





Prospective Associations of Circulating Bile Acids and Short-Chain Fatty Acids With Incident Colorectal Cancer

Erikka Loftfield, PhD, MPH ^{1,*} Roni T. Falk, MS ¹ Joshua N. Sampson, PhD,² Wen-Yi Huang, PhD ¹ Autumn Hullings, MPH,¹ Gwen Murphy, PhD, MPH,¹ Stephanie J. Weinstein, PhD,¹ Demetrius Albanes, MD,¹ Neal D. Freedman, PhD, MPH ¹ Rashmi Sinha, PhD¹

¹Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute, National Institutes of Health, Rockville, MD, USA; and ²Biostatistics Branch, DCEG, National Cancer Institute, National Institutes of Health, Rockville, MD, USA

*Correspondence to: Erikka Loftfield, PhD, MPH, Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute, National Institutes of Health, NCI 9609 Medical Center Dr, Rockville, MD 20850, USA (e-mail: erikka.loftfield@nih.gov).

Abstract

Background: Human studies investigating the prospective relationship between microbial metabolites and colorectal cancer (CRC) risk are lacking. We tested whether higher serum bile acids (BAs) and lower short-chain fatty acids (SCFAs) were associated with CRC risk. **Methods:** In baseline serum collected more than 30 years before a CRC diagnosis, we quantified concentrations of 15 BAs and 6 SCFAs using targeted liquid chromatography with tandem mass spectrometry assays in 1:1 matched cases and controls from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial (men: n = 262 cases; women: n = 233 cases) and the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (men: n = 598 cases). We estimated odds ratios (ORs) and 95% confidence intervals (CIs) for BA and SCFA quartiles and summary measures with CRC overall and by anatomic location using multivariable conditional logistic regression models. PLCO analyses were stratified by sex. All statistical tests were 2-sided. **Results:** In PLCO women, 7 BAs were strongly associated with increased CRC risk, including the secondary BAs, deoxycholic (OR_{Q4 v Q1} = 2.85, 95% CI = 1.45 to 5.60, $Q_{trend} = 0.011$), glycodeoxycholic (OR_{Q4 v Q1} = 3.45, 95% CI = 1.79 to 6.64, $Q_{trend} = 0.006$), taurodeoxycholic (OR_{Q4 v Q1} = 2.36, 95% CI = 1.22 to 4.55, $Q_{trend} = 0.023$), and glycolithocholic acid (OR_{Q4 v Q1} = 2.71, 95% CI = 1.41 to 5.22, $Q_{trend} = 0.015$). Women in the highest compared with lowest quartile of total SCFAs had a 45% lower risk of CRC (OR = 0.55, 95% CI = 0.31 to 0.98, $P_{trend} = .03$). Associations for total BAs and SCFAs were strongest among women with proximal colon cancer. No statistically significant associations were observed for BA or SCFA measures among men. **Conclusions:** Serum concentrations of BAs, particularly downstream microbial metabolites of cholic acid, were strongly associated with increased risk of CRC among women.

Accumulating evidence from observational studies indicates that the human microbiota contributes to colorectal cancer (CRC) development and that decreased microbial diversity in the colon, assessed in fecal samples, is cross-sectionally associated with CRC (1). Unfortunately, prediagnostic fecal samples for studying prospective associations of the microbiome with CRC risk are not available in most cohorts, but bacteria are metabolically active, and microbial metabolites, which likely mediate associations between the gut microbiota and CRC, can be measured in blood.

Secondary bile acids (BAs) and short-chain fatty acids (SCFAs) are known microbial metabolites that have been linked to dietary intake and colorectal carcinogenesis, but in different

ways. High-fat diets increase secretion of primary BAs, and secondary BAs—namely, deoxycholic (DCA) and lithocholic acid—are formed from primary BAs by *Clostridia* species (2) in the distal small intestine and colon. Experimental evidence indicates that secondary BAs are carcinogenic to the colon (3). In contrast, high-fiber diets increase microbial fermentation products, including the SCFA butyrate. Experimental studies have demonstrated an inhibitory effect of butyrate on colorectal carcinogenesis through suppression of inflammation, inhibition of neoplastic cell proliferation and migration, restriction of tumor angiogenesis, and induction of neoplastic apoptosis (4). Furthermore, dietary fiber has been linked to lower colorectal adenoma and CRC risk in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening

Trial cohort (5). Despite abundant experimental evidence, human studies investigating prospective relationships between BAs, SCFAs, and CRC risk are lacking.

A previous untargeted metabolomics case-control study, nested in PLCO, found that higher serum levels of glycochenodeoxycholic acid were associated with CRC risk among women but not men (6). However, this study was semiquantitative and did not assess SCFAs or associations by anatomic location. A study in the European Prospective Investigation into Cancer and Nutrition found that higher levels of several glycine-conjugated BAs were associated with colon cancer risk, with stronger associations in women, but this study was also semiquantitative and did not include rectal cancer cases or SCFAs (7). Herein, we explore the prospective associations of serum BA and SCFA concentrations with CRC risk overall and by anatomic location.

Materials and Methods

Study Design

The PLCO nested case-control study includes participants from the screening arm of the study, which enrolled more than 39 000 women and more than 38 000 men, aged 55-74 years, in the United States from 1993 to 2001. With follow-up for cancer incidence through December 31, 2014, we identified 495 first primary incident CRC cases (ICD-0-3 codes: C180, C182-C189, C199, C209, C260, excluding ICD morphologies not in the range of 8240-8249, which are carcinoid or neuroendocrine tumors) among those who completed a baseline questionnaire; consented for biospecimen use; were CRC free at study entry according to screening sigmoidoscopy; were free of any cancer, Crohn's disease, ulcerative colitis, familial polyposis, Gardner's syndrome, or colorectal polyps according to baseline questionnaire; did not have a rare cancer during follow-up; had more than 6 months of follow-up; and had unthawed baseline serum ($n = 37\,412$). Cases were individually matched to controls on randomization age (± 5 years), sex, race, randomization year, and blood draw season.

The ATBC nested case-control study includes participants from a Finnish cohort of more than 29 000 male smokers, aged 50-69 years, at enrollment (1985-1988). With follow-up for cancer incidence through December 31, 2015, we identified 598 incident first primary CRC cases (ICD9: 153.0-153.4, 153.6-153.7, and 154.0-154.1) with unthawed serum and without a subsequent diagnosis of a rare cancer. Cases were individually matched to controls on randomization age (± 5 years), serum draw date (± 30 days), and fasting status (< 8 or ≥ 8 hours fasting).

In both studies, eligible controls were alive and cancer free at CRC diagnosis date, and data on risk factors were collected at baseline via self-reported questionnaires; blood was drawn, processed to serum according to standard procedures, and stored at -80°C . The institutional review boards of the US National Cancer Institute and the 10 screening centers approved the PLCO study, and all participants provided informed consent. The institutional review boards of the US National Cancer Institute and the National Public Health Institute of Finland approved the ATBC Study, and all participants provided informed consent.

Metabolomics Analysis

Study and replicate quality control (QC) samples (rate of 5%-10%) were blinded; cases and their matched controls were placed

adjacent but in random order. In PLCO, technical reproducibility was acceptable with inter-batch coefficients of variation (CVs) less than 18% for all BAs with concentrations above the lower limit of quantification (LLOQ), except glycodeoxycholic acid (GDCA; CV = 22%), and all SCFAs, except propionic acid (CV = 38%). In PLCO, concentrations of ursodeoxycholic and tauroursodeoxycholic acid consistently fell below the LLOQ in QC samples. In ATBC, technical reproducibility was also acceptable, with interbatch CVs less than 12% for all BAs and all SCFAs, except isovaleric acid (CV = 23%).

BA and SCFA panels were run on separate aliquots. Serum samples were spiked with corresponding labeled internal standards for each measured compound, subjected to protein precipitation with organic solvent, and centrifuged. Sample analysis was carried out in a 96-well plate containing 2 calibration curves and 6 additional Metabolon QC samples per plate to monitor assay performance. For the BA panel, a portion of the organic supernatant was evaporated to dryness in a stream of nitrogen at 40°C . The dried extract was reconstituted, and an aliquot was injected onto an Agilent 1290 Infinity II/Sciex QTRAP 6500+ or Agilent 1290 Infinity/Sciex QTRAP 6500 liquid chromatography with tandem mass spectrometry system equipped with a C18 reverse-phase ultra-high performance liquid chromatography column. For the SCFA panel, an aliquot of the supernatant was derivatized forming the corresponding SCFA aryl hydrazides. The reaction mixture was diluted and injected onto an Agilent 1290/AB Sciex QTrap 5500 LC MS/MS system equipped with a C18 reversed-phase ultra-high performance liquid chromatography column. For both panels, the mass spectrometer was operated in negative mode using electrospray ionization, and the peak area of each parent or product ion was measured against the peak area of the respective internal standard parent or product ion.

Quantitation was performed with Metabolon, Inc. (Morrisville, North Carolina) using a weighted and linear least squares regression analysis generated from fortified calibration standards prepared immediately before each run. Analyte concentrations that fell below and above the LOQ were extrapolated and flagged. Analytes that could not be extrapolated below the LOQ were assigned a value of one-half the lowest observed concentration for a given analyte within a given study.

Statistical Analysis

The distributions of putative CRC risk factors, BAs, and SCFAs were summarized according to sex and case-control status. Spearman rank correlation coefficients (ρ) between BA and SCFA concentrations were estimated among controls.

Conditional logistic regression models were used to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for CRC across quartiles of analyte concentrations, based on the distribution among controls, and for continuously measured analytes. To compare results from PLCO with ATBC and because previous studies found differences between BA and CRC or colon cancer risk (6), 7 by sex, we stratified PLCO analyses by sex and tested the multiplicative interaction between sex and each analyte by including the cross-product term in logistic regression models; P values for interaction were assessed by χ^2 tests. For categorical analyses, if more than 25% of control samples for a given analyte had concentrations less than the LLOQ, the analyte was defined as quantifiable or not quantifiable. Trends across quartiles were calculated by modeling the ordinal variable; P values were corrected for multiple comparisons using a

false discovery rate–corrected Q value (8). For continuous analyses, because concentrations tended to be right-skewed, analytes were \log_2 -transformed. Thus, a 1-unit increase is interpreted as a doubling in concentration. All models were conditioned on study-specific matching factors and adjusted for suspected CRC risk factors.

In secondary analyses, total measures were created by summing 15 BAs (total BA), 2 unconjugated and 4 conjugated primary BAs (total primary BA), and 2 unconjugated and 4 conjugated secondary BAs (total secondary BA); and 6 SCFAs (total SCFA). Odds ratios and 95% confidence intervals for CRC were estimated using the main analytic model overall and by anatomic location. Models were run excluding cases that occurred within 5 years of blood draw. Finally, we used a stepwise selection process, with an alpha of .05 to enter and .05 to remain in the model, to test each of the 15 BAs and 6 SCFAs and determine the set of metabolites that were independently associated with CRC risk in PLCO, by sex, and in ATBC. Statistical tests were 2-sided. Differences were considered statistically significant if Q values (main analyses) or P values (secondary analyses) were less than .05. SAS 9.4 (Cary, NC, USA) and RStudio Version 1.2.5033 were used for analysis.

Results

In PLCO, cases occurred throughout the follow-up period with 13%, 42%, and 45% of cases occurring within 5 years of baseline, 5–10 years after baseline, and more than 10 years after baseline, respectively. Approximately 69% of cancers among women and men were in the proximal colon, whereas 18% and 13% were in the distal colon and rectum, respectively. The mean time from blood draw to CRC diagnosis was 10 years. Female and male CRC cases were less likely than controls to report having graduated from college or being never smokers. Among female cases, total energy intake and body mass index tended to be higher than controls, and among male cases, unprocessed red and processed meat intake, total energy intake, and body mass index tended to be higher than among controls. Distributions of other putative risk factors appeared similar. Among female cases, mean BA concentrations were consistently higher than among controls (Table 1). In ATBC, cases also occurred throughout the follow-up period, with 8%, 16%, and 76% of cases occurring within 5 years of baseline, 5–10 years after baseline, and more than 10 years after baseline, respectively. In contrast to PLCO, 30% of cancers occurred in the proximal colon, and 24% and 46% occurred in the distal colon and rectum, respectively. The mean time from blood draw to CRC diagnosis was 15 years. Distributions of putative risk factors generally appeared similar between the 2 groups. Mean BA concentrations tended to be higher among cases than controls (Table 1).

In PLCO, BAs were generally positively correlated ($\rho > 0.57$ for BAs within the same classification). All SCFAs were positively correlated ($0.12 \leq \rho \leq 0.63$). Acetic acid was negatively correlated with most BAs (Supplementary Figure 1, available online). In ATBC, most BAs were also generally positively correlated with each other ($\rho \geq 0.50$ for BAs within the same classification), and all SCFAs were positively correlated. As in PLCO, acetic acid was negatively correlated with most BAs (Supplementary Figure 2, available online).

In PLCO women, higher concentrations of 7 BAs were associated with higher CRC risk (all $Q_{\text{trend}} < 0.05$; Figure 1; Supplementary Table 1, available online). Notably, women in the highest quartile of DCA had nearly threefold higher odds of CRC compared with those in the lowest quartile (DCA; $OR_{Q4 \text{ v } Q1} = 2.85$,

95% CI = 1.45 to 5.60, $Q_{\text{trend}} = 0.011$). The strongest and most consistent associations were observed for conjugated secondary BAs, with those in the highest quartile of GDCA ($OR_{Q4 \text{ v } Q1} = 3.45$, 95% CI = 1.79 to 6.64, $Q_{\text{trend}} = 0.006$), taurodeoxycholic acid (TDCA; $OR_{Q4 \text{ v } Q1} = 2.36$, 95% CI = 1.22 to 4.55, $Q_{\text{trend}} = 0.023$), and glycolithocholic acid ($OR_{Q4 \text{ v } Q1} = 2.71$, 95% CI = 1.41 to 5.22, $Q_{\text{trend}} = 0.015$) having two- to threefold higher odds of CRC. Additionally, a quantifiable concentration of tauroolithocholic acid was also associated with higher odds of CRC (tauroolithocholic acid; $OR_{\text{Quantifiable v not quantifiable}} = 1.84$, 95% CI = 1.16 to 2.93, $Q_{\text{trend}} = 0.047$). Finally, 2 primary conjugated BAs, taurochenodeoxycholic acid ($OR_{Q4 \text{ v } Q1} = 2.98$, 95% CI = 1.42 to 6.24, $Q_{\text{trend}} = 0.047$) and glycocholic acid ($OR_{Q4 \text{ v } Q1} = 2.26$, 95% CI = 1.21 to 4.22, $Q_{\text{trend}} = 0.022$), were associated with higher odds of CRC. Consistent findings were observed for total BA concentrations, with women in the highest quartile having twofold higher odds of CRC ($OR_{Q4 \text{ v } Q1} = 1.95$, 95% CI = 1.04 to 3.66, $P_{\text{trend}} = .01$); associations for total primary BAs and total secondary BAs with CRC were stronger (Figure 2; Supplementary Table 2, available online). Stratification by cancer anatomic location revealed that associations with total BA concentrations were strongest among women with proximal colon cancer (Table 2).

Individual BAs were not associated with CRC among men in either PLCO (Figure 1; Supplementary Table 1, available online) or ATBC (Figure 1; Supplementary Table 3, available online) after correcting for multiple comparisons; however, a borderline statistically significant association was observed for GDCA in ATBC ($OR_{Q4 \text{ v } Q1} = 1.54$, 95% CI = 1.07 to 2.21, $Q_{\text{trend}} = 0.06$). No statistically significant associations between BA concentration totals and CRC were observed among men in PLCO (Figure 2; Supplementary Table 2, available online), but in ATBC, men in the highest quartile of total secondary BAs had higher odds of CRC ($OR_{Q4 \text{ v } Q1} = 1.47$, 95% CI = 1.04 to 2.07, $P_{\text{trend}} = .03$) (Figure 2; Supplementary Table 2, available online). In women and men, associations of BA concentration totals were generally similar when excluding cases that occurred within 5 years of blood draw in both cohorts (Supplementary Tables 4 and 5, available online).

In both PLCO and ATBC, individual SCFAs were not associated with CRC (all $Q_{\text{trend}} \geq 0.20$; Figure 3; Supplementary Table 6, available online). However, women in the highest quartile of total SCFAs had 45% lower odds of CRC ($OR_{Q4 \text{ v } Q1} = 0.55$, 95% CI = 0.31 to 0.98, $P_{\text{trend}} = .03$) (Figure 2; Supplementary Table 2, available online). Again, stratification by anatomic location revealed that the inverse associations for total SCFAs were strongest among women with proximal colon cancer (Table 2). In ATBC, associations for total SCFA concentrations with CRC were not statistically significant and generally reflected those among men in PLCO (Supplementary Table 2, available online). Total SCFA associations were generally similar when excluding cases that occurred within 5 years of blood draw in both cohorts (Supplementary Tables 4 and 5, available online).

In a secondary analysis using a stepwise selection process (Supplementary Table 7, available online), we found that among PLCO women, GDCA ($OR_{\log_2[\text{GDCA}]} = 1.29$, 95% CI = 1.10 to 1.50, $P = .003$) and isovaleric acid ($OR_{\log_2[\text{isovaleric acid}]} = 0.68$, 95% CI = 0.48 to 0.96, $P = .03$) were independently associated with CRC risk. Among men in ATBC, GDCA was also positively and independently associated with CRC risk ($OR_{\log_2[\text{GDCA}]} = 1.10$, 95% CI = 1.12 to 1.18, $P = .01$). In contrast, among men in PLCO, the strongest associations were for TDCA, which, like its glycine conjugate, was positively associated with CRC risk ($OR_{\log_2[\text{TDCA}]} = 1.06$, 95% CI = 1.00 to 1.13, $P = .04$), and ursodeoxycholic acid, which was inversely associated with CRC risk ($OR_{\log_2[\text{UDCA}]} = 0.94$, 95% CI = 0.89 to 0.99, $P = .02$).

Table 1. Baseline characteristics of CRC cases and controls in the PLCO Cancer Screening Trial Cohort by sex and the ATBC Cohort^a

Characteristics	PLCO Women (n = 466)		PLCO Men (n = 524)		ATBC Men (n = 1196)	
	Controls (n = 233)	Cases (n = 233)	Controls (n = 262)	Cases (n = 262)	Controls (n = 598)	Cases (n = 598)
Mean age at baseline (SD), y ^{b,c}	64.1 (5.4)	64.2 (5.3)	64.0 (5.1)	64.0 (5.0)	56.6 (4.8)	56.6 (4.8)
Race, % (No.)						
Asian	0.9 (2)	0.9 (2)	5.0 (13)	5.0 (13)	0 (0)	0 (0)
Black, Non-Hispanic	5.6 (13)	5.6 (13)	3.8 (10)	3.8 (10)	0 (0)	0 (0)
Hispanic, American Indian, or Pacific Islander	1.3 (3)	1.3 (3)	2.7 (7)	2.7 (7)	0 (0)	0 (0)
Non-Hispanic white ^p	92.3 (215)	92.3 (215)	88.5 (232)	88.5 (232)	100 (598)	100 (598)
College graduate, % (No.)	34.4 (80)	23.3 (54)	43.3 (114)	37.6 (99)	4.4 (26)	5.1 (31)
Current smoker, % (No.)	5.3 (12)	9.1 (21)	8.4 (22)	11.4 (30)	100 (598)	100 (598)
Self-reported diabetes, % (No.)	4.9 (11)	4.8 (11)	7.3 (19)	10.8 (28)	3.1 (19)	2.5 (15)
Family history of colorectal cancer, % (No.)	14.3 (33)	12.7 (30)	9.1 (24)	13.2 (35)	—	—
Vigorous activity, 4+ h/wk, % (No.)	22.5 (52)	21.2 (49)	24.5 (64)	24.5 (64)	—	—
Leisure time physical activity, 3+ times/wk, % (No.)	—	—	—	—	19.9 (119)	17.7 (106)
Never HRT user (women only), % (No.)	31.1 (72)	32.7 (76)	—	—	—	—
Alcohol drinker, % (No.)	75.1 (175)	70.3 (164)	83.3 (218)	86.4 (226)	90.8 (543)	94.7 (566)
Mean alcohol intake (SD), drinks/d	0.4 (0.8)	0.5 (1.1)	1.3 (1.9)	1.4 (2.9)	1.3 (1.5)	1.5 (1.6)
Mean whole grain intake (SD), oz/d/1000 kcal	0.6 (0.4)	0.6 (0.4)	0.5 (0.4)	0.5 (0.3)	2.1 (1.4)	2.0 (1.5)
Mean unprocessed red meat intake (SD), g/d/1000 kcal	22.2 (13.9)	22.0 (14.5)	29.6 (16.7)	33.0 (18.9)	27.1 (11.8)	28.0 (11.8)
Mean processed meat intake (SD), g/d/1000 kcal	4.3 (4.5)	4.3 (5.8)	6.9 (6.6)	8.6 (9.3)	28.1 (17.9)	28.4 (17.8)
Mean energy intake (SD), kcal/d	1674.9 (595.3)	1753.7 (675.0)	2254.9 (948.2)	2439.5 (925.4)	2725.4 (688.6)	2691.7 (704.4)
Mean body mass index (SD), kg/m ²	26.8 (5.4)	27.6 (5.4)	27.1 (3.8)	28.6 (4.7)	26.1 (3.7)	26.4 (3.7)
Bile acids, Mean (SD)						
Chenodeoxycholic acid, ng/mL	61.2 (117.4)	115.4 (557.6)	123.4 (329.9)	107.2 (207.3)	132.0 (257.0)	151.1 (377.8)
Cholic acid, ng/mL	28.1 (78.6)	55.2 (242.1)	70.1 (209.7)	43.9 (92.9)	94.7 (192.8)	118.2 (462.5)
Deoxycholic acid, ng/mL	115.7 (123.5)	154.5 (203.6)	138.1 (210.2)	135.1 (145.5)	111.7 (156.9)	124.5 (164.7)
Lithocholic acid, ng/mL	4.2 (5.5)	5.3 (8.4)	4.5 (5.5)	5.8 (7.6)	3.0 (2.9)	3.3 (4.1)
Ursodeoxycholic acid, ng/mL	17.1 (31.7)	32.8 (165.1)	20.4 (40.6)	19.0 (34.8)	22.6 (38.9)	23.7 (41.9)
Glycochenodeoxycholic acid, ng/mL	263.5 (290.0)	359.3 (451.3)	388.5 (534.0)	409.9 (558.1)	206.1 (237.8)	234.6 (305.5)
Glycocholic acid, ng/mL	114.9 (199.0)	145.7 (408.1)	120.2 (180.0)	168.4 (366.3)	51.5 (95.0)	65.2 (121.5)
Glycodeoxycholic acid, ng/mL	140.4 (179.7)	205.8 (259.1)	170.9 (291.6)	177.8 (244.8)	53.9 (79.7)	65.1 (97.6)
Glycolithocholic acid, ng/mL	7.2 (13.2)	9.4 (14.6)	10.0 (17.6)	9.3 (13.1)	3.9 (5.2)	4.1 (5.0)
Glycoursodeoxycholic acid, ng/mL	48.6 (80.2)	60.9 (121.2)	47.3 (65.5)	49.3 (75.4)	36.6 (45.6)	37.2 (49.4)
Taurochenodeoxycholic acid, ng/mL	54.8 (81.3)	68.0 (115.8)	48.8 (79.7)	72.7 (165.5)	26.0 (32.5)	36.6 (128.2)
Taurocholic acid, ng/mL	27.9 (80.5)	30.7 (105.2)	14.9 (31.9)	32.0 (111.3)	8.5 (28.0)	14.9 (91.3)
Taurodeoxycholic acid, ng/mL	30.7 (47.7)	40.8 (61.7)	21.2 (41.6)	29.5 (57.5)	9.0 (21.0)	10.3 (16.4)
Tauroolithocholic acid, ng/mL	1.2 (2.3)	1.4 (2.8)	1.2 (2.6)	1.5 (3.1)	1.4 (1.5)	1.4 (0.7)
Tauroursodeoxycholic acid, ng/mL	2.2 (4.6)	2.6 (6.2)	1.3 (3.6)	2.3 (8.9)	1.7 (1.5)	1.7 (1.9)
Short-chain fatty acids, Mean (SD)						
Acetic acid, ng/mL	1747.5 (1455.0)	1470.4 (1227.8)	1661.1 (1458.3)	2264.4 (5601.9)	3223.2 (8269.5)	3165.2 (7084.3)
Butyric acid, ng/mL	31.5 (24.2)	36.3 (31.2)	31.1 (27.7)	39.9 (74.4)	38.8 (19.0)	39.0 (23.1)
Propionic acid, ng/mL	41.7 (40.5)	48.7 (51.4)	42.6 (40.7)	49.6 (69.3)	75.6 (36.0)	76.9 (41.3)
Hexanoic acid, ng/mL	45.4 (21.1)	45.8 (22.2)	40.7 (17.0)	42.3 (18.8)	84.5 (23.8)	86.3 (26.5)
Isobutyric acid, ng/mL	72.6 (84.3)	82.8 (122.5)	71.7 (74.6)	76.3 (112.9)	95.5 (25.0)	96.1 (26.3)
Isovaleric acid, ng/mL	59.2 (30.5)	56.5 (29.2)	72.6 (32.8)	75.4 (40.2)	84.9 (30.4)	86.0 (33.8)

^aValues are standardized to the age distribution of the study population. ATBC = Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; CRC = colorectal cancer; PLCO = Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial.

^bMatching factor in PLCO.

^cMatching factor in ATBC.

Discussion

Compelling experimental evidence links BAs and SCFAs, 2 classes of gut microbial metabolites, to CRC etiology (3,4,9,10). Secondary BAs have been shown to cause DNA damage (9) and to induce colon cancer stem cells (11), whereas butyrate and other SCFAs have been shown to inhibit carcinogenesis via multiple mechanisms (4). Supporting such potential causal relationships, we observed associations for higher serum concentrations of nearly all downstream metabolites of cholic acid, including DCA, the bacteria-derived secondary BA, and its glycine and taurine conjugated counterparts with increased

CRC risk among women. Total serum SCFA concentrations were associated with lower CRC risk among women. However, we observed limited evidence for associations for BAs or SCFAs among men. Tests for multiplicative interactions between sex and SCFAs or BAs in PLCO were generally not statistically significant, but sex differences in CRC etiology, particularly by anatomic location, are supported by the literature.

Foremost, 2 previous studies found evidence that the associations between prediagnostic serum BA levels and CRC vary by sex. First, an untargeted metabolomics case-control study in PLCO, which overlapped with but had fewer CRC cases than this

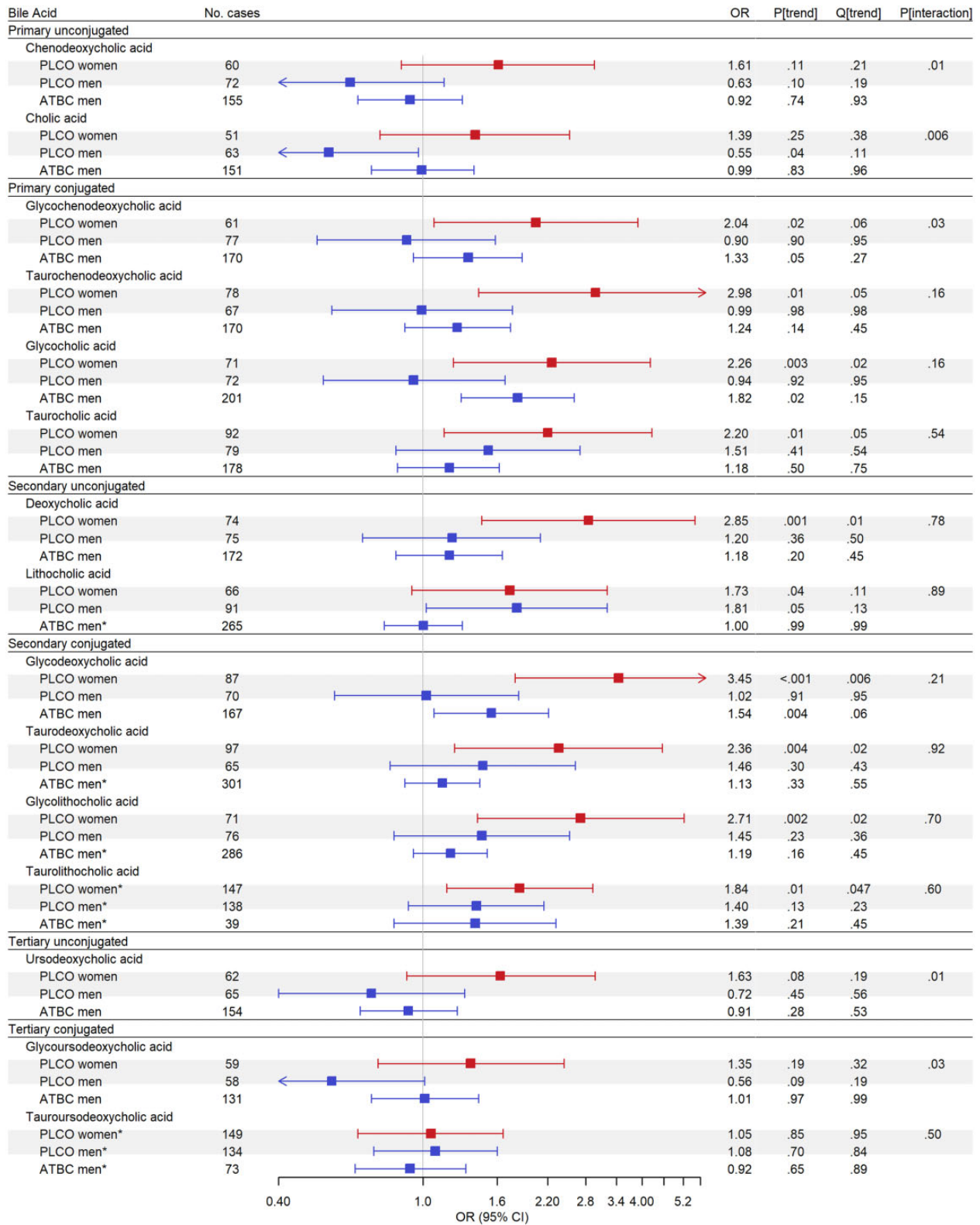


Figure 1. Odds ratios (ORs) for bile acids (BA) with colorectal cancer (CRC) risk, estimated by means of conditional logistic regression to account for matching, for individuals in the highest quartile (Q4) of BA concentrations compared with those in the lowest quartile (Q1) based on the BA distribution in controls for either Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) or Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC). Asterisks indicate odds ratios for individuals with quantifiable concentrations compared with those with concentrations below the limit of quantification for analytes with more than 25% of concentrations flagged as not quantifiable among controls. **Red squares** show odds ratios for women, **blue squares** show odds ratios for men, and **grey shaded boxes** highlight analyses run in PLCO. Analyses in PLCO were adjusted for age at baseline (years), body mass index (BMI; kg/m²), smoking status (never, former, current), family history of CRC (no, yes), diabetes diagnosis (no, yes), alcohol drinker (no, yes), alcoholic intake (average number of drinks per day), hours spent in vigorous physical

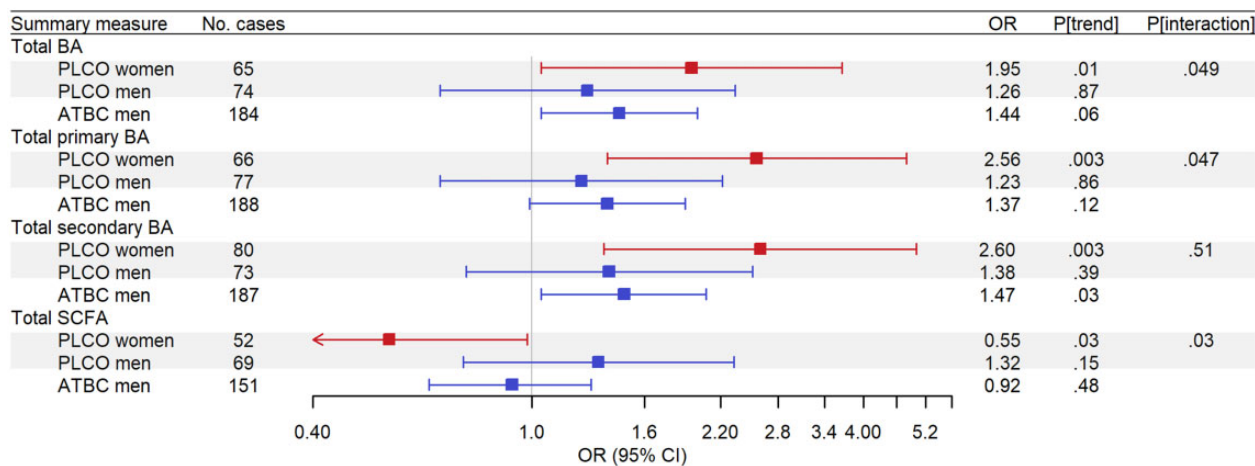


Figure 2. Odds ratios (ORs) for summary bile acid (BA) and short-chain fatty acid (SCFA) concentrations with colorectal cancer (CRC) risk, estimated by means of conditional logistic regression to account for matching, for individuals in the highest quartile (Q4) of a given summary measure compared with those in the lowest quartile (Q1), based on the distribution in controls for either Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial cohort (PLCO) or Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC). **Red squares** show odds ratios for women, **blue squares** show odds ratios for men, and **gray shaded boxes** highlight analyses run in the PLCO. Analyses in the PLCO were adjusted for age at baseline (years), body mass index (BMI; kg/m²), smoking status (never, former, current), family history of CRC (no, yes), diabetes diagnosis (no, yes), alcohol drinker (no, yes), alcoholic intake (average number of drinks per day), hours spent in vigorous physical activity (none, <1h/wk, 1h/wk, 2h/wk, 3h/wk, ≥4h/wk), total unprocessed red meat (g/d/1000 kcal), total processed meat (g/d/1000 kcal), whole grains (oz/d/1000 kcal), total energy (kcal/d), education (high school graduate or less, some college or post high school training, college graduate or postgraduate), and hormone replacement therapy status (women only; never, current, former). Analyses in ATBC were adjusted for age at baseline (years), BMI (kg/m²), smoking intensity (cigarettes per day), diabetes diagnosis (no, yes), average number of alcoholic drinks per day, frequency of physical activity in leisure time (none, <1/wk, 1–2/wk, ≥3/wk), total unprocessed red meat (g/d/1000 kcal), total processed meat (g/day/1000 kcal), whole grains (oz/d/100 kcal), total energy (kcal/d), and education (less than elementary school, some junior high school, completed junior high school, some senior high school, senior high school graduate).

study, reported an association between glycochenodeoxycholic acid and CRC risk among women but not men (6). Second, a nested case-control study in the European Prospective Investigation into Cancer and Nutrition found that higher concentrations of several glycine-conjugated BAs were associated with increased risk of colon cancer, with stronger associations in women; odds ratios from sex-stratified analyses were similar in magnitude to odds ratios in our study (7).

Sex differences in CRC are well documented regarding cancer incidence, anatomic location, molecular features, and risk factors (12–14). In the United States, the age-adjusted CRC incidence rate is higher among men than women (15). The largest disparity is for rectal cancer, with a male-to-female incidence rate ratio of 1.62, and the smallest is for proximal colon cancer (incidence rate ratio = 1.07) (12). This is explained, in part, by the higher proportion of proximal tumors in women and the higher proportions of distal and rectal tumors in men (12,16). Moving from the proximal colon, which derives from the embryonic midgut, to the distal colon and rectum, which derive from the hindgut, researchers have found that tumor pathologic and molecular features change, with frequencies of high levels of microsatellite instability, CpG island methylator phenotype, and BRAF mutations decreasing more than 10-fold from the proximal colon to rectum (17), and proximal tumors more often being mucinous and having serrated pathway signatures (18).

Higher rates of microsatellite instability, CpG island methylator phenotype, and BRAF-mutated tumors have similarly been observed in women (13). Regarding BA specifically, a small-nested case-control study found that higher fecal CDCA concentrations were associated with a more than 6-fold increased odds in proximal but not distal colon cancer (19). Such differences by anatomic site indicate not only that CRC is a heterogeneous disease but also that the contents of the bowel, including food waste, microbes, and microbial metabolites, which vary across subsites, likely contribute to different etiologies (17).

A compelling line of evidence for sex and site differences in BA-CRC associations comes from cohort studies of cholecystectomy and CRC risk. The gallbladder stores primary BAs, thereby reducing the BA concentrations that escape enterohepatic circulation into the intestine and reducing the rate of formation of secondary BAs (20). A meta-analysis of 10 cohort studies found that cholecystectomy was associated with a 30% increase in colon cancer but not rectal cancer risk, and in stratified analyses, the statistically significant associations between cholecystectomy and CRC and colon cancer risk were observed only among women (21).

Despite numerous observational studies of fiber and CRC risk (22–25) and experimental studies on the anticarcinogenic potential of SCFAs (3,4,26), we observed limited evidence for an association of serum SCFAs with CRC in women and no

activity (none, <1 h/wk, 1 h/wk, 2 h/wk, 3 h/wk, ≥4 h/wk), total unprocessed red meat (g/d/1000 kcal), total processed meat (g/d/1000 kcal), whole grains (oz/d/1000 kcal), total energy (kcal/d), education (high school graduate or less, some college or post high school training, college graduate or postgraduate), and hormone replacement therapy status (women only; never, current, former). Analyses in ATBC were adjusted for age at baseline (years), BMI (kg/m²), smoking intensity (cigarettes per d), diabetes diagnosis (no, yes), average number of alcoholic drinks per day, frequency of physical activity in leisure time (none, <1/wk, 1–2/wk, ≥3/wk), total red meat (g/d/1000 kcal), total processed meat (g/d/1000 kcal), whole grains (oz/d/100 kcal), total energy (kcal/d), and education (less than elementary school, some junior high school, completed junior high school, some senior high school, senior high school graduate).

Table 2. Association of summary BA and SCFA concentrations with CRC overall and by tumor site in the PLCO Study and ATBC Study

	CRC						Proximal colon cancer			Distal colon cancer			Rectal cancer		
	No. cases	OR (95% CI) [measure]	P _{int} for sex ^c	No. cases	OR (95% CI) [measure]	P _{int} for sex ^c	No. cases	OR (95% CI) [measure]	P _{int} for sex ^c	No. cases	OR (95% CI) [measure]	P _{int} for sex ^c	No. cases	OR (95% CI) [measure]	P _{int} for sex ^c
Total BA															
PLCO women ^a	233	1.31 (1.11 to 1.53)	.049	160	1.40 (1.13 to 1.74)	.009	42	0.82 (0.52 to 1.31)	.06	29	1.13 (0.84 to 1.52)	.29	29	1.13 (0.84 to 1.52)	.29
PLCO men ^a	262	1.05 (0.91 to 1.21)	—	180	1.15 (0.96 to 1.39)	—	44	1.05 (0.67 to 1.62)	—	35	0.86 (0.66 to 1.12)	—	35	0.86 (0.66 to 1.12)	—
ATBC men ^b	598	1.10 (1.00 to 1.20)	—	181	1.20 (0.99 to 1.45)	—	143	0.98 (0.80 to 1.19)	—	274	1.09 (0.94 to 1.27)	—	274	1.09 (0.94 to 1.27)	—
Total primary BA															
PLCO women ^a	233	1.22 (1.06 to 1.41)	.047	160	1.27 (1.06 to 1.53)	.03	42	0.80 (0.51 to 1.24)	.07	29	1.23 (0.93 to 1.62)	.05	29	1.23 (0.93 to 1.62)	.05
PLCO men ^a	262	1.01 (0.89 to 1.15)	—	180	1.09 (0.93 to 1.28)	—	44	0.97 (0.64 to 1.45)	—	35	0.81 (0.63 to 1.03)	—	35	0.81 (0.63 to 1.03)	—
ATBC men ^b	598	1.07 (0.99 to 1.16)	—	181	1.19 (1.01 to 1.41)	—	143	1.01 (0.84 to 1.22)	—	274	1.03 (0.91 to 1.16)	—	274	1.03 (0.91 to 1.16)	—
Total secondary BA															
PLCO women ^a	233	1.14 (1.02 to 1.27)	.51	160	1.16 (1.00 to 1.33)	.14	42	1.12 (0.79 to 1.58)	.33	29	1.00 (0.83 to 1.19)	.54	29	1.00 (0.83 to 1.19)	.54
PLCO men ^a	262	1.07 (0.98 to 1.18)	—	180	1.10 (0.98 to 1.23)	—	44	1.27 (0.86 to 1.88)	—	35	1.03 (0.85 to 1.24)	—	35	1.03 (0.85 to 1.24)	—
ATBC men ^b	598	1.09 (1.00 to 1.19)	—	181	1.07 (0.90 to 1.27)	—	143	0.93 (0.77 to 1.12)	—	274	1.23 (1.07 to 1.41)	—	274	1.23 (1.07 to 1.41)	—
Total SCFA															
PLCO women ^a	233	0.77 (0.61 to 0.98)	.03	160	0.68 (0.50 to 0.93)	.002	42	1.31 (0.56 to 3.03)	.72	29	1.01 (0.66 to 1.53)	.43	29	1.01 (0.66 to 1.53)	.43
PLCO men ^a	262	1.20 (0.97 to 1.48)	—	180	1.12 (0.88 to 1.44)	—	44	1.88 (0.64 to 5.47)	—	35	0.93 (0.60 to 1.44)	—	35	0.93 (0.60 to 1.44)	—
ATBC men ^b	598	0.98 (0.87 to 1.09)	—	181	0.82 (0.64 to 1.04)	—	143	1.13 (0.87 to 1.47)	—	274	1.00 (0.85 to 1.20)	—	274	1.00 (0.85 to 1.20)	—

^aAdjusted for age at baseline (years), BMI (kg/m²), smoking status (never, former, current), family history of CRC (no, yes), alcohol intake (average number of drinks per day), hours spent in vigorous physical activity (none, <1 h/wk, 1 h/wk, 2 h/wk, 3 h/wk, ≥4 h/wk), not processed red meat (g/d/1000 kcal), processed meat (g/d/1000 kcal), whole grains (oz/d/1000 kcal), total energy (kcal/d), education (high school graduate or less, some college or post high school training, college graduate or postgraduate), hormone replacement therapy status (women only, never, current, former), ATBC = Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study; BA = bile acids; BMI = body mass index; CI = confidence interval; CRC = colorectal cancer; OR = odds ratio; PLCO = Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial; SCFA = short-chain fatty acids.

^bAdjusted for age at baseline (years), BMI (kg/m²), smoking intensity (cigarettes/day), diabetes diagnosis (no, yes), average number of alcoholic drinks per day, frequency of physical activity in leisure time (none, <1/wk, 1–2/wk, ≥3/wk), total red meat (g/d/1000 kcal), total processed meat (g/d/1000 kcal), whole grains (oz/d/1000 kcal), total energy (kcal/d), education (less than elementary school, some junior high school, completed junior high school, some senior high school, senior high school graduate).

^cP_{interaction} (ie, P_{int}) between log_e-transformed bile acid measure and sex estimated in unconditional logistic regression model including sex as a covariate in the model in PLCO analyses only.

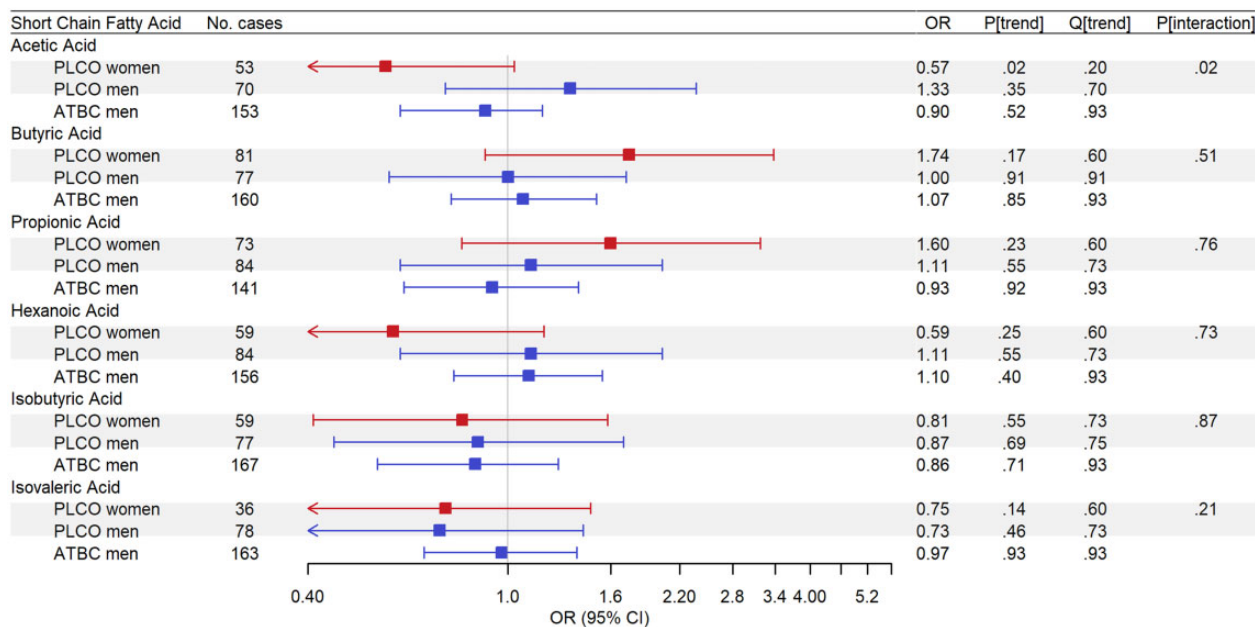


Figure 3. Odds ratios (ORs) for short-chain fatty acids (SCFA) with colorectal cancer (CRC) risk, estimated by means of conditional logistic regression to account for matching, for individuals in the highest quartile (Q4) of SCFA concentrations compared with those in those in the lowest quartile (Q1), based on the SCFA distribution in controls for either Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) or Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC). **Red squares** show odds ratios for women, **blue squares** show odds ratios for men, and **grey shaded boxes** highlight analyses run in the PLCO Cancer Screening Trial cohort. Analyses in PLCO were adjusted for age at baseline (years), body mass index (BMI; kg/m²), smoking status (never, former, current), family history of CRC (no, yes), diabetes diagnosis (no, yes), alcohol drinker (no, yes), alcoholic intake (average number of drinks per day), hours spent in vigorous physical activity (none, <1 h/wk, 1 h/wk, 2 h/wk, 3 h/wk, ≥4 h/wk), total unprocessed red meat (g/d/1000 kcal), total processed meat (g/d/1000 kcal), whole grains (oz/d/1000 kcal), total energy (kcal/d), education (high school graduate or less, some college or post high school training, college graduate or postgraduate), and hormone replacement therapy status (women only; never, current, former). Analyses in ATBC were adjusted for age at baseline (years), BMI (kg/m²), smoking intensity (cigarettes per day), diabetes diagnosis (no, yes), average number of alcoholic drinks per day, frequency of physical activity in leisure time (none, <1/wk, 1–2/wk, ≥3/wk), total unprocessed red meat (g/d/1000 kcal), total processed meat (g/day/1000 kcal), whole grains (oz/d/100 kcal), total energy (kcal/d), and education (less than elementary school, some junior high school, completed junior high school, some senior high school, senior high school graduate).

evidence in men. Supportive, albeit limited, evidence indicates that fecal SCFA concentrations are lower among those with colorectal adenoma cases (27,28) and that a lower abundance of butyrate-producing bacteria are found in CRC cases compared with healthy controls (29). The lack of association observed in this study for the individual SCFAs, in contrast to the presence of associations for BAs, may be explained in part by weak correlations between fecal and serum SCFAs and stronger associations between fecal and serum BAs, particularly secondary BAs. A recent study found no or weak correlations between serum and fecal SCFAs, ranging from -0.05 for acetic acid to 0.24 for valeric acid, but moderate correlations between serum and fecal secondary (ie, DCA, $r=0.61$) and tertiary (ie, ursodeoxycholic acid, $r=0.50$) BAs, which are microbially derived (30).

Strengths of our study include its prospective nature, which allowed us to evaluate the temporality between metabolites and CRC onset. Similar results for BA concentrations after excluding cases that occurred during the first 5 years of follow-up are consistent with a possible role for circulating BAs in CRC independent of underlying disease. Furthermore, PLCO participants were screened using flexible sigmoidoscopy at baseline and 3–5 years later; those with a screen-detected polyp or mass were referred for diagnostic follow-up (31). CRC screening in PLCO likely influenced the distribution of cancers by anatomic location. With 495 CRC cases in PLCO, we were able to stratify by sex and explore potential sex differences by tumor location. We found no association for serum BAs and SCFAs with CRC risk among men, indicating that sex differences in CRC etiology may extend to BAs and SCFAs, as well as more broadly to diet,

the microbiome, and other metabolites. Nevertheless, future research is needed to replicate our findings in diverse populations and to explore possible associations between BAs and CRC risk factors such as insulin resistance as well as the role of endogenous hormones, such as estrogen, in BA synthesis and metabolism. Unfortunately, as in most cohort studies, fecal samples were not collected; thus, we were unable to measure fecal metabolites or the fecal microbiome.

In summary, we found that among women, total circulating BA and SCFA concentrations were associated with increased and decreased risk of CRC, respectively. The individual BAs most strongly associated with increased CRC risk comprised the downstream metabolites of CA, including the microbially derived DCA, which has been shown repeatedly to be carcinogenic in experimental studies.

Funding

This work was supported by the NCI Intramural Research Program.

Notes

Role of the funder: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation of the manuscript; and decision to submit the manuscript for publication.

Disclosures: The authors have no potential conflicts of interest to disclose.

Author contributions: Conceptualization: EL, JNS, WYH, GM, SJW, DA, NDF, and RS; Methodology: EL, JNS; Software: EL; Formal analysis: EL, RT, JNS; Investigation: EL, WYH, NDF, DA, SJW; Data curation: EL, WYH, SJW; Writing—Original Draft: EL; Writing—Review and Editing: all authors; Visualization: EL; Supervision: RS, NDF; Project administration: EL; Funding acquisition: EL, RS, NDF.

Acknowledgements: The authors would like to acknowledge that cancer incidence data have been provided by the Alabama Statewide Cancer Registry, Arizona Cancer Registry, Colorado Central Cancer Registry, District of Columbia Cancer Registry, Georgia Cancer Registry, Hawaii Cancer Registry, Cancer Data Registry of Idaho, Maryland Cancer Registry, Michigan Cancer Surveillance Program, Minnesota Cancer Surveillance System, Missouri Cancer Registry, Nevada Central Cancer Registry, Ohio Cancer Incidence Surveillance System, Pennsylvania Cancer Registry, Texas Cancer Registry, Utah Cancer Registry, Virginia Cancer Registry, and Wisconsin Cancer Reporting System. All are supported in part by funds from the Center for Disease Control and Prevention, National Program for Central Registries, local states or by the National Cancer Institute, Surveillance, Epidemiology, and End Results program.

Disclaimers: The results reported here, and the conclusions derived are the sole responsibility of the authors.

Data Availability

Data can be requested via PLCO (<https://cdas.cancer.gov/plco/>) and ATBC (<https://atbcstudy.cancer.gov/pts/a/>) data access systems. Analytic methods and study materials will be available upon reasonable request to corresponding author.

References

- Ahn J, Sinha R, Pei Z, et al. Human gut microbiome and risk for colorectal cancer. *J Natl Cancer Inst.* 2013;105(24):1907–1911.
- Wells JE, Hylemon PB. Identification and characterization of a bile acid 7 α -dehydroxylation operon in *Clostridium* sp. strain TO-931, a highly active 7 α -dehydroxylating strain isolated from human feces. *Appl Environ Microbiol.* 2000;66(3):1107–1113.
- Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol.* 2014;12(10):661–672.
- O’Keefe SJ. Diet, microorganisms and their metabolites, and colon cancer. *Nat Rev Gastroenterol Hepatol.* 2016;13(12):691–706.
- Kunzmann AT, Coleman HG, Huang WY, et al. Dietary fiber intake and risk of colorectal cancer and incident and recurrent adenoma in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *Am J Clin Nutr.* 2015;102(4):881–890.
- Cross AJ, Moore SC, Boca S, et al. A prospective study of serum metabolites and colorectal cancer risk. *Cancer.* 2014;120(19):3049–3057.
- Kuhn T, Stepien M, Lopez-Nogueroles M, et al. Prediagnostic plasma bile acid levels and colon cancer risk: a prospective study. *J Natl Cancer Inst.* 2020;112(5):516–524.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B-Stat Methodol.* 1995;57(1):289–300.
- Bernstein H, Bernstein C, Payne CM, et al. Bile acids as carcinogens in human gastrointestinal cancers. *Mutat Res.* 2005;589(1):47–65.
- Ocvirk S, Wilson AS, Appolonia CN, et al. Fiber, fat, and colorectal cancer: new insight into modifiable dietary risk factors. *Curr Gastroenterol Rep.* 2019;21(11):62.
- Farhana L, Nangia-Makker P, Arbit E, et al. Bile acid: a potential inducer of colon cancer stem cells. *Stem Cell Res Ther.* 2016;7(1):181.
- Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. *CA A Cancer J Clin.* 2020;70(3):145–164.
- White A, Ironmonger L, Steele RJC, et al. A review of sex-related differences in colorectal cancer incidence, screening uptake, routes to diagnosis, cancer stage and survival in the UK. *BMC Cancer.* 2018;18(1):906.
- Koo JH, Leong RW. Sex differences in epidemiological, clinical and pathological characteristics of colorectal cancer. *J Gastroenterol Hepatol.* 2010;25(1):33–42.
- Howlander N, Noone AM, Krapcho M, et al., eds. *SEER Cancer Statistics Review, 1975–2018.* Bethesda, MD: National Cancer Institute; 2020. https://seer.cancer.gov/csr/1975_2018/. Accessed April 2021.
- Murphy G, Devesa SS, Cross AJ, et al. Sex disparities in colorectal cancer incidence by anatomic subsite, race and age. *Int J Cancer.* 2011;128(7):1668–1675.
- Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut.* 2012;61(6):847–854.
- Missiaglia E, Jacobs B, D’Ario G, et al. Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features. *Ann Oncol.* 2014;25(10):1995–2001.
- Haines A, Hill MJ, Thompson MH, et al. A prospective study of faecal bile acids and colorectal cancer. *Eur J Cancer Prev.* 2000;9(5):317–323.
- Turumin JL, Shanturov VA, Turumina HE. The role of the gallbladder in humans. *Rev Gastroenterol Mex.* 2013;78(3):177–187.
- Zhang Y, Liu H, Li L, et al. Cholecystectomy can increase the risk of colorectal cancer: a meta-analysis of 10 cohort studies. *PLoS One.* 2017;12(8):e0181852.
- Bingham SA, Day NE, Luben R, et al., European Prospective Investigation into Cancer and Nutrition. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an observational study. *Lancet.* 2003;361(9368):1496–1501.
- Hullings AG, Sinha R, Liao LM, et al. Whole grain and dietary fiber intake and risk of colorectal cancer in the NIH-AARP Diet and Health Study cohort. *Am J Clin Nutr.* 2020;112(3):603–612.
- Aune D, Chan DS, Lau R, et al. Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *BMJ.* 2011;343:d6617.
- Oh H, Kim H, Lee DH, et al. Different dietary fibre sources and risks of colorectal cancer and adenoma: a dose-response meta-analysis of prospective studies. *Br J Nutr.* 2019;122(6):605–615.
- Scheppach W, Bartram HP, Richter F. Role of short-chain fatty acids in the prevention of colorectal cancer. *Eur J Cancer.* 1995;31A(7–8):1077–1080.
- Chen HM, Yu YN, Wang JL, et al. Decreased dietary fiber intake and structural alteration of gut microbiota in patients with advanced colorectal adenoma. *Am J Clin Nutr.* 2013;97(5):1044–1052.
- Kelly JJ, Alberts SR, Sacco F, et al. Colorectal cancer in Alaska native people, 2005–2009. *Gastrointest Cancer Res.* 2012;5(5):149–154.
- Wang T, Cai G, Qiu Y, et al. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J.* 2012;6(2):320–329.
- Farhat Z, Sampson JN, Hildesheim A, et al. Reproducibility, temporal variability, and concordance of serum and fecal bile acids and short chain fatty acids in a population-based study. *Cancer Epidemiol Biomarkers Prev.* 2021;30(10):1875–1883.
- Weissfeld JL, Schoen RE, Pinsky PF, et al. Flexible sigmoidoscopy in the PLCO cancer screening trial: results from the baseline screening examination of a randomized trial. *J Natl Cancer Inst.* 2005;97(13):989–997.