0 (3.4,7.2)	0.67
	Quartile 4	143 (24.9)	5.0 (3.5,7.1)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, sex, fasting status, smoking status, education level, physical activity level, BMI, height, and consumption of alcohol, red and processed meat, fat, and energy;
*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);
†Two-sided P value for linear trend from Generalized Linear Models;

Supplementary Table 8 Geometric means (95% confidence intervals) of plasma bile acid concentrations across quartiles of fat consumption

Bile Acid*	Fat consumption	Frequency	Geometric Mean (95% CI)	P†
CA	Quartile 1	143 (24.9)	72.2 (37.0,141.0)	0.68
	Quartile 4	143 (24.9)	65.9 (35.0,123.9)	
CDCA	Quartile 1	143 (24.9)	153.7 (84.4,280.2)	0.67
	Quartile 4	143 (24.9)	121.9 (69.1,215.0)	
GCA	Quartile 1	143 (24.9)	140.7 (91.1,217.3)	0.95
	Quartile 4	143 (24.9)	150.7 (99.9,227.2)	
TCA	Quartile 1	143 (24.9)	17.3 (10.3,29.0)	0.93
	Quartile 4	143 (24.9)	18.0 (11.0,29.4)	
GCDCA	Quartile 1	143 (24.9)	657.8 (444.1,974.4)	0.69
	Quartile 4	143 (24.9)	606.8 (418.5,879.8)	
TCDCA	Quartile 1	143 (24.9)	70.6 (44.7,111.5)	0.25
	Quartile 4	143 (24.9)	72.8 (47.3,112.2)	
DCA	Quartile 1	143 (24.9)	173.2 (102.6,292.5)	0.22
	Quartile 4	143 (24.9)	127.7 (77.8,209.5)	
GDCA	Quartile 1	143 (24.9)	148.8 (91.5,242.0)	0.86
	Quartile 4	143 (24.9)	126.0 (79.6,199.5)	
TDCA	Quartile 1	143 (24.9)	22.0 (13.2,36.9)	0.69
	Quartile 4	143 (24.9)	19.8 (12.2,32.2)	
GLCA	Quartile 1	143 (24.9)	17.2 (11.0,26.8)	0.81
	Quartile 4	143 (24.9)	15.5 (10.2,23.7)	
UDCA	Quartile 1	143 (24.9)	17.6 (9.7,32.2)	0.77
	Quartile 4	143 (24.9)	14.1 (8.0,25.0)	
GUDCA	Quartile 1	143 (24.9)	78.9 (51.1,121.9)	0.69
	Quartile 4	143 (24.9)	71.1 (47.1,107.3)	
TUDCA	Quartile 1	143 (24.9)	3.7 (2.6,5.3)	0.71
	Quartile 4	143 (24.9)	3.6 (2.5,5.0)	
HCA	Quartile 1	143 (24.9)	6.9 (4.1,11.4)	0.92
	Quartile 4	143 (24.9)	6.2 (3.9,10.0)	
GHCA	Quartile 1	143 (24.9)	10.4 (7.8,13.8)	0.44
	Quartile 4	143 (24.9)	12.4 (9.5,16.3)	
THCA	Quartile 1	143 (24.9)	2.3 (1.8,3.0)	0.84
	Quartile 4	143 (24.9)	2.5 (1.9,3.2)	
TaMCA	Quartile 1	143 (24.9)	4.5 (3.0,6.6)	0.55
	Quartile 4	143 (24.9)	5.3 (3.7,7.7)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, sex, fasting status, smoking status, education level, physical activity level, BMI, height, and consumption of alcohol, fiber, red and processed meat, and energy;
*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);
†Two-sided P value for linear trend from Generalized Linear Models;

Supplementary Table 9 Geometric means (95% confidence intervals) of plasma bile acid concentrations across strata of alcohol intake

Bile Acid*	Alcohol Intake	Frequency	Geometric Mean (95% CI)	P†
CA	Non Consumer	111 (19.3)	90.8 (48.1,171.5)	0.43
	<0-12 g/day	257 (44.8)	72.6 (40.2,131.2)	
	12-24 g/day	94 (16.4)	65.3 (35.0,121.8)	
	>24 g/day	112 (19.5)	65.2 (34.9,121.7)	
CDCA	Non Consumer	111 (19.3)	184.4 (104.4,325.8)	0.23
	<0-12 g/day	257 (44.8)	137.8 (81.1,234.0)	
	12-24 g/day	94 (16.4)	131.1 (74.9,229.2)	
	>24 g/day	112 (19.5)	130.5 (74.6,228.4)	
GCA	Non Consumer	111 (19.3)	174.2 (115.7,262.4)	0.02
	<0-12 g/day	257 (44.8)	134.2 (91.7,196.5)	
	12-24 g/day	94 (16.4)	123.4 (82.6,184.5)	
	>24 g/day	112 (19.5)	164.4 (109.9,245.9)	
TCA	Non Consumer	111 (19.3)	19.8 (12.1,32.3)	0.05
	<0-12 g/day	257 (44.8)	15.8 (10.0,25.0)	
	12-24 g/day	94 (16.4)	14.6 (9.0,23.7)	
	>24 g/day	112 (19.5)	21.1 (13.0,34.1)	
GCDCA	Non Consumer	111 (19.3)	685.2 (472.5,993.7)	0.20
	<0-12 g/day	257 (44.8)	585.3 (414.0,827.3)	
	12-24 g/day	94 (16.4)	564.0 (391.5,812.5)	
	>24 g/day	112 (19.5)	676.0 (469.1,974.1)	
TCDCA	Non Consumer	111 (19.3)	76.4 (49.5,117.9)	0.02
	<0-12 g/day	257 (44.8)	60.6 (40.5,90.7)	
	12-24 g/day	94 (16.4)	60.2 (39.3,92.1)	
	>24 g/day	112 (19.5)	83.2 (54.3,127.4)	
DCA	Non Consumer	111 (19.3)	213.2 (129.8,350.3)	0.04
	<0-12 g/day	257 (44.8)	167.2 (105.3,265.5)	
	12-24 g/day	94 (16.4)	138.2 (84.9,225.1)	
	>24 g/day	112 (19.5)	130.9 (80.3,213.2)	
GDCA	Non Consumer	111 (19.3)	174.5 (110.3,276.1)	0.07
	<0-12 g/day	257 (44.8)	145.9 (95.1,223.7)	
	12-24 g/day	94 (16.4)	115.2 (73.4,180.9)	
	>24 g/day	112 (19.5)	128.7 (82.0,202.1)	
TDCA	Non Consumer	111 (19.3)	24.2 (14.9,39.5)	0.16
	<0-12 g/day	257 (44.8)	20.1 (12.7,31.6)	
	12-24 g/day	94 (16.4)	16.8 (10.4,27.1)	
	>24 g/day	112 (19.5)	21.0 (13.0,33.9)	
GLCA	Non Consumer	111 (19.3)	18.3 (12.0,27.9)	0.22
	<0-12 g/day	257 (44.8)	18.3 (12.4,27.1)	
	12-24 g/day	94 (16.4)	14.5 (9.6,21.9)	
	>24 g/day	112 (19.5)	15.3 (10.1,23.1)	

Table continued on next page;

Supplementary Table 9 (continued) Geometric means (95% confidence intervals) of plasma bile acid concentrations across strata of alcohol intake

Bile Acid*	Alcohol Intake	Frequency	Geometric Mean (95% CI)	P†
UDCA	Non Consumer	111 (19.3)	23.2 (13.1,40.9)	0.04
	<0-12 g/day	257 (44.8)	15.2 (9.0,25.9)	
	12-24 g/day	94 (16.4)	16.3 (9.3,28.5)	
	>24 g/day	112 (19.5)	14.2 (8.1,24.8)	
GUDCA	Non Consumer	111 (19.3)	92.9 (61.7,139.9)	0.02
	<0-12 g/day	257 (44.8)	66.6 (45.5,97.4)	
	12-24 g/day	94 (16.4)	67.2 (45.0,100.5)	
	>24 g/day	112 (19.5)	76.9 (51.4,114.9)	
TUDCA	Non Consumer	111 (19.3)	4.2 (3.0,5.9)	0.03
	<0-12 g/day	257 (44.8)	3.4 (2.4,4.6)	
	12-24 g/day	94 (16.4)	3.2 (2.3,4.4)	
	>24 g/day	112 (19.5)	3.8 (2.7,5.3)	
HCA	Non Consumer	111 (19.3)	8.0 (5.0,12.9)	0.18
	<0-12 g/day	257 (44.8)	6.2 (4.0,9.6)	
	12-24 g/day	94 (16.4)	5.8 (3.6,9.3)	
	>24 g/day	112 (19.5)	6.0 (3.8,9.7)	
GHCA	Non Consumer	111 (19.3)	12.2 (9.3,16.0)	0.29
	<0-12 g/day	257 (44.8)	11.1 (8.7,14.3)	
	12-24 g/day	94 (16.4)	10.5 (8.0,13.6)	
	>24 g/day	112 (19.5)	11.8 (9.1,15.4)	
THCA	Non Consumer	111 (19.3)	2.5 (2.0,3.2)	0.57
	<0-12 g/day	257 (44.8)	2.3 (1.9,2.9)	
	12-24 g/day	94 (16.4)	2.3 (1.8,2.9)	
	>24 g/day	112 (19.5)	2.4 (1.9,3.1)	
TaMCA	Non Consumer	111 (19.3)	6.3 (4.4,9.1)	0.004
	<0-12 g/day	257 (44.8)	5.4 (3.8,7.6)	
	12-24 g/day	94 (16.4)	4.3 (3.0,6.1)	
	>24 g/day	112 (19.5)	4.1 (2.8,5.8)	

Geometric means (95 % confidence intervals) from Generalized Linear Models adjusted for age, sex, fasting status, smoking status, education level, physical activity level, BMI, height, and consumption of fiber, red and processed meat, fat, and energy;

*CA (Cholic Acid); CDCA (Chenodeoxycholic Acid); DCA (Deoxycholic Acid); GCA (Glycocholic Acid); GCDCA (Glycochenodeoxycholic Acid); GDCA (Glycodeoxycholic Acid); GHCA (Glycohyocholic Acid); GLCA (Glycolithocholic Acid); GUDCA (Glycoursodeoxycholic Acid); HCA (Hyocholic Acid); TaMCA (Tauro-Alpha-Muricholic Acid); TCA (Taurocholic Acid); TCDCA (Taurochenodeoxycholic Acid); TDCA (Taurodeoxycholic Acid); THCA (Taurohyocholic Acid); TUDCA (Tauroursodeoxycholic Acid); UDCA (Ursodeoxycholic Acid);

†Two-sided P value for linear trend from Generalized Linear Models;

Supplementary Table 10 Odds ratios (95 % confidence intervals) for colon cancer across quartiles of *relative* concentrations of bile acids

Bile acid metabolite		Quartile 1	Quartile 2	Quartile 3	Quartile 4	p_{trend}*
Cholic Acid	N Cases (%)	150 (26.4)	122 (21.4)	152 (26.7)	145 (25.5)	
(CA)	OR (95%CI)	Reference	0.81 (0.56-1.17)	1.07 (0.74-1.53)	0.95 (0.64-1.39)	0.77
Chenodeoxycholic Acid	N Cases (%)	144 (25.3)	143 (25.1)	132 (23.2)	150 (26.4)	
(CDCA)	OR (95%CI)	Reference	0.93 (0.65-1.33)	0.90 (0.61-1.33)	0.98 (0.66-1.44)	0.43
Hyocholic Acid	N Cases (%)	137 (24.1)	175 (30.8)	111 (19.5)	146 (25.7)	
(HCA)	OR (95%CI)	Reference	0.93 (0.66-1.31)	0.90 (0.60-1.35)	0.92 (0.63-1.36)	0.42
Tauro-Alpha-Muricholic Acid	N Cases (%)	110 (19.3)	153 (26.9)	152 (26.7)	154 (27.1)	
(TaMCA)	OR (95%CI)	Reference	0.88 (0.61-1.28)	0.96 (0.65-1.41)	0.98 (0.65-1.46)	0.28
Glycocholic Acid	N Cases (%)	123 (21.6)	124 (21.8)	154 (27.1)	168 (29.5)	
(GCA)	OR (95%CI)	Reference	1.10 (0.77-1.57)	1.27 (0.89-1.81)	1.48 (1.04-2.12)	0.09
Taurocholic Acid	N Cases (%)	139 (24.4)	126 (22.1)	135 (23.7)	169 (29.7)	
(TCA)	OR (95%CI)	Reference	0.96 (0.68-1.37)	0.98 (0.68-1.41)	1.29 (0.91-1.82)	0.45
Glycochenodeoxycholic Acid	N Cases (%)	127 (22.3)	152 (26.7)	146 (25.7)	144 (25.3)	
(GCDCA)	OR (95%CI)	Reference	1.08 (0.76-1.52)	1.08 (0.75-1.56)	1.09 (0.75-1.57)	0.57
Taurochenodeoxycholic Acid	N Cases (%)	121 (21.3)	159 (27.9)	148 (26.0)	141 (24.8)	
(TCDCA)	OR (95%CI)	Reference	1.28 (0.90-1.82)	1.29 (0.90-1.83)	1.14 (0.79-1.65)	0.62
Glycohyocholic Acid	N Cases (%)	158 (27.8)	125 (22.0)	136 (23.9)	150 (26.4)	
(GHCA)	OR (95%CI)	Reference	0.76 (0.54-1.09)	0.83 (0.57-1.19)	0.98 (0.68-1.42)	0.27
Deoxycholic Acid	N Cases (%)	135 (23.7)	136 (23.9)	171 (30.1)	127 (22.3)	
(DCA)	OR (95%CI)	Reference	1.02 (0.71-1.47)	1.22 (0.87-1.72)	0.90 (0.63-1.28)	0.30
Glycodeoxycholic Acid	N Cases (%)	144 (25.3)	143 (25.1)	132 (23.2)	150 (26.4)	
(GDCA)	OR (95%CI)	Reference	0.93 (0.65-1.33)	0.90 (0.61-1.33)	0.98 (0.66-1.44)	0.40
Taurodeoxycholic Acid	N Cases (%)	129 (22.7)	149 (26.2)	124 (21.8)	167 (29.3)	
(TDCA)	OR (95%CI)	Reference	0.81 (0.57-1.15)	0.98 (0.67-1.42)	1.12 (0.77-1.64)	0.65
Glycolithocholic Acid	N Cases (%)	159 (27.9)	154 (27.1)	132 (23.2)	124 (21.8)	
(GLCA)	OR (95%CI)	Reference	0.94 (0.67-1.31)	0.73 (0.51-1.04)	0.78 (0.54-1.13)	0.12
Ursodeoxycholic Acid	N Cases (%)	147 (25.8)	138 (24.3)	147 (25.8)	137 (24.1)	
(UDCA)	OR (95%CI)	Reference	0.85 (0.60-1.22)	1.04 (0.73-1.49)	0.89 (0.61-1.30)	0.99
Glycoursodeoxycholic Acid	N Cases (%)	147 (25.8)	168 (29.5)	136 (23.9)	118 (20.7)	
(GUDCA)	OR (95%CI)	Reference	1.12 (0.79-1.58)	1.02 (0.72-1.47)	0.79 (0.54-1.17)	0.44
Tauroursodeoxycholic Acid	N Cases (%)	130 (22.8)	212 (37.3)	114 (20.0)	113 (19.9)	
(TUDCA)	OR (95%CI)	Reference	0.96 (0.69-1.35)	1.03 (0.69-1.54)	0.87 (0.58-1.30)	0.63

Results from conditional logistic regression analyses;
Conditioned on age, sex, study center, follow-up time, time of day at blood collection, fasting status, menopausal status (women), current use of exogenous hormones (women), and phase of menstrual cycle (women), and adjusted for education level, physical activity level, smoking, waist circumference, height, and consumption of red and processed meat, fat, fiber, and energy;
*Two-sided P value for linear trend from logistic regression model with bile acid values on the log2 scale as the independent variable;

Supplementary Table 11 Odds ratios (95 % confidence intervals) for colon cancer across quartiles of bile acids, stratified by anatomic site

Bile acid metabolite		Quartile 1	Quartile 2	Quartile 3	Quartile 4	p_{trend}*	$p_{hetero.}$†
Cholic Acid (CA)							
PROXIMAL	N Cases (%)	58 (25.7)	45 (19.9)	59 (26.1)	64 (28.3)		0.79
	OR (95%CI)	Reference	0.57 (0.31-1.04)	1.02 (0.53-1.96)	1.25 0.68-2.31)	0.26	
DISTAL	N Cases (%)	48 (23.1)	53 (25.5)	53 (25.5)	54 (26.0)		
	OR (95%CI)	Reference	2.0 (1.08-3.72)	1.50 (0.79-2.87)	1.88 (0.92-3.83)	0.20	
Chenodeoxycholic Acid (CDCA)							
PROXIMAL	N Cases (%)	58 (25.7)	55 (24.3)	51 (22.6)	62 (27.4)		0.34
	OR (95%CI)	Reference	0.72 (0.38-1.36)	0.89 (0.48-1.67)	1.29 (0.67-2.48)	0.23	
DISTAL	N Cases (%)	51 (24.5)	52 (25.0)	52 (25.0)	53 (25.5)		
	OR (95%CI)	Reference	1.04 (0.56-1.91)	1.30 (0.66-2.54)	1.31 (0.67-2.59)	0.86	
Hyocholic Acid (HCA)							
PROXIMAL	N Cases (%)	57 (25.2)	54 (23.9)	67 (29.6)	48 (21.2)		0.36
	OR (95%CI)	Reference	1.13 (0.61-2.10)	1.58 (0.88-2.84)	1.14 (0.59-2.18)	0.09	
DISTAL	N Cases (%)	64 (30.8)	46 (22.1)	53 (25.5)	45 (21.6)		
	OR (95%CI)	Reference	0.57 (0.31-1.04)	0.85 (0.45-1.62)	0.88 (0.43-1.82)	0.73	
Glycocholic Acid (GCA)							
PROXIMAL	N Cases (%)	41 (18.1)	57 (25.2)	54 (23.9)	74 (32.7)		0.29
	OR (95%CI)	Reference	1.23 (0.63-2.40)	1.06 (0.54-2.10)	1.83 (0.96-3.52)	0.43	
DISTAL	N Cases (%)	38 (18.3)	47 (22.6)	51 (24.5)	72 (34.6)		
	OR (95%CI)	Reference	1.32 (0.69-2.53)	1.61 (0.80-3.23)	2.07 (1.04-4.12)	0.03	
Taurocholic Acid (TCA)							
PROXIMAL	N Cases (%)	44 (19.5)	55 (24.3)	54 (23.9)	73 (32.3)		0.70
	OR (95%CI)	Reference	1.07 (0.56-2.05)	1.41 (0.75-2.63)	1.45 (0.80-2.60)	0.49	
DISTAL	N Cases (%)	38 (18.3)	53 (25.5)	47 (22.6)	70 (33.7)		
	OR (95%CI)	Reference	1.44 (0.76-2.74)	0.84 (0.43-1.65)	1.53 (0.82-2.89)	0.25	
Glycochenodeoxycholic Acid (GCDCA)							
PROXIMAL	N Cases (%)	41 (18.1)	63 (27.9)	57 (25.2)	65 (28.8)		0.93
	OR (95%CI)	Reference	2.12 (1.14-3.92)	1.27 (0.67-2.40)	2.05 (1.03-4.06)	0.41	
DISTAL	N Cases (%)	40 (19.2)	61 (29.3)	42 (20.2)	65 (31.3)		
	OR (95%CI)	Reference	1.33 (0.71-2.51)	0.79 (0.38-1.65)	1.30 (0.63-2.70)	0.52	
Taurochenodeoxycholic Acid (TCDCA)							
PROXIMAL	N Cases (%)	40 (17.7)	65 (28.8)	59 (26.1)	62 (27.4)		0.80
	OR (95%CI)	Reference	1.30 (0.73-2.32)	1.54 (0.85-2.77)	1.50 (0.83-2.74)	0.34	
DISTAL	N Cases (%)	41 (19.7)	54 (26.0)	49 (23.6)	64 (30.8)		
	OR (95%CI)	Reference	1.65 (0.86-3.14)	0.86 (0.45-1.63)	1.37 (0.70-2.68)	0.62	
Glycohyocholic Acid (GHCA)							
PROXIMAL	N Cases (%)	56 (24.8)	46 (20.4)	69 (30.5)	55 (24.3)		0.24
	OR (95%CI)	Reference	1.34 (0.73-2.45)	1.48 (0.83-2.61)	1.47 (0.78-2.80)	0.56	
DISTAL	N Cases (%)	53 (25.5)	34 (16.3)	57 (27.4)	64 (30.8)		
	OR (95%CI)	Reference	1.88 (0.90-3.91)	1.91 (0.98-3.69)	2.42 (1.21-4.86)	0.04	

Table continued on next page;

Supplementary Table 11 (continued) Odds ratios (95 % confidence intervals) for colon cancer across quartiles of bile acids, stratified by anatomic site

Bile acid metabolite		Quartile 1	Quartile 2	Quartile 3	Quartile 4	p_{trend}*	$p_{hetero.}$†
Tauro-Alpha-Murocholic Acid (TaMCA)							
PROXIMAL	N Cases (%)	89 (39.4)	20 (8.8)	61 (27.0)	56 (24.8)		0.17
	OR (95%CI)	Reference	0.90 (0.43-1.89)	0.95 (0.55-1.64)	1.05 (0.59-1.87)	0.95	
DISTAL	N Cases (%)	57 (27.4)	26 (12.5)	60 (28.8)	65 (31.3)		
	OR (95%CI)	Reference	1.51 (0.69-3.30)	1.74 (0.97-3.11)	2.07 (1.11-3.86)	0.05	
Deoxycholic Acid (DCA)							
PROXIMAL	N Cases (%)	57 (25.2)	53 (23.5)	54 (23.9)	62 (27.4)		0.47
	OR (95%CI)	Reference	0.84 (0.47-1.47)	0.82 (0.44-1.52)	1.14 (0.61-2.14)	0.79	
DISTAL	N Cases (%)	37 (17.8)	48 (23.1)	71 (34.1)	52 (25.0)		
	OR (95%CI)	Reference	1.35 (0.68-2.65)	2.32 (1.22-4.42)	1.36 (0.70-2.66)	0.23	
Glycodeoxycholic Acid (GDCA)							
PROXIMAL	N Cases (%)	49 (21.7)	59 (26.1)	52 (23.0)	66 (29.2)		0.13
	OR (95%CI)	Reference	1.24 (0.68-2.26)	1.02 (0.54-1.93)	1.41 (0.73-2.71)	0.85	
DISTAL	N Cases (%)	28 (13.5)	50 (24.0)	65 (31.3)	65 (31.3)		
	OR (95%CI)	Reference	1.28 (0.64-2.59)	1.66 (0.83-3.29)	1.83 (0.86-3.91)	0.07	
Taurodeoxycholic Acid (TDCA)							
PROXIMAL	N Cases (%)	46 (20.4)	65 (28.8)	60 (26.5)	55 (24.3)		0.34
	OR (95%CI)	Reference	1.26 (0.68-2.33)	1.42 (0.78-2.61)	1.0 (0.51-1.94)	0.95	
DISTAL	N Cases (%)	32 (15.4)	49 (23.6)	61 (29.3)	66 (31.7)		
	OR (95%CI)	Reference	1.19 (0.60-2.39)	1.10 (0.56-2.14)	1.66 (0.81-3.42)	0.19	
Glycolithocholic Acid (GLCA)							
PROXIMAL	N Cases (%)	54 (23.9)	66 (29.2)	58 (25.7)	48 (21.2)		0.18
	OR (95%CI)	Reference	1.34 (0.71-2.52)	0.85 (0.46-1.54)	0.62 (0.33-1.19)	0.02	
DISTAL	N Cases (%)	49 (23.6)	56 (26.9)	46 (22.1)	57 (27.4)		
	OR (95%CI)	Reference	0.78 (0.43-1.43)	0.73 (0.37-1.46)	0.87 (0.43-1.77)	0.83	
Ursodeoxycholic Acid (UDCA)							
PROXIMAL	N Cases (%)	55 (24.3)	56 (24.8)	49 (21.7)	66 (29.2)		0.39
	OR (95%CI)	Reference	0.74 (0.40-1.37)	0.71 (0.38-1.31)	1.67 (0.89-3.14)	0.17	
DISTAL	N Cases (%)	61 (29.3)	45 (21.6)	54 (26.0)	48 (23.1)		
	OR (95%CI)	Reference	0.63 (0.34-1.18)	1.06 (0.56-1.97)	0.83 (0.44-1.57)	0.91	
Glycoursodeoxycholic Acid (GUDCA)							
PROXIMAL	N Cases (%)	56 (24.8)	66 (29.2)	37 (16.4)	67 (29.6)		0.65
	OR (95%CI)	Reference	0.95 (0.54-1.70)	0.54 (0.28-1.02)	1.47 (0.78-2.76)	0.96	
DISTAL	N Cases (%)	43 (20.7)	72 (34.6)	37 (17.8)	56 (26.9)		
	OR (95%CI)	Reference	1.96 (1.07-3.59)	0.95 (0.47-1.96)	1.26 (0.65-2.44)	0.52	
Tauroursodeoxycholic Acid (TUDCA)							
PROXIMAL	N Cases (%)	62 (27.4)	67 (29.6)	34 (15.0)	63 (27.9)		0.81
	OR (95%CI)	Reference	1.34 (0.77-2.32)	0.68 (0.35-1.33)	1.34 (0.76-2.38)	0.63	
DISTAL	N Cases (%)	56 (26.9)	55 (26.4)	39 (18.8)	58 (27.9)		
	OR (95%CI)	Reference	0.80 (0.44-1.46)	0.90 (0.44-1.82)	0.99 (0.52-1.87)	0.92	

Results from conditional logistic regression analyses;
Conditioned on age, sex, study center, follow-up time, time of day at blood collection, fasting status, menopausal status (women), current use of exogenous hormones (women), and phase of menstrual cycle (women), and adjusted for education level, physical activity level, smoking, waist circumference, height, and consumption of red and processed meat, fat, fiber, and energy;
*Two-sided P value for linear trend from logistic regression model with bile acid values on the log2 scale as the independent variable;
† Two-sided P value for heterogeneity from X^2 tests

Supplementary Table 12 Odds ratios (95 % confidence intervals) for colon cancer across quartiles of bile acids, stratified by sex

Bile acid metabolite		Quartile 1	Quartile 2	Quartile 3	Quartile 4	p$_{trend}$*	p$_{hetero.}$†
Cholic Acid (CA)							
WOMEN	N Cases (%)	76 (21.3)	91 (25.6)	91 (25.6)	98 (27.5)		0.40
	OR (95%CI)	Reference	1.24 (0.78-1.96)	1.17 (0.71-1.93)	1.41 (0.86-2.32)	0.13	
MEN	N Cases (%)	53 (24.9)	43 (20.2)	68 (31.9)	49 (23.0)		
	OR (95%CI)	Reference	0.72 (0.39-1.32)	1.32 (0.70-2.49)	0.77 (0.40-1.47)	0.75	
Chenodeoxycholic Acid (CDCA)							
WOMEN	N Cases (%)	81 (22.8)	86 (24.2)	99 (27.8)	90 (25.3)		0.22
	OR (95%CI)	Reference	1.06 (0.67-1.68)	1.2 (0.76-1.90)	1.18 (0.73-1.89)	0.34	
MEN	N Cases (%)	51 (23.9)	46 (21.6)	68 (31.9)	48 (22.5)		
	OR (95%CI)	Reference	0.7 (0.35-1.40)	1.14 (0.57-2.27)	0.79 (0.40-1.58)	0.23	
Hyocholic Acid (HCA)							
WOMEN	N Cases (%)	88 (24.7)	79 (22.2)	110 (30.9)	79 (22.2)		0.07
	OR (95%CI)	Reference	0.8 (0.49-1.31)	1.31 (0.82-2.08)	0.94 (0.57-1.53)	0.45	
MEN	N Cases (%)	53 (24.9)	52 (24.4)	60 (28.2)	48 (22.5)		
	OR (95%CI)	Reference	0.88 (0.48-1.64)	1.04 (0.55-1.95)	0.91 (0.46-1.80)	0.84	
Glycocholic Acid (GCA)							
WOMEN	N Cases (%)	67 (18.8)	65 (18.3)	87 (24.4)	137 (38.5)		0.07
	OR (95%CI)	Reference	0.89 (0.53-1.48)	1.52 (0.91-2.55)	2.52 (1.52-4.20)	<0.001	
MEN	N Cases (%)	48 (22.5)	63 (29.6)	44 (20.7)	58 (27.2)		
	OR (95%CI)	Reference	1.41 (0.79-2.53)	1.01 (0.54-1.90)	1.25 (0.67-2.34)	0.89	
Taurocholic Acid (TCA)							
WOMEN	N Cases (%)	64 (18.0)	86 (24.2)	81 (22.8)	125 (35.1)		0.11
	OR (95%CI)	Reference	1.36 (0.83-2.25)	1.28 (0.78-2.10)	2.46 (1.49-4.08)	0.006	
MEN	N Cases (%)	48 (22.5)	53 (24.9)	50 (23.5)	62 (29.1)		
	OR (95%CI)	Reference	1.12 (0.62-2.03)	1.05 (0.56-1.97)	1.18 (0.65-2.15)	0.80	
Glycochenodeoxycholic Acid (GCDCA)							
WOMEN	N Cases (%)	63 (17.7)	91 (25.6)	85 (23.9)	117 (32.9)		0.53
	OR (95%CI)	Reference	1.46 (0.92-2.34)	1.19 (0.73-1.94)	2.22 (1.33-3.71)	0.007	
MEN	N Cases (%)	55 (25.8)	56 (26.3)	58 (27.2)	44 (20.7)		
	OR (95%CI)	Reference	1.11 (0.59-2.06)	0.96 (0.51-1.79)	0.78 (0.40-1.54)	0.68	
Taurochenodeoxycholic Acid (TCDCA)							
WOMEN	N Cases (%)	59 (16.6)	95 (26.7)	80 (22.5)	122 (34.3)		0.26
	OR (95%CI)	Reference	1.49 (0.93-2.40)	1.32 (0.81-2.15)	2.23 (1.37-3.65)	0.009	
MEN	N Cases (%)	46 (21.6)	57 (26.8)	56 (26.3)	54 (25.4)		
	OR (95%CI)	Reference	1.22 (0.67-2.24)	1.11 (0.61-2.03)	0.99 (0.53-1.85)	0.99	
Glycohyocholic Acid (GHCA)							
WOMEN	N Cases (%)	105 (29.5)	40 (11.2)	108 (30.3)	103 (28.9)		0.34
	OR (95%CI)	Reference	1.04 (0.61-1.79)	1.51 (0.98-2.32)	1.71 (1.06-2.77)	0.02	
MEN	N Cases (%)	47 (22.1)	65 (30.5)	50 (23.5)	51 (23.9)		
	OR (95%CI)	Reference	1.66 (0.92-3.01)	1.13 (0.60-2.12)	1.36 (0.70-2.65)	0.95	

Table continued on next page;

Supplementary Table 12 (continued) Odds ratios (95 % confidence intervals) for colon cancer across quartiles of bile acids, stratified by sex

Bile acid metabolite		Quartile 1	Quartile 2	Quartile 3	Quartile 4	p_{trend}*	$p_{hetero.}$[†]
Tauro-Alpha-Murocholic Acid (TaMCA)							
WOMEN	N Cases (%)	48 (13.5)	117 (32.9)	103 (28.9)	88 (24.7)		0.51
	OR (95%CI)	Reference	1.62 (1.00-2.61)	1.81 (1.06-3.08)	1.52 (0.88-2.63)	0.46	
MEN	N Cases (%)	32 (15.0)	60 (28.2)	64 (30.0)	57 (26.8)		
	OR (95%CI)	Reference	1.15 (0.59-2.21)	1.72 (0.85-3.47)	1.65 (0.79-3.44)	0.07	
Deoxycholic Acid (DCA)							
WOMEN	N Cases (%)	87 (24.5)	78 (22.0)	103 (29.0)	87 (24.5)		0.16
	OR (95%CI)	Reference	0.89 (0.57-1.40)	1.13 (0.72-1.78)	0.91 (0.56-1.47)	0.71	
MEN	N Cases (%)	41 (19.2)	54 (25.4)	59 (27.7)	59 (27.7)		
	OR (95%CI)	Reference	1.31 (0.71-2.41)	1.72 (0.92-3.24)	1.20 (0.63-2.31)	0.69	
Glycodeoxycholic Acid (GDCA)							
WOMEN	N Cases (%)	65 (18.3)	80 (22.5)	111 (31.2)	100 (28.1)		0.23
	OR (95%CI)	Reference	1.35 (0.82-2.22)	1.69 (1.05-2.73)	1.58 (0.92-2.71)	0.19	
MEN	N Cases (%)	40 (18.8)	61 (28.6)	44 (20.7)	68 (31.9)		
	OR (95%CI)	Reference	1.78 (0.94-3.37)	1.11 (0.55-2.24)	2.20 (1.06-4.60)	0.46	
Taurodeoxycholic Acid (TDCA)							
WOMEN	N Cases (%)	67 (18.8)	93 (26.1)	97 (27.2)	99 (27.8)		0.66
	OR (95%CI)	Reference	1.54 (0.95-2.50)	1.39 (0.86-2.24)	1.54 (0.91-2.62)	0.22	
MEN	N Cases (%)	46 (21.6)	41 (19.2)	65 (30.5)	61 (28.6)		
	OR (95%CI)	Reference	0.91 (0.48-1.74)	1.25 (0.68-2.30)	1.37 (0.69-2.72)	0.44	
Glycolithocholic Acid (GLCA)							
WOMEN	N Cases (%)	94 (26.4)	83 (23.3)	100 (28.1)	79 (22.2)		0.11
	OR (95%CI)	Reference	0.74 (0.47-1.18)	1.01 (0.62-1.64)	0.76 (0.46-1.26)	0.41	
MEN	N Cases (%)	44 (20.7)	70 (32.9)	44 (20.7)	55 (25.8)		
	OR (95%CI)	Reference	1.50 (0.80-2.81)	1.16 (0.60-2.24)	1.50 (0.75-2.99)	0.68	
Ursodeoxycholic Acid (UDCA)							
WOMEN	N Cases (%)	107 (30.1)	66 (18.5)	82 (23.0)	101 (28.4)		0.06
	OR (95%CI)	Reference	0.49 (0.30-0.78)	0.70 (0.44-1.10)	0.87 (0.54-1.38)	0.65	
MEN	N Cases (%)	40 (18.8)	78 (36.6)	53 (24.9)	42 (19.7)		
	OR (95%CI)	Reference	2.03 (0.98-4.18)	1.26 (0.65-2.45)	1.00 (0.51-1.98)	0.58	
Glycoursodeoxycholic Acid (GUDCA)							
WOMEN	N Cases (%)	77 (21.6)	104 (29.2)	73 (20.5)	102 (28.7)		0.78
	OR (95%CI)	Reference	1.48 (0.94-2.33)	0.97 (0.60-1.56)	1.40 (0.87-2.25)	0.16	
MEN	N Cases (%)	43 (20.2)	84 (39.4)	36 (16.9)	50 (23.5)		
	OR (95%CI)	Reference	1.96 (1.02-3.75)	0.93 (0.46-1.86)	1.23 (0.61-2.49)	0.41	
Tauroursodeoxycholic Acid (TUDCA)							
WOMEN	N Cases (%)	103 (28.9)	98 (27.5)	48 (13.5)	107 (30.1)		0.63
	OR (95%CI)	Reference	1.02 (0.68-1.55)	0.79 (0.45-1.38)	1.33 (0.86-2.06)	0.14	
MEN	N Cases (%)	50 (23.5)	62 (29.1)	54 (25.4)	47 (22.1)		
	OR (95%CI)	Reference	1.31 (0.72-2.38)	1.09 (0.58-2.07)	0.95 (0.49-1.83)	0.53	

Results from conditional logistic regression analyses;
Conditioned on age, sex, study center, follow-up time, time of day at blood collection, fasting status, menopausal status (women), current use of exogenous hormones (women), and phase of menstrual cycle (women), and adjusted for education level, physical activity level, smoking, waist circumference, height, and consumption of red and processed meat, fat, fiber, and energy;
*Two-sided P value for linear trend from logistic regression model with bile acid values on the log2 scale as the independent variable;
† Two-sided P value for heterogeneity from X^2 tests

Supplementary Table 13 Odds ratios (95 % confidence intervals) for colon cancer among participants in the highest quartile of bile acid levels before and after correction for measurement error

Bile acid metabolite	Pearson's r	Uncorrected Odds Ratio	Corrected Odds Ratio
Primary unconjugated bile acids			
Cholic Acid (CA)	0.55	1.32 (0.90, 1.94)	1.55 (0.81, 3.35)
Chenodeoxycholic Acid (CDCA)	0.63	1.12 (0.77, 1.64)	1.24 (0.70, 2,26)
Hyocholic Acid (HCA)	0.57	0.89 (0.61, 1.31)	0.73 (0.35, 1.38)
Primary conjugated bile acids			
Glycocholic Acid (GCA)	0.45	2.22 (1.52, 3.26)	4.78 (2.01, 29.89)
Taurocholic Acid (TCA)	0.56	1.78 (1.23, 2.58)	2.68 (1.42, 6,47)
Glycochenodeoxycholic Acid (GCDCA)	0.46	1.68 (1.13, 2.48)	2.96 (1.31, 12.24)
Taurochenodeoxycholic Acid (TCDCA)	0.51	1.62 (1.11, 2.36)	2.56 (1.28, 7.33)
Glycohyocholic Acid (GHCA)	0.28	1.65 (1.13, 2.40)	Not estimated*
Tauro-Alpha, Muricholic Acid (TaMCA)	0.50	1.38 (0.98, 1.94)	1.66 (0.87, 3.93)
Secondary unconjugated bile acids			
Deoxycholic Acid (DCA)	0.73	1.11 (0.76, 1.61)	1.23 (0.77, 2.03)
Secondary conjugated bile acids			
Glycodeoxycholic Acid (GDCA)	0.57	1.68 (1.12, 2.54)	2.68 (1.35, 6.70)
Taurodeoxycholic Acid (TDCA)	0.63	1.54 (1.02, 2.31)	2.06 (1.11, 4.26)
Glycolithocholic Acid (GLCA)	0.14	0.91 (0.62, 1.35)	Not estimated*
Tertiary unconjugated bile acids			
Ursodeoxycholic Acid (UDCA)	0.28	0.98 (0.68, 1.41)	Not estimated*
Tertiary conjugated bile acids			
Glycoursodeoxycholic Acid (GUDCA)	0.59	1.36 (0.94, 1.96)	1.41 (0.78, 2.76)
Tauroursodeoxycholic Acid (TUDCA)	0.43	1.26 (0.88, 1.79)	1.50 (0.68, 4.48)

*Correlations at Pearson's coefficients below 0.3 were considered as too low for measurement error correction;