Metabolites of tobacco smoking and colorectal cancer risk

Amanda J.Cross*, Simina Boca[1], Neal D.Freedman,
Neil E.Caporaso[2], Wen-Yi Huang[3], Rashmi Sinha,
Joshua N.Sampson[1] and Steven C.Moore

Nutritional Epidemiology Branch, [1]Biostatistics Branch, [2]Genetic
Epidemiology Branch and [3]Occupational and Environmental Epidemiology
Branch, Division of Cancer Epidemiology and Genetics, National Cancer
Institute, National Institutes of Health, Department of Health and Human
Services, Rockville, MD 20850, USA

*To whom correspondence should be addressed. Department of Health and
Human Services, Nutritional Epidemiology Branch, Division of Cancer
Epidemiology and Genetics, National Cancer Institute, National Institutes of
Health, 9609 Medical Center Drive, Rockville, MD 20850, USA.
Tel: +1 301 496 4378; Fax: +1 301 496 6829;
Email: crossa@mail.nih.gov

Colorectal cancer is not strictly considered a tobacco-related
malignancy, but modest associations have emerged from large
meta-analyses. Most studies, however, use self-reported data,
which are subject to misclassification. Biomarkers of tobacco
exposure may reduce misclassification and provide insight into
metabolic variability that potentially influences carcinogen-
esis. Our aim was to identify metabolites that represent smok-
ing habits and individual variation in tobacco metabolism, and
investigate their association with colorectal cancer. In a nested
case-control study of 255 colorectal cancers and 254 matched
controls identified in the Prostate, Lung, Colorectal and Ovarian
cancer screening trial, baseline serum was used to identify metab-
olites by ultra-high-performance liquid-phase chromatography
and mass spectrometry, as well as gas chromatography with tan-
dem mass spectrometry. Odds ratios (OR) and 95% confidence
intervals (CI) were estimated by logistic regression. Self-reported
current smoking was associated with serum cotinine, O-cresol
sulfate and hydroxycotinine. Self-reported current smoking of
any tobacco (OR = 1.90, 95% CI: 1.02–3.54) and current cigarette
smoking (OR = 1.51, 95% CI: 0.75–3.04) were associated with
elevated colorectal cancer risks, although the latter was not statis-
tically significant. Individuals with detectable levels of hydroxy-
cotinine had an increased colorectal cancer risk compared with
those with undetectable levels (OR = 2.68, 95% CI: 1.33–5.40).
Although those with detectable levels of cotinine had a suggestive
elevated risk of this malignancy (OR = 1.81, 95% CI: 0.98–3.33),
those with detectable levels of O-cresol sulfate did not (OR = 1.16,
95% CI: 0.57–2.37). Biomarkers capturing smoking behavior and
metabolic variation exhibit stronger associations with colorectal
cancer than self-report, providing additional evidence for a role
for tobacco in this malignancy.

Introduction

It is well established that smoking is a risk factor for cancers of the
respiratory tract (1,2); however, the effect of smoking on colorec-
tal cancer has been less clear. In 2004, the International Agency for
Research on Cancer concluded that there was insufficient evidence
for a causal relationship between smoking and colorectal cancer
(2) and the United States Surgeon General's report concluded that
the association was only 'suggestive' (1). A few studies since then
have observed elevated risks with tobacco smoking (3–7) but most
do not reach statistical significance (8–14) or they are null (15–32).
However, in 2008, a meta-analysis of case-control and cohort stud-
ies reported a borderline statistically significant 7% increased risk of

colorectal cancer for current smokers [relative risk (RR) = 1.07, 95%
confidence interval (CI): 0.99–1.16] (33). Two further meta-analyses
published in 2009 provided additional support; one meta-analyses of
36 prospective studies reported a 17% increased risk for colorectal
cancer in current smokers (RR = 1.17, 95% CI: 0.97–1.40) (34), and
another meta-analyses of 28 cohorts reported a 20% increased risk
in current smokers (RR = 1.20, 95% CI: 1.10–1.30) (35). Based on
the available data in 2009, the International Agency for Research on
Cancer included colorectal cancer as a smoking-related malignancy
(36). Nevertheless, the magnitude of the association between smoking
and colorectal cancer in these observational studies is small, leading
to concerns about whether these associations were causal, or simply
reflected confounding or bias.

Tobacco smoking causes many physiologic changes and there are
several metabolic intermediates of tobacco. The epidemiologic lit-
erature, however, consists almost entirely of studies that examined
tobacco smoking by self-report. Self-report is invariably subject to
misclassification and is also not able to consider metabolic processes
that affect internal biologic exposure. Although self-report is effec-
tive at detecting strong associations between tobacco smoking and a
number of cancers, such as those of the lung, larynx, esophagus and
bladder, misclassification may have a particularly important effect on
cancers with weaker associations. Furthermore, using metabolites of
tobacco, rather than self-reported data, enable the incorporation of
both exposure and individual variability in carcinogen metabolism.
The aim of our study was first to identify biomarkers of tobacco using
non-targeted metabolomics and second to use the biomarkers iden-
tified to reevaluate the relationship between smoking and colorectal
cancer.

Methods

The Prostate, Lung, Colorectal and Ovarian cancer screening trial

We conducted a nested case-control study within the screening arm of the
Prostate, Lung, Colorectal and Ovarian cancer screening trial, which is a
large randomized controlled trial to test the efficacy of screening methods
for these four cancers (37–39). At baseline (1993–2001), ~155 000 men and
women from 10 United States centers, aged 55–74 years who had no history
of prostate, lung, colorectal or ovarian cancer, were enrolled and randomly
assigned to the screened or the non-screened arm. Those in the screened arm
(n = 77 445) were offered a flexible sigmoidoscopy at study entry to examine the
distal colorectum (60 cm), of which 83% (n = 64 658) were compliant and 89%
(n = 57 559) of these were considered successful (insertion to at least 50 cm
with >90% of mucosa visible or a suspect lesion identified). If neoplas-
tic lesions were detected during flexible sigmoidoscopy, participants were
referred for a colonoscopy. All participants in the screening arm required a
follow-up flexible sigmoidoscopy either 3 or 5 years after baseline.

Medical and pathologic reports during follow-up were obtained and
abstracted. The institutional review boards of the United States National
Cancer Institute and the 10 screening centers approved the study, and all par-
ticipants provided informed consent.

Study sample

Within the screening arm of the trial, our sample was drawn from those who
completed the baseline risk factor questionnaire and the dietary questionnaire,
provided consent for biospecimens to be used in etiologic studies and did not
have colorectal cancer at baseline (n = 52 705). We excluded individuals who
had a self-reported personal history of cancer (except basal cell skin cancer;
n = 4924), had <6 months of follow-up (an additional 168 individuals), had a
rare cancer during follow-up (an additional 1074), had self-reported Crohn's
disease, ulcerative colitis, familial polyposis, Gardner's syndrome or colorec-
tal polyps (an additional 6429) and those who did not have any serum avail-
able from baseline (an additional 2866 individuals); some individuals fell into
multiple exclusion categories.

After these inclusion/exclusion criteria, we selected the 255 first primary
incident colorectal cancers [International Classification of Diseases for
Oncology (third edition) codes: C180–189, C199, C209, C260] (40) whose
International Classification of Diseases for Oncology morphologies were not

in the range of 8240–8249 and were identified at least 6 months after baseline through February 2011. Controls ($n = 254$) were free from any cancer at the time the matched case was diagnosed and were incidence-density sampled and matched to the cases on age at randomization (5 year intervals), gender, race, year of randomization and season of blood draw.

Questionnaire data

At baseline, participants completed a risk factor questionnaire, which queried whether they had ever smoked cigarettes regularly for 6 months or longer and whether they were currently regularly smoking cigarettes. Furthermore, they were asked at what age they started smoking cigarettes, at what age they quit smoking (if they quit), how many cigarettes they usually smoked per day and whether they usually smoked filtered or non-filtered cigarettes. Finally, the questionnaire asked whether they had ever smoked a pipe or cigar regularly.

Metabolite assessment

Non-fasting serum samples from baseline that had not previously been thawed were arranged in batches so that matched cases and controls were consecutive samples within a batch, and the order of case versus control was counterbalanced within each batch. Blinded quality control samples of pooled serum were inserted at a level of 10% randomly throughout each batch, in addition to an unblinded standard every sixth sample.

The metabolomics methods used have been described previously in detail (41). In brief, a non-targeted single methanol extraction was performed, followed by protein precipitation, to allow identification of a range of metabolites approximately under 1000 Daltons. Ultra-high-performance liquid-phase chromatography and mass spectrometry, as well as gas chromatography coupled with tandem mass spectrometry, was used to identify peaks, as described previously (41). Using a chemical reference library generated from 2500 standards, mass spectral peaks, retention times and mass-to-charge ratios were used to identify individual metabolites as well as their relative quantities. A total of 446 metabolites were identified, including amino acids, carbohydrates, fatty acids, androgens and xenobiotics. Using the quality control samples to assess technical reliability, the median coefficients of variation across all of the metabolites was 0.10 (inter-quartile range: 0.04–0.21).

Statistical analysis

We describe demographic information in the case and control groups, with *P*-values calculated by either Fisher's exact test (categorical variables) or Wilcoxon rank test (continuous) for characteristics not used in matching individuals. Each metabolite value was batch normalized by dividing it by the batch median (of the non-missing values). Unless otherwise stated, values were log transformed, with values less than the detection threshold being set to the minimum observed value.

We first identified metabolites associated with smoking. We modeled the effect of current smoking status on metabolite level by linear regression adjusted for age, gender, race, study center and body mass index (BMI). To account for multiple comparisons, we used a Bonferroni corrected *P*-value of 0.05/446 as the threshold of statistical significance. We then considered whether there was a dose-response relationship for the study participants who were current cigarette smokers (we did not have dose information for cigars or pipes) by modeling the effect of self-reported maximum number of cigarettes smoked per day on metabolite levels by linear regression.

We then compared the relationships between both self-reported smoking and smoking-related metabolites with colorectal cancer. For self-reported responses, we estimated odds ratios (OR) and 95% CIs for colorectal cancer, as compared with never smokers, for each grouping by conditional logistic regression, conditioned on the matching factors, and additionally adjusted for study center and BMI. When former smokers were omitted from the analysis, we estimated the OR by unconditional logistic regression adjusted for age, gender, race, study center and BMI. For smoking-related metabolites, we characterized individuals as exposed (i.e. detectable levels) or unexposed (i.e. below detection threshold) and estimated the OR between the two groups by conditional logistic regression adjusted for study center and BMI. To assess the dose response, the exposed group was further divided by whether their unnormalized level was above the median. For all analyses, the *P*-trend was calculated by assigning ordered numeric values 0, 1, … , K, to the K categories (e.g. K = 3 when categorized as unexposed, below median and above median) and then performing a likelihood ratio test comparing models with and without that ordinal variable.

Results

The median follow-up time from serum collection to diagnosis of colorectal cancer was 7.8 years (25th and 75th percentiles were 5.6 and 10.1 years). None of the baseline characteristics were significantly different between cases and controls (Table I). There were a total of 55 individuals who were current smokers of any tobacco (cigarettes, cigars or pipes).

After correcting for multiple testing, we identified eight metabolites that were significantly associated with self-reported current smoking of any tobacco among controls (Table II). The metabolites most strongly associated with self-reported current smoking were cotinine, O-cresol sulfate and hydroxycotinine; these associations were evident among both the cases and the controls. Other metabolites associated with self-reported current smoking included trigonelline (N´-methyl nicotinate) and N-(2-furoyl)glycine in both cases and controls, as well as 1-methylxanthine, 1-methylurate and paraxanthine in the controls only. To reduce the possibility of chance findings, we only carried forward cotinine, O-cresol sulfate and hydroxycotinine to the main analysis because these were the most significantly associated with self-reported current smoking. There was modest correlation between cotinine and both O-cresol sulfate ($r = 0.63$) and hydroxycotinine ($r = 0.76$), and the lowest correlation was between hydroxycotinine and O-cresol sulfate ($r = 0.49$).

Defining current smokers using the self-reported questionnaire data, 50 of the 55 current smokers had detectable levels of cotinine in their serum, 35 had O-cresol sulfate and 37 had hydroxycotinine; therefore, cotinine appeared to be the most sensitive marker of current smoking. In current cigarette smokers, we further investigated whether serum cotinine, O-cresol sulfate and hydroxycotinine increased with self-reported cigarettes smoked per day, and we observed evidence for a dose response with cotinine (*P*-trend = 0.010; Figure 1A), O-cresol sulfate (*P*-trend = 0.058; Figure 1B) and hydroxycotinine (*P*-trend = 0.010; Figure 1C). There was only one individual at the highest end of exposure who smoked 60 cigarettes per day; if this data point was excluded, the dose response was still evident for cotinine and hydroxycotinine (*P*-trend = 0.012, and *P*-trend = 0.037, respectively), but not for O-cresol sulfate (*P*-trend = 0.187).

Using only the questionnaire-based data, current users of any tobacco had an elevated risk of colorectal cancer compared with non-current users (OR = 1.90, 95% CI: 1.02–3.54; *P*-trend = 0.042; Table III). Examining cigarette smoking specifically, current cigarette smokers had an elevated, but not statistically significant, risk of colorectal cancer, compared with those who had never smoked cigarettes (OR = 1.51, 95% CI: 0.75–3.04). Within current cigarette smokers, there was no association by cigarettes per day, years smoked, pack years smoked or preference of filtered versus non-filtered cigarettes and colorectal cancer; furthermore, within former smokers, there was no effect of time since quitting (Table III).

Using the metabolite data, there were 59 individuals who had detectable levels of cotinine in their serum, and compared with those with undetectable levels of cotinine, they had an elevated risk of colorectal cancer, but the estimate did not quite reach statistical significance (OR = 1.81, 95% CI: 0.98–3.33) and was of a similar magnitude to that obtained from self-reported current smoking of any tobacco (Table IV). Defining cotinine exposure as low or high, compared with those with undetectable levels of cotinine in their serum, did not change the risk estimates obtained (OR = 1.84, 95% CI: 0.80–4.23 and OR = 1.78, 95% CI: 0.78–4.07, respectively). There were 40 individuals who had detectable levels of O-cresol sulfate in their serum, but this metabolite was not associated with colorectal cancer (OR for exposed versus unexposed = 1.16, 95% CI: 0.57–2.37; OR for highest category of exposure versus no exposure = 1.99, 0.72–5.49; Table IV). Hydroxycotinine was present in the serum of 47 individuals at baseline, and this conferred an increased risk for colorectal cancer during follow-up (OR for exposed versus unexposed = 2.68, 95% CI: 1.33–5.40); this risk was further elevated among those in the highest exposure category compared with those unexposed (OR = 3.19, 95% CI: 1.13–9.04; *P*-trend = 0.006; Table IV).

Five cases, but zero controls, tested positive for hydroxycotinine but not cotinine; whereas 17 individuals tested positive for cotinine (nine cases and eight controls) but not hydroxycotinine. In exploratory analyses, we investigated the ratio of serum cotinine to hydroxycotinine in relation to colorectal cancer; however, we did not

Table I. Participant characteristics (mean ± SD)[a]

		Cases (*n* = 255)	Controls (*n* = 254)	*P*-value[b]
Gender	Men	143	143	
	Women	112	111	Matched
Age (years)		64.3 (5.1)	64.3 (5.1)	Matched
Race	White	227	226	
	Black	13	13	
	Other	15	15	Matched
Cigarette smoking status	Current	26	19	
	Former	120	121	
	Never	109	114	0.57
Cigar smoking status	Current	5	2	
	Former	28	35	
	Never	220	217	0.36
Pipe smoking status	Current	4	2	
	Former	40	41	
	Never	208	211	0.79
Cigarettes per day (current smokers)	10	6	3	
	20	11	8	
	≥30	8	8	0.79
Pack years (current smokers)	0.25 to <41	10	4	
	41–60	7	8	
	>60	8	7	0.41
Years since quit cigarettes (former smokers)	>20	62	72	
	>10 to 20	33	24	
	≤10	22	22	0.50
Education	High school or less	85	84	
	Posthigh school training/some college	85	75	
	College or postgraduate	85	94	0.59
BMI (kg/m^2)	<25	78	88	
	25 to <30	103	112	
	30 to <35	54	41	
	35+	19	10	0.14
Vigorous physical activity (h/week)	<1	75	78	
	1–3	112	95	
	4+	65	75	0.35
Menopausal hormone use[c]	Never	38	28	
	Former	19	21	
	Current	53	62	0.31
Regular aspirin use	No	124	131	
	Yes	131	123	0.54
Regular ibuprofen use	No	174	183	
	Yes	80	71	0.44
Alcohol (g/day)		12.2 (23.0)	13.2 (23.6)	0.44

[a]May not add to total due to some missing values.
[b]Fisher's exact test (categorical variables) or Wilcoxon rank test (continuous data).
[c]Among women only.

observe any significant associations. Because cotinine and hydroxycotinine are both metabolites of nicotine, we examined a combined exposure by summing normalized (mean = 0, standard deviation = 1) versions of these two variables. The results we obtained from this combined cotinine/hydroxycotinine variable revealed similar correlations with self-reported tobacco smoking within cases and controls combined (*r* = 0.78, *P* = 5.68e-112), within controls only (*r* = 0.73, *P*-value = 1.07e-47), and within cases only (*r* = 0.82, *P* = 3.86e-58). The combined cotinine/hydroxycotinine variable was associated with colorectal cancer when comparing those exposed to unexposed (multivariable OR = 2.10, 95% CI: 1.17–3.75; *P*-trend = 0.012), and when comparing those in the highest category of exposure (above the median among those who were exposed) compared with the unexposed individuals (multivariable OR = 2.66, 95% CI: 1.15–6.14; *P*-trend = 0.009).

Discussion

The association between self-reported smoking and colorectal cancer has been inconsistent in previous literature and the magnitude of the risks observed has been small. In this study, we used metabolomics to identify tobacco-related metabolites in serum that represent an internal biologic measure of tobacco exposure that incorporates both exposure and metabolism. Using these circulating biological exposure measures, we observed that one of these biomarkers, serum hydroxycotinine, had a stronger association with colorectal cancer than did self-reported tobacco use alone.

The International Agency for Research on Cancer determined that colorectal cancer is a smoking-associated malignancy and that the colorectum may be exposed to numerous tobacco-related carcinogens via the circulatory system (42). Yet, associations between smoking and colorectal cancer have been modest and inconsistent across previous epidemiologic studies. Three meta-analyses have determined that there is a positive association between smoking and colorectal cancer, but the risks were only bordering on statistical significance (RR for current versus never smokers = 1.07, 95% CI: 0.99–1.16 (33); 1.17, 95% CI: 0.97–1.40 (34); 1.20, 95% CI: 1.10–1.30 (35), respectively). Studies performed subsequent to these meta-analyses have also been inconsistent showing both positive and null associations (6,7,14).

Nearly all epidemiologic studies of tobacco smoking have relied on self-reported assessment, which is subject to reporting bias and measurement error. Self-reported data have effectively been used to detect associations between tobacco smoking and a number of other malignancies; however, measurement error may be more problematic for

Table II. Metabolites associated with self-reported current smoking of cigarettes, pipes or cigars[a]

Cases and controls			Controls			Cases		
Metabolite	r	P-value	Metabolite	r	P-value	Metabolite	r	P-value
Cotinine	0.88	2.64e-158	Cotinine	0.84	1.1e-60	Cotinine	0.91	3.45e-94
O-Cresol sulfate	0.72	3.98e-78	O-Cresol sulfate	0.69	2.63e-35	O-Cresol sulfate	0.74	1.22e-42
Hydroxycotinine	0.67	2.56e-60	Hydroxycotinine	0.66	1.33e-28	Hydroxycotinine	0.68	5.41e-32
4-Vinylphenol sulfate	0.29	6.26e-10	1-Methylxanthine	0.34	9.23e-07	4-Vinylphenol sulfate	0.35	5.52e-09
N-(2-furoyl)glycine	0.25	2.19e-09	1-Methylurate	0.32	1.42e-06	N-(2-furoyl)glycine	0.23	8.67e-06
1-Methylurate	0.27	3.86e-09	Trigonelline (N´-methyl nicotinate)	0.28	2.84e-05	Catechol sulfate	0.23	5.72e-05
1-Methylxanthine	0.29	8.87e-09	N-(2-furoyl)glycine	0.27	9.37e-05	Trigonelline (N´-methyl nicotinate)	0.23	7.96e-05
Trigonelline (N´-methyl nicotinate)	0.25	3.83e-08	Paraxanthine	0.23	9.81e-05	Threonate	-0.29	0.000125
Threonate	-0.25	5.05e-08	1,3-Dimethylurate	0.27	0.000147	21-Hydroxypregnenolone disulfate	0.30	0.000167
Catechol sulfate	0.25	2.03e-07	Threonate	-0.20	0.000163	Pregnenolone sulfate	0.30	2e-04
Quinate	0.22	4.45e-06	Catechol sulfate	0.27	0.000624	1-Methylurate	0.24	0.000369
4-Ethylphenylsulfate	0.21	4.61e-06	1,3,7-Trimethylurate	0.23	0.000719	Quinate	0.19	0.000388
Scyllo-inositol	-0.19	2.36e-05	Quinate	0.26	0.000813	1-Methylxanthine	0.25	0.000493
Paraxanthine	0.19	6.2e-05	4-Ethylphenylsulfate	0.22	0.00101	Andro steroid monosulfate 2	0.23	0.000967
3-Hydroxyhippurate	0.17	7.07e-05	Theophylline	0.20	0.00114	4-Ethylphenylsulfate	0.20	0.0012
21-Hydroxypregnenolone disulfate	0.22	0.00014	1-Palmitoylglycerol	-0.23	0.00292	Carnitine	0.16	0.00184
Piperine	-0.14	0.000174	Hippurate	0.21	0.00329	Scyllo-inositol	-0.25	0.00197
1,3,7-Trimethylurate	0.17	0.000371	3-Hydroxyhippurate	0.21	0.00339	3-Carboxy-4-methyl-5-propyl-2-furanpropanoate	-0.21	0.00256
Pregnenolone sulfate	0.22	0.000645	Caffeine	0.17	0.00376	2-Hydroxystearate	0.15	0.00272
1,3-Dimethylurate	0.18	0.000824	4-Vinylphenol sulfate	0.22	0.00585	Dehydroisoandrosterone sulfate	0.25	0.00276

Shaded cells reach the Bonferroni corrected P-value of 0.05/446.
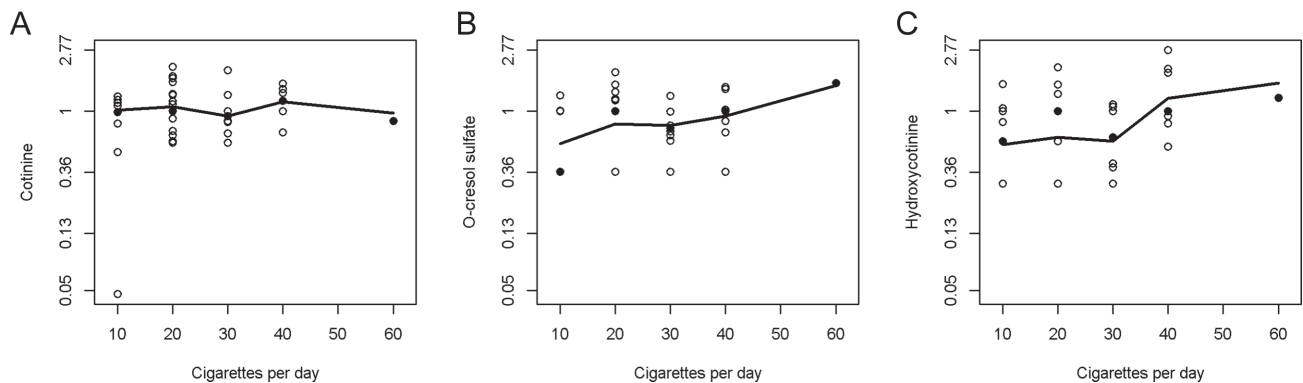[a]Adjusted for age, gender, race, study center and BMI.

Fig. 1. (**A**) Serum cotinine and self-reported cigarettes smoked per day among current cigarette smokers, with a lowess smoother, P-trend = 0.010; if the outlier of 60 cigarettes per day is removed, the P-trend = 0.012. (**B**) Serum O-cresol sulfate and self-reported cigarettes smoked per day among current cigarette smokers, with a lowess smoother, P-trend = 0.058; if the outlier of 60 cigarettes per day is removed, the P-trend = 0.187. (**C**) Serum hydroxycotinine and self-reported cigarettes smoked per day among current cigarette smokers, with a lowess smoother, P-trend = 0.010; if the outlier of 60 cigarettes per day is removed, the P-trend = 0.037.

cancers with modest associations. Biomarkers can potentially serve as objective markers of tobacco exposure that also incorporate individual variability in carcinogen metabolism to facilitate both our understanding of cancer etiology and tobacco carcinogenesis.

We found a number of small metabolites associated with tobacco smoking. Serum cotinine, O-cresol sulfate and hydroxycotinine were all associated with self-reported smoking in both cases and controls; moreover, they showed evidence for a dose response according to cigarettes smoked per day among current smokers. Our findings revealed that serum hydroxycotinine was a stronger predictor of colorectal cancer risk than cotinine and O-cresol sulfate. Different associations between particular metabolites and cancer could potentially reflect the pharmacokinetics of each.

Approximately 70–80% of nicotine is metabolized to cotinine by the liver; this is further metabolized into several other metabolites, with

the majority (33–40%) becoming hydroxycotinine (43). Although we conducted exploratory analyses of the ratio of cotinine to hydroxycotinine, our study had insufficient power to investigate this combined exposure. Although the half-life of nicotine in the blood is only ~2 h, its metabolites, such as cotinine and hydroxycotinine, have longer half-lives of around 5–20 h (43). The clearance of nicotine metabolites may be pertinent to cancer risk.

A previous study showed that among ~1500 people who reported to have smoked 'today or yesterday', 5.2% did not have measurable levels of cotinine in their serum; furthermore, ~60% did not have cotinine in their serum if they had not smoked 'today or yesterday' but had smoked within the last month (44). This provides evidence that not all smokers have nicotine metabolites in their blood at any given time, although this previous study did not investigate hydroxycotinine. In addition, it is clear that variants of the *CYP2A6* gene result

in varying rates of nicotine metabolism. A recent study showed that genetic variation within the nicotinic acetylcholine receptor gene cluster had a stronger influence on circulating cotinine levels than self-reported cigarettes per day (45); this study also did not investigate hydroxycotinine.

In our study, neither serum cotinine nor O-cresol sulfate was statistically significantly associated with colorectal cancer. Cotinine present in blood or urine has been used as a biomarker for tobacco smoking for several years (46). Previous prospective studies of tobacco-related malignancies have investigated plasma or urinary cotinine and reported an elevated risk of pancreatic (47) and lung cancer (48),

Table III. Self-reported smoking variables and colorectal cancer risk

	OR (95% CI)[a]
Any tobacco use (cigarette, cigar and pipe)	
Never/former any tobacco	1.0
Current any tobacco	1.90 (1.02–3.54)
P-trend	0.042
Cigarette smoking status	
Never	1.0
Former	1.09 (0.72–1.65)
Current	1.51 (0.75–3.04)
P-trend	0.294
Cigarettes per day (current smokers)	
Never	1.0
10	1.66 (0.34–8.18)
20	1.57 (0.57–4.34)
≥30	1.11 (0.34–3.62)
P-trend	0.488
Years smoked (current smokers)	
Never	1.0
0.5–40	3.04 (0.84–11.05)
>40	1.09 (0.47–2.55)
P-trend	0.543
Pack years (current smokers)	
Never	1.0
0.25–40	2.63 (0.71–9.78)
41–60	0.95 (0.31–2.87)
>60	1.30 (0.38–4.42)
P-trend	0.599
Filtered cigarettes	
Never	1.0
Filtered	1.36 (0.90–2.06)
Unfiltered	0.74 (0.44–1.25)
P-trend	0.592
Years since quit cigarettes (former smokers)	
Never	1.0
>20	0.94 (0.59–1.48)
>10 to 20	1.54 (0.83–2.86)
≤10	1.05 (0.53–2.07)
P-trend	0.210

[a]Conditional logistic regression with models adjusted for study center and BMI.

respectively. Although there have been other studies of cotinine, very little is known about the association between O-cresol sulfate and cancer risk. O-cresol sulfate is a known metabolite of toluene, which has been identified in tobacco smoke (49) but has not been extensively studied.

Our study revealed that individuals with hydroxycotinine in their serum had an increased risk of colorectal cancer. To our knowledge, no previous epidemiologic study has investigated serum hydroxycotinine in relation to colorectal cancer. Previous studies have shown that urinary excretion of hydroxycotinine is induced by tobacco smoking (50). Hydroxycotinine is a metabolite of cotinine and the formation of hydroxycotinine is dependent on the activity of the polymorphic hepatic enzyme CYP2A6 (43).

Serum hydroxycotinine may identify those who are less efficient at eliminating nicotine metabolites and, therefore, those most at risk. Using these serologic biomarkers could improve exposure assessment to enable the investigation of other cancers where associations with smoking have been unclear. Serologic measures of tobacco exposure may also be a superior means of controlling for confounding by tobacco because they incorporate metabolism and all sources of exposure to tobacco. Because hydroxycotinine was present in individuals who had undetectable levels of cotinine and who were not currently smoking according to the self-report data (n = 3 former smokers and n = 2 never smokers), there may be other exposures to tobacco that could explain the increased risk of disease observed with this metabolite, or it may be that this metabolite is capturing other exposures.

Strengths of our study include the prospective design whereby self-reported data, as well as blood samples, were obtained prior to the onset of disease; this is important because the disease state can result in changes in behavior, as well as perturbations in metabolite levels. Nevertheless, there are some limitations of our study. The association between smoking and colorectal cancer may be stronger among certain subtypes of this malignancy, such as those characterized by microsatellite instability (16,51); however, we were unable to investigate this due to a lack of information on this subtype in our study. In addition, our metabolomics analyses were limited to metabolites <1000 Daltons, metabolites that were detectable by mass-spectrometry-based platforms, and the data obtained from the platforms were in relative, rather than absolute concentrations. Our agnostic analysis did not identify markers of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. Previous studies of lung cancer have conducted a targeted analysis of the 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone metabolite: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol and its glucuronides and reported an increased risk of lung cancer with a unit standard deviation increase in serum total 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (52); however, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol require metabolic activation to exert their carcinogenic effects and therefore in any individual their effects can vary. Finally, we were limited by the

Table IV. Serum metabolites associated with self-reported smoking in relation to colorectal cancer

Metabolite	Exposed cases	Exposed controls	Unexposed cases	Unexposed controls	Exposed versus unexposed OR (95% CI)[b]	*P*-value	Low[a] metabolite level versus unexposed OR (95% CI)[b]	High[a] metabolite level versus unexposed OR (95% CI)[b]	*P*-trend
Self-reported current smoker[c]	33	22	219	232	1.90 (1.02–3.54)	0.042	—	—	—
Cotinine	36	23	219	231	1.81 (0.98–3.33)	0.058	1.84 (0.80–4.23)	1.78 (0.78–4.07)	0.078
O-Cresol sulfate	21	19	234	235	1.16 (0.57–2.37)	0.685	0.68 (0.25–1.85)	1.99 (0.72–5.49)	0.384
Hydroxycotinine	32	15	223	239	2.68 (1.33–5.40)	0.006	2.32 (0.91–5.88)	3.19 (1.13–9.04)	0.006

[a]Categorical variable using the median level as a cutpoint among those with values >0.
[b]Conditional logistic regression with models adjusted for study center and BMI.
[c]Current smoker of any tobacco (cigarettes, pipes and cigars) versus everyone else is also given in Table III, but listed again here as comparison.

measurement of metabolites at a single timepoint and, therefore, we are unable to quantify past or lifetime exposure; and due to small numbers, it is possible that our observations could be due to chance.

In conclusion, we identified serum cotinine, O-cresol sulfate and hydroxycotinine as biomarkers of self-reported current smoking. Hydroxycotinine was significantly associated with colorectal cancer, and the risk observed was stronger than that for self-reported tobacco use and this malignancy; thus, our findings provide additional evidence for an association between tobacco smoking and colorectal cancer.

References

1. Office of the Surgeon General (2004) The health consequences of smoking: a report of the surgeon general. Centers for Disease Control and Prevention, Atlanta, GA.
2. IARC (2004) Tobacco smoke and involuntary smoking. *IARC Monographs on the Evaluation of Carcinogenic Risks to Human*, Vol. **83**. International Agency for Research on Cancer, Lyon, pp. 161–263.
3. Chyou,P.H. *et al.* (1996) A prospective study of colon and rectal cancer among Hawaii Japanese men. *Ann. Epidemiol.*, **6**, 276–282.
4. Driver,J.A. *et al.* (2007) Development of a risk score for colorectal cancer in men. *Am. J. Med.*, **120**, 257–263.
5. Stürmer,T. *et al.* (2000) Lifetime cigarette smoking and colorectal cancer incidence in the Physicians' Health Study I. *J. Natl Cancer Inst.*, **92**, 1178–1181.
6. Hannan,L.M. *et al.* (2009) The association between cigarette smoking and risk of colorectal cancer in a large prospective cohort from the United States. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 3362–3367.
7. Limsui,D. *et al.* (2010) Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. *J. Natl Cancer Inst.*, **102**, 1012–1022.
8. Giovannucci,E. *et al.* (1994) A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in U.S. men. *J. Natl Cancer Inst.*, **86**, 183–191.
9. Tulinius,H. *et al.* (1997) Risk factors for malignant diseases: a cohort study on a population of 22,946 Icelanders. *Cancer Epidemiol. Biomarkers Prev.*, **6**, 863–873.
10. Wu,A.H. *et al.* (1987) Alcohol, physical activity and other risk factors for colorectal cancer: a prospective study. *Br. J. Cancer*, **55**, 687–694.
11. Sanjoaquin,M.A. *et al.* (2004) Nutrition, lifestyle and colorectal cancer incidence: a prospective investigation of 10998 vegetarians and non-vegetarians in the United Kingdom. *Br. J. Cancer*, **90**, 118–121.
12. Akhter,M. *et al.* (2007) Cigarette smoking and the risk of colorectal cancer among men: a prospective study in Japan. *Eur. J. Cancer Prev.*, **16**, 102–107.
13. Berndt,S.I. *et al.* (2006) Genetic variation in the nucleotide excision repair pathway and colorectal cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **15**, 2263–2269.
14. Leufkens,A.M. *et al.* (2011) Cigarette smoking and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition study. *Clin. Gastroenterol. Hepatol.*, **9**, 137–144.
15. Nordlund,L.A. *et al.* (1997) Cancer incidence in female smokers: a 26-year follow-up. *Int. J. Cancer*, **73**, 625–628.
16. Poynter,J.N. *et al.* (2009) Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 2745–2750.
17. Sandler,R.S. *et al.* (1988) Cigarette smoking and the risk of colorectal cancer in women. *J. Natl Cancer Inst.*, **80**, 1329–1333.
18. Giovannucci,E. *et al.* (1994) A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in U.S. women. *J. Natl Cancer Inst.*, **86**, 192–199.
19. Engeland,A. *et al.* (1996) Smoking habits and risk of cancers other than lung cancer: 28 years' follow-up of 26,000 Norwegian men and women. *Cancer Causes Control*, **7**, 497–506.
20. Nyrén,O. *et al.* (1996) Smoking and colorectal cancer: a 20-year follow-up study of Swedish construction workers. *J. Natl Cancer Inst.*, **88**, 1302–1307.
21. Kato,I. *et al.* (1997) Prospective study of diet and female colorectal cancer: the New York University Women's Health Study. *Nutr. Cancer*, **28**, 276–281.
22. Knekt,P. *et al.* (1998) Smoking and risk of colorectal cancer. *Br. J. Cancer*, **78**, 136–139.
23. Terry,P.D. *et al.* (2002) Prospective cohort study of cigarette smoking and colorectal cancer risk in women. *Int. J. Cancer*, **99**, 480–483.
24. Limburg,P.J. *et al.* (2003) Cigarette smoking and colorectal cancer: long-term, subsite-specific risks in a cohort study of postmenopausal women. *Clin. Gastroenterol. Hepatol.*, **1**, 202–210.
25. Wakai,K. *et al.* (2003) Smoking and colorectal cancer in a non-Western population: a prospective cohort study in Japan. *J. Epidemiol.*, **13**, 323–332.
26. Jee,S.H. *et al.* (2004) Smoking and cancer risk in Korean men and women. *Cancer Causes Control*, **15**, 341–348.
27. Lüchtenborg,M. *et al.* (2005) Cigarette smoking and colorectal cancer: APC mutations, hMLH1 expression, and GSTM1 and GSTT1 polymorphisms. *Am. J. Epidemiol.*, **161**, 806–815.
28. Yun,Y.H. *et al.* (2005) Cigarette smoking and cancer incidence risk in adult men: National Health Insurance Corporation Study. *Cancer Detect. Prev.*, **29**, 15–24.
29. Kim,H.J. *et al.* (2006) Smoking and colorectal cancer risk in the Korean elderly. *J. Prev. Med. Public Health*, **39**, 123–129.
30. Shankar,A. *et al.* (2008) Morbidity and mortality in relation to smoking among women and men of Chinese ethnicity: the Singapore Chinese Health Study. *Eur. J. Cancer*, **44**, 100–109.
31. Tiemersma,E.W. *et al.* (2002) Meat consumption, cigarette smoking, and genetic susceptibility in the etiology of colorectal cancer: results from a Dutch prospective study. *Cancer Causes Control*, **13**, 383–393.
32. Weijenberg,M.P. *et al.* (2008) Cigarette smoking and KRAS oncogene mutations in sporadic colorectal cancer: results from the Netherlands Cohort Study. *Mutat. Res.*, **652**, 54–64.
33. Botteri,E. *et al.* (2008) Smoking and colorectal cancer: a meta-analysis. *JAMA*, **300**, 2765–2778.
34. Liang,P.S. *et al.* (2009) Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *Int. J. Cancer*, **124**, 2406–2415.
35. Tsoi,K.K. *et al.* (2009) Cigarette smoking and the risk of colorectal cancer: a meta-analysis of prospective cohort studies. *Clin. Gastroenterol. Hepatol.*, **7**, 682–688.e1.
36. Secretan,B. *et al.* (2009) A review of human carcinogens–Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol.*, **10**, 1033–1034.
37. Gohagan,J.K. *et al.* (2000) The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. *Control. Clin. Trials*, **21** (suppl. 6), 251S–272S.
38. Hayes,R.B. *et al.* (2005) Methods for etiologic and early marker investigations in the PLCO trial. *Mutat. Res.*, **592**, 147–154.
39. Prorok,P.C. *et al.* (2000) Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control. Clin. Trials*, **21**(suppl. 6), 273S–309S.
40. Fritz,A. *et al.* (2000) *International Classification of Diseases for Oncology (ICD-O)*. 3rd edn. World Health Organization, Geneva, Switzerland.
41. Evans,A.M. *et al.* (2009) Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem.*, **81**, 6656–6667.
42. Yamasaki,E. *et al.* (1977) Concentration of mutagens from urine by absorption with the nonpolar resin XAD-2: cigarette smokers have mutagenic urine. *Proc. Natl Acad. Sci. USA*, **74**, 3555–3559.
43. Hukkanen,J. *et al.* (2005) Metabolism and disposition kinetics of nicotine. *Pharmacol. Rev.*, **57**, 79–115.
44. Vartiainen,E. *et al.* (2002) Validation of self reported smoking by serum cotinine measurement in a community-based study. *J. Epidemiol. Community Health*, **56**, 167–170.
45. Munafò,M.R. *et al.* (2012) Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J. Natl Cancer Inst.*, **104**, 740–748.
46. Connor Gorber,S. *et al.* (2009) The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine Tob. Res.*, **11**, 12–24.
47. Leenders,M. *et al.* (2012) Plasma cotinine levels and pancreatic cancer in the EPIC cohort study. *Int. J. Cancer*, **131**, 997–1002.

48. Hecht,S.S. *et al.* (2013) Tobacco smoke biomarkers and cancer risk among male smokers in the Shanghai cohort study. *Cancer Letters,* **334**, 34–38.

49. Pieraccini,G. *et al.* (2008) Identification and determination of mainstream and sidestream smoke components in different brands and types of cigarettes by means of solid-phase microextraction-gas chromatography-mass spectrometry. *J. Chromatogr. A*, **1180**, 138–150.

50. Benowitz,N.L. *et al.* (2000) Effects of cigarette smoking and carbon monoxide on nicotine and cotinine metabolism. *Clin. Pharmacol. Ther.*, **67**, 653–659.

51. Slattery,M.L. *et al.* (2000) Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. *J. Natl Cancer Inst.*, **92**, 1831–1836.

52. Church,T.R. *et al.* (2009) A prospectively measured serum biomarker for a tobacco-specific carcinogen and lung cancer in smokers. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 260–266.