Title: Case-Control Study of Cumulative Cigarette Tar Exposure and Lung and Upper Aerodigestive Tract Cancers

Authors and affiliations: Travis J. Meyers[1], Shen-Chih Chang[2], Po-Yin Chang[3], Hal Morgenstern[4], Donald P. Tashkin[5], Jian-Yu Rao[6], Wendy Cozen[7], Thomas M. Mack[8], and Zuo-Feng Zhang[9]*

Author Affiliations:

[1]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California.

[2]Division of Neonatal and Developmental Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, California; Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California.

[3]Division of Epidemiology, Department of Public Health Sciences, University of California, Davis School of Medicine, Davis, California.

[4]Departments of Epidemiology and Environmental Health Sciences, School of Public Health and Comprehensive Cancer Center, University of Michigan, Ann Arbor, Michigan.

[5]Division of Pulmonary and Critical Care Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California.

[6]Department of Pathology, David Geffen School of Medicine at UCLA, Los Angeles, California; Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California.

[7]Departments of Preventive Medicine and Pathology, Keck School of Medicine of the University of Southern California, Los Angeles, California.

[8]Departments of Preventive Medicine and Pathology, Keck School of Medicine of the University of Southern California, Los Angeles, California.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version record. Please cite this article as doi:10.1002/ijc.30632.

[9]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California; Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, California.

***Corresponding author:** Zuo-Feng Zhang, Department of Epidemiology, UCLA Fielding School of Public Health, 71-225 CHS, Box 951772, 650 Charles E. Young Drive South, Los Angeles, CA 90095-1772. Phone: 310-825-8418; Fax: 310-206-6039; E-mail: zfzhang@ucla.edu

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

Keywords: lung cancer; UADT cancer; epidemiology; tobacco tar; case-control studies

Abbreviations: UADT: Upper aerodigestive tract; FTC: US Federal Trade Commission; IQR: Interquartile range; OR: Odds ratio; CL: Confidence limit; SBOR: Semi-Bayes odds ratio; PL: Semi-Bayes posterior limit; ROC-AUC: Area under the curve of the receiver operating characteristic; CSI: comprehensive smoking index

Article category: Cancer epidemiology

Manuscript Word Count: 4,018

Abstract Word Count: 249

Tables: 9

Figures: 2 equations

Novelty and Impact Statement:

Tobacco use is the leading preventable risk factor for cancer mortality worldwide. Standard exposure estimates fail to account for different emissions between products. In this analysis, we estimated cumulative cigarette tar exposure from 39 government reports for participants of a case-control study. Cumulative tar was associated with lung cancer-especially small and large cell subtypes-even after adjusting for pack-years. Incorporating the composition of tobacco carcinogens in lifetime smoking exposure may improve lung cancer risk estimation.

Abstract:

The development of comprehensive measures for tobacco exposure is crucial to specify effects on disease and inform public health policy. In this population-based case-control study, we evaluated the associations between cumulative lifetime cigarette tar exposure and cancers of the lung and upper aerodigestive tract (UADT). The study included 611 incident cases of lung cancer; 601 cases of UADT cancers (oropharyngeal, laryngeal, and esophageal cancers); and 1,040 cancer-free controls. We estimated lifetime exposure to cigarette tar based on tar concentrations abstracted from government cigarette records and self-reported smoking histories derived from a standardized questionnaire. We analyzed the associations for cumulative tar exposure with lung and UADT cancer, overall and according to histological subtype. Cumulative tar exposure was highly correlated with pack-years among ever smoking controls (Pearson coefficient=0.90). The adjusted odds ratio (95% confidence limits) for the estimated effect of about 1 kilogram increase in tar exposure (approximately the interquartile range in all controls) was 1.61 (1.50, 1.73) for lung cancer and 1.21 (1.13, 1.29) for UADT cancers. In general, tar exposure was more highly associated with small, squamous, and large cell lung cancer than to adenocarcinoma. With additional adjustment for pack-years, positive associations between tar and lung cancer were evident, particularly for small cell and large cell subtypes. Therefore, incorporating the composition of tobacco carcinogens in lifetime smoking exposure may improve lung cancer risk estimation. This study does not support the claim of a

null or inverse association between ‘low exposure’ to tobacco smoke and risk of these cancer types.

Introduction:

Tobacco smoking has been identified as a causal factor for 15 organ sites, including the lung and upper aerodigestive tract (UADT) (1). In addition, smoking is associated with all major histological subtypes of lung cancer, although a higher association has been reported for small cell cancer and squamous cell carcinoma than for large cell lung cancer and adenocarcinoma (2, 3). With respect to UADT cancer, smoking is associated with squamous cell carcinoma of the head and neck, and with both squamous cell carcinoma and adenocarcinoma of the esophagus (1). Tobacco smoke is a complex mixture of over 7,000 compounds, of which 81 are considered carcinogenic or potentially carcinogenic in humans (4-8). ‘Tar’ is a common term for the total particulate matter in tobacco smoke- excluding nicotine and water- that contains these putative carcinogens such as benzo[a]pyrene (4). While standard measures of tobacco exposure (e.g. pack-years) treat tobacco smoke as homogeneous, emissions have been shown to vary not only between filtered and unfiltered cigarettes, but also between brands of each type (4, 9-11).

Reviewing the extensive literature on the relationship between cigarette tar content and health, reports have concluded that ‘low-tar’ cigarettes do not reduce risk for lung cancer and should not be recommended as healthy alternatives (11, 12). Only seven studies have investigated the association between cancer and cumulative tar exposure, an index accounting for changing smoking behaviors over time as well as tar content for different brands (13-19). Positive associations with cumulative tar exposure were reported for lung (13, 15, 16, 19), pancreatic (14), and oral cancer (17), but not for bladder cancer (18). Limitations of these prior studies include very few years of measured tar content (13-18), hospital-based control selection (13, 15-17), and limited or no analysis by histological subtype for lung cancer (13, 15, 16, 19). In addition, only two studies have adjusted for other measures of tobacco exposure (13, 18). In this study, we modified and applied the cumulative tar index to evaluate associations with cancers of the lung and UADT in a population-based case-control study conducted in Los Angeles County. In addition, we compared the associations with these cancers between cumulative tar and pack-years, as well as between histological subtypes. Furthermore, we measured the associations for cumulative tar and cancer after adjusting for pack-years to evaluate this additional information on cigarette composition.

Material and Methods:

Study Design and Population: Investigators conducted a population-based case-control study of lung and UADT cancers in Los Angeles County from 1999 to 2004. The Institutional Review Boards of the University of California, Los Angeles (UCLA) and the University of Southern California (USC) approved the study, and all participants provided their written informed consent. Further details of the original study design are available in earlier references (20, 21). In brief, newly diagnosed lung and UADT cancer patients were recruited from the USC Cancer Surveillance Program for Los Angeles County (USC CSP), a National Cancer Institute Surveillance, Epidemiology, and End-Results (SEER) Program cancer registry, through a rapid ascertainment system. Participants met the following inclusion criteria: (1) Residence in Los Angeles County at the time of diagnosis; (2) diagnosis age of 18–65 during the study period; (3) either English or Spanish speaking or accompanied by a translator during the interview. Among eligible patients, the recruitment rates for cases were 39% (611 of 1,556) for lung and 46% (601 of 1,301) for UADT cancer cases. The USC CSP collects pathology reports (over 95% of patients) and other diagnostic methods including magnetic resonance imaging and computed tomography scan with cancer reporting. In addition, the USC CSP classifies cancer diagnoses according to the International Classification of Diseases for Oncology, Third Edition (ICD-O-3). Among the 601 recruited UADT cases, there were 497 (82.7%) patients with squamous cell carcinoma of the oropharynx, larynx, and esophagus. In addition, 74 UADT patients were diagnosed with adenocarcinoma, all confined to the esophagus. The 611 lung cancer cases consisted of 508 (83.1%) patients with non-small cell lung cancer and 75 patients with small cell lung cancer. Non-small cell lung cancer includes adenocarcinoma (n=290), squamous cell carcinoma (n=95), and large cell carcinoma (n=115). Neighborhood-ascertained controls were matched to cases on sex and age (within five-years) and had a 79% recruitment rate among identified eligible matches.

Research staff interviewed each participant in-person using a standardized questionnaire. Questionnaire items included demographic characteristics; lifetime history of exposure to tobacco, alcohol, marijuana, and other recreational drugs; medical and occupational histories; and family history of cancer. Cigarette smoking information was collected on a yearly basis, including age at starting and quitting, brand and sub-brand details, number and frequency (i.e., cigarettes smoked per day/week/month/year), usual length of unsmoked cigarette (explained below), and smoke inhalation depth (deep, moderate, shallow, did not inhale). Participants also reported details for the lifetime use of cigars, pipes, chewing tobacco, and snuff.

Exposure estimation: We defined 'ever-smokers' as participants who smoked at least 100 cigarettes in their lifetime. To estimate their cumulative tar exposure, we first created a historical database of machine-measure tar yields from 39 reports of the Federal Trade Commission (FTC) between 1967 and 2000. We ascertained these reports from the University of California, San Francisco online archive, the Truth Tobacco Industry Documents, formerly known as the Legacy Tobacco Documents Library (22). The FTC started collecting these ratings in 1967 according to the standardized machine smoking protocol of the Cambridge filter method (23). Next, we used this longitudinal database to estimate cumulative tar exposure for all smoking participants. For each reported sub-brand in the questionnaire, we identified the closest match from the FTC report with respect to calendar year, size (Regular, King, 100mm, 120mm), design (Filter/Non-Filter), additive (Menthol/Non-menthol), and flavor (Full flavor, Light, Ultra-light). We calculated the average values for the ratings of multiple matches in the FTC report and for reports covering the same testing period. Missing tar ratings for years of reported exposure were imputed with the most recent rating in the database. For example, for pre-1967 smoking histories, we imputed tar ratings with values from 1967, when reporting began. Then, we modified Zang and Wynder's cumulative exposure index for tar by accounting for cigarette portion size and tar ratings by calendar year (13). We have reproduced the original index below:

[See Attached TIF Image- Equation 1]

where T is cumulative tar exposure (in kilograms), t is tar level per cigarette sub-brand (mg), D is days of smoking, C is cigarettes smoked per day, and B is all of the cigarette sub-brands smoked during the participant's lifetime. We summed tar exposure per year across all years of smoking to estimate cumulative tar. Study participants reported the portion of the unsmoked cigarette including the butt as ‘less than one-quarter’, ‘about one-quarter’, ‘about one-third’, and ‘about one-half or more’, which we specified as consumed portions of 7/8, 3/4, 2/3, and 1/4, respectively. While marijuana smoke also contains tar with many of the same components that are found in tobacco tar, including pro-carcinogenic polycyclic aromatic hydrocarbons (24), we have not observed clear associations between marijuana smoking and cancer risk in this study population (21) and did not estimate tar exposure from this source.

Statistical analysis: First, we calculated pack-years of cigarette smoking, cumulative tar exposure, and drink-years of alcohol consumption, lagged one year before the diagnosis year or reference year for controls. Then, we estimated the associations of cumulative tar exposure on risk for cancers of the lung and UADT in continuous and categorical analyses. The continuous measure was one interquartile range (IQR) increase in cumulative tar exposure; the categorical analysis used never smokers as the reference group and tertiles of exposure in ever smokers. Both the IQR and tertiles were based on tar distribution in the controls. We also analyzed associations with pack-years to compare cumulative tar exposure with the conventional measure of cumulative tobacco exposure. We used unconditional logistic regression to estimate odds ratios (OR's) and 95% confidence limits (CL’s), adjusting for potential covariates: age and sex (the matching variables), race/ethnicity, education level, and alcohol drink-years. In addition, we repeated the analyses for histological subtypes in both cancer groups. For lung cancer, we analyzed the common subtypes: squamous cell, small cell, large cell, and adenocarcinoma. For UADT cancer, we separately analyzed squamous cell carcinoma and esophageal adenocarcinoma. The logistic regression equation takes the general form:

[See Attached TIF Image- Equation 2]

where Y is the natural log odds of disease status (case vs. control), β_0 is the intercept when all predictors are zero, and β_i is the regression coefficient for each predictor i multiplied by some value X of the predictor. Furthermore, β_i is the natural logarithm of the odds ratio for the association between disease status with a unit increase in the covariate. Each coefficient is conditional on other coefficients in the model. Models estimating the association for cumulative tar exposure and each cancer subtype included beta coefficients for cumulative tar exposure, age, sex, race/ethnicity, and alcohol drink-years. We also modeled the comprehensive smoking index (CSI), a function of smoking duration, intensity, and time since cessation, including a half-life parameter (τ) of 10 years for the smoking effect (25). We calculated the area under the curve (AUC) for the receiver operating characteristic (ROC) to compare models for different measures of smoking exposure, including the linear combination of pack-years and tar (the sum of each coefficient multiplied by the record value). Furthermore, we evaluated the modification of the association between cumulative tar exposure and cancer by race/ethnicity and smoking status (current/former). We tested for a residual association of cumulative tar exposure in models adjusted for key covariates as well as pack-years (an additional beta coefficient in the logistic regression model) and according to subtype. In order to correct for potential false-positive findings, we re-ran these adjusted models using semi-Bayes ‘shrinkage’ estimation (26-29). In this analysis, prior coefficients with null associations are updated with coefficients from observed data to shrink associations in the logistic regression model toward the null. We assigned independent normal priors for targeted coefficients of tar exposure and cancer risk, with mean zero and variance 0.5 (corresponding to OR=1, 95% prior limits= 0.25, 4). Then, we combined the prior data with the observed data to calculate posterior estimates and 95% posterior limits. We performed our analyses using SAS version 9.4 (SAS Institute, Cary, NC, USA).

Results:

Distributions of socio-demographic characteristics, cigarette smoking and alcohol consumption are presented in Table 1. Age and sex are matched overall but lung and UADT cancers have different distributions. Sixty-five percent of all study participants reported smoking at least 100 cigarettes in their lifetime. As expected, we observed positive associations between cigarette pack-years and risk of both lung and UADT cancers, stronger for lung than for UADT cancer. We observed positive associations between former vs. never smoking and both cancer types. For current smoking, we observed a positive association for lung cancer and an inverse association for UADT cancer (OR=0.59, 95% CL's=0.41, 0.85). The association for former smoking was higher than current smoking for both cancer types, possibly due to the induction time between smoking exposure and cancer onset. In addition, we observed a positive association between the highest level of drink-years (>80 drink-years vs. never drinkers) and UADT cancer (OR = 2.28, 95% CL's = 1.54, 3.37) after adjusting for covariates including cigarette pack-years. The test for trend across categories of drink-years also suggested a positive association (p-trend < 0.0001), consistent with published reports (30). We did not observe associations for consumption of cigars, pipes, snuff, marijuana, or chewing tobacco, or for passive smoking duration in either cancer type probably due to sparse exposure data (data not shown).

Distributions of cumulative tar exposure and pack-years for both cancer types, as well as controls, are shown in Table 2. Overall mean (standard deviation) tar exposure (in kilograms) was 0.86 (1.58) in controls, 2.90 (2.86) in lung cancer cases, and 2.19 (2.78) in UADT cancer cases. Corresponding statistics for pack-years were 9.09 (15.39), 30.50 (24.33), and 21.88 (23.69). Point estimates for the mean of both measures were higher in men than women for all three groups. Cumulative tar exposure was highly correlated with pack-years, based on Pearson r=0.90 in ever smoking controls. Nearly 25% of the 1,468 smokers reported a brand that was unknown or unlisted in the FTC Reports; 76 smokers had completely unknown/unlisted brand information; 5 smokers had missing information for portion size (length of unsmoked cigarette).

Tables 3-1, 3-2, and 4 display adjusted odds ratios (OR’s) and 95% confidence limits (CL’s) for associations of lung and UADT cancers, overall and by histological subtype, with cumulative tar exposure and pack-years. Models were adjusted for age (in fine categories), race/ethnicity, sex, years of education, and drink-years. Positive associations were evident in both cancer types and corresponding subtypes, by trend tests (all $p<0.05$) and by OR estimates for categorical and continuous analyses. OR's for one-IQR increase of tar exposure (0.96kg) and pack-years (12.81) did not vary much by subtype within each cancer group. For lung cancer and subtypes (Table 3-1: overall, small cell, squamous cell; Table 3-2: adenocarcinoma, large cell), the OR estimates for pack-years were greater than tar exposure by one-IQR increase (overall: pack-years- OR= 2.16, 95% CL's= 1.96, 2.39; tar- OR= 1.61, 95% CL's= 1.50, 1.73). However, the confidence intervals for corresponding tertiles of smoking exposures overlapped, suggesting no obvious difference. Compared to never smokers, the second and third tertiles of exposure were associated with overall lung cancer and with the major histological subtypes. For squamous cell lung cancer, the second tertile excluded the null for tar (OR= 5.09, 95% CL's= 2.08, 12.41) but not for pack-years (OR= 2.49, 95% CL's= 0.93, 6.65). Furthermore, associations in the higher tertiles for tar and pack-years were generally higher for small cell, squamous, and large cell carcinoma of the lung than for lung adenocarcinoma.

For overall UADT cancer (Table 4), estimated associations for tar and pack-years were generally less than for lung cancer (one IQR increased tar exposure- OR= 1.21, 95% CL's= 1.13, 1.29). With respect to overall disease and subtypes, the third tertile for tar and pack-years were associated with increased cancer risk. The second tertile for tar exposure, not pack-years, was associated with overall disease (OR= 1.53, 95% CL's= 1.11, 2.10), UADT squamous cell carcinoma (OR= 1.43, 95% CL's= 1.02, 2.01), and esophageal adenocarcinoma (OR=2.52, 95% CL’s= 1.21, 5.25). The subtypes did not appear to differ by tertiles of tar or pack-years. For the ROC analysis, the area under the curve (AUC) was lower for cumulative tar compared to pack-years, except for esophageal adenocarcinoma (p>0.05; Supplementary Table 1). The AUC’s for the CSI were higher than pack-years for overall lung (77.3% vs. 76.7%) and UADT cancers (66.6% vs. 66.1%). However, the AUC’s for combined cumulative tar and pack-years were no different than pack-years alone.

Table 5 displays estimates for associations between cumulative tar exposure and cancer, with additional adjustment for pack-years in maximum-likelihood and semi-Bayes corrected models. Associations with lung cancer were evident in the second exposure tertiles even after semi-Bayes adjustment: overall lung cancer (semi-Bayes odds ratio-SBOR=1.55, 95% posterior limits-PL’s=1.07, 2.24); small cell (SBOR=2.73, 95% PL’s=1.33, 5.61); adenocarcinoma (SBOR=1.56, 95% PL’s =1.02, 2.37); and in the third tertile of large cell lung cancer (SBOR= 2.51, 95% PL's = 1.20, 5.25). We also observed positive trends for cumulative tar in small cell and large cell lung cancer in these models (p-trend < 0.05). However, we did not observe associations between tar and cancer by per-IQR increase in exposure. Moreover, we did not observe positive associations for cumulative tar exposure in UADT cancer or subtypes after adjusting for pack-years in maximum likelihood or semi-Bayes models.

We observed a higher association between tar and overall lung cancer for Whites than non-Whites in the highest tertile of exposure compared to never smokers (Supplementary Table 2: OR=16.65, 95% CL's= 10.22, 27.13 vs. OR=5.89, 95% CL's = 3.43, 10.10; p for multiplicative interaction <0.05). The association did not differ for overall UADT cancer (Supplementary Table 2: p for multiplicative interaction > 0.05). After we adjusted for pack-years in semi-Bayes corrected models, the positive trend between cumulative tar and lung cancer was apparent in Whites (overall, small cell, adenocarcinoma, and large cell disease, p-trend <0.05) but not non-Whites. We did not observe modification of the association between cumulative tar and either cancer type by smoking status (current vs. former smokers), comparing the highest tertile of exposure to the first tertile (Supplementary Table 3: p for multiplicative interaction > 0.05).

Discussion:

Our study of 611 lung cancer patients, 601 UADT cancer patients, and 1,040 controls found that cumulative tar exposure is highly correlated with pack-years and is positively associated with lung and UADT cancers. An increase of about one kilogram lifetime cumulative tar exposure was associated with approximately a 61% increased risk of lung cancer and about a 21% increased risk of UADT cancer. This concurs with prior evidence that tobacco smoking is a stronger risk factor for lung cancer than UADT cancer (31). To estimate cumulative tar exposure, we modified Zang and Wynder's cumulative lifetime tar index by incorporating historical tar values and cigarette portion size. We believe that the modification is closer to the true tobacco exposure. The major advantage of using this index compared to pack-years is that cumulative tar accounts for the attributable risk of particulate carcinogens and could potentially sort out the remaining risk by other carcinogens not directly associated with tar exposure, such as those in gas-phase. Our reported positive association between cumulative tar and lung cancer risk is consistent with other reports (13, 15, 16, 19). While most studies of cumulative tar and cancer risk have relied on one or two years of reported cigarette tar ratings (13-18), one study of lung cancer in Tasmania estimated cumulative tar based on 17 reports of machine-measured tar yields published between 1961and 1996 (19). Our study collected yields from 39 reports covering 25 testing years between 1967 and 2000. Furthermore, this appears to be the first association reported for overall UADT cancer.

We detected associations between cumulative cigarette tar exposure and lung cancer subtypes after adjusting for pack-years, the standard measure of cumulative cigarette smoking. Zang and Wynder (13) previously noted a positive trend for cumulative tar and lung cancer even after restricting to higher levels of pack-years. However, the residual association we observed could be a result of the strong correlation between cumulative tar and pack-years (r=0.90). We applied semi-Bayes shrinkage estimation to reduce the potential for false positive findings from variance inflation or multiple comparisons. Cumulative tar neither improved risk models for any case group compared to pack-years in the ROC analysis, nor was cumulative tar associated with UADT cancer after we adjusted for pack-years.

Furthermore, we detected positive associations between cumulative tar exposure and major subtypes of lung cancer, with higher estimates for small cell, squamous cell, and large cell cancer than for adenocarcinoma. This concurs with previous reports (13, 15, 16) for a higher association for cumulative tar in “Kreyberg type I” cancers (squamous, epidermoid, oat, small, and large cell) compared to “Kreyberg type II” (adenocarcinoma). This also concurs with two published meta-analyses of cigarette smoking and lung cancer, which reported greater estimates for small cell and squamous cell cancer than for adenocarcinoma and large cell carcinoma (2, 3). However, we had limited sample size to detect small differences between lung and UADT cancer subtypes.

In the United States, although the adult smoking prevalence decreased by 60% between 1965 and 2014, the risk of smoking-related lung cancer and mortality has increased (11, 32, 33). Increasing risk in both sexes may partly be attributed to changes in cigarette design and composition in the past 50 years (11, 12, 34-38). For example, filtered cigarettes, which were introduced in the 1950’s, are associated with deeper inhalation and smaller particle size of smoke, which may increase the deposition of tobacco carcinogens throughout the airway (12, 36, 37, 39, 40). In addition, cigarette content of tobacco-specific nitrosamines increased between 17% and 73% from 1978 to 1995 when measured under standard FTC smoking conditions (11, 12, 37). Furthermore, these changes may explain the shift in smoking-related lung cancer incidence from squamous cell to adenocarcinoma (11, 12, 34-38). However, while smoking-related lung cancer has increased, average sales-weighted tar yield decreased by 44% between 1968 and 1998 (22; Document ID yqpk0154). Harris (41) provided a possible explanation for this disparity when he reported weak associations between FTC tar rating and tobacco-specific nitrosamines (r^2=0.38 and 0.76 for NNN and NNK, respectively; both p<0.01). In spite of this disadvantage, the cumulative tar index allowed us to simultaneously account for changes in tar yield and smoking behavior (e.g., cigarettes per day) over time.

This population-based case-control study included histologically confirmed cases, relatively good statistical power for the main cancer types, and information from both a lifetime exposure questionnaire and longitudinal federal government reports. The first limitation to this study was that machine-testing of cigarette yields has been shown to underestimate smoking exposures (12). Smokers compensate their breathing to achieve a steady nicotine dose (12). In addition, cigarettes have been engineered to produce misleadingly lower yields under machine smoking conditions, such as perforated filters which smokers cover with their fingers (11, 12). Therefore, we most likely underestimated the cumulative tar index and biased the odds ratios non-differentially toward the null. Second, the fact that we observed null associations with cancer in the lowest tertile of tar exposure/pack-years was likely due to insufficient power to detect small associations. Third, residual confounding by unmeasured human papillomavirus (HPV) infection may have biased the associations with UADT cancer, especially squamous cell carcinoma. The direction of bias is difficult to evaluate because the reports on the association between tobacco smoking and HPV infection have been inconsistent (42-45), although our reported associations appear to be biased toward the null. Fourth, adjusting the tar-cancer associations for pack-years may have biased the associations toward the null. Smokers who switch to lower-tar brands likely increase their smoking intensity (compensation), and adjusting for an intermediate variable generally biases estimates toward the null (12, 46). Fifth, selection bias may have occurred if tobacco exposure (measured as pack-years or cumulative tar) was associated with participation differentially for eligible cases and controls. Ten percent of eligible UADT cancer cases and 25% of eligible lung cancer cases died before they could be interviewed. Selective-survival bias could have occurred because smoking is associated with shorter survival time for these cancers (47, 48). Therefore, we would expect a downward bias in OR estimation in this scenario that nonparticipation was selectively greater in more highly exposed cases. Given these limitations, this study does not support the claim of a null or inverse association between 'low exposure' to tobacco smoke and risk of these cancer types.

In conclusion, our study suggests that cumulative tar exposure is associated with cancer risk and is associated with small and large cell lung cancer after adjusting for pack-years. The Family Smoking Prevention and Tobacco Control Act of 2009 (USA) requires tobacco product manufacturers and importers to report harmful and potentially harmful constituents (HPHC's) to the FDA, including 93 carcinogens, toxicants, and additive substances measured from a machine smoking regimen (49). Although machine smoking protocols have limited ability to reflect real exposure to smokers, researchers have found that they "may be the limit of current scientific assessment of differences between brands that can be used for regulatory assessment of product toxicity" (10). It is possible that novel exposure measures incorporating smoking duration and intensity, as well as constituent levels by tobacco product could help to identify people at high risk for cancer who would benefit from screening and/or tobacco cessation intervention. The present study suggests that cumulative tar, a crude estimate of total smoke constituent exposure, may improve exposure assessment and risk estimation particularly for small cell and large cell lung subtypes. Biomarkers of tobacco smoke constituents should also continue to be identified to improve cancer risk assessment (50). Public health messages should meanwhile focus on abstaining from all tobacco products, regardless of tar content (11, 12). However, tobacco products should also be strictly regulated to deliver lower doses of carcinogens, in terms of total particulate matter or specific harmful constituents such as tobacco-specific nitrosamines (38).

Acknowledgements: The authors thank all of the Los Angeles study participants for their time and effort. This research was supported by the US National Institutes of Health (Grant Numbers ES06718, ES011667, CA90833, CA077954, CA96134, DA11386] and the Alper Research Center for Environmental Genomics of the UCLA Jonsson Comprehensive Cancer Center. The first author would also like to acknowledge his support from a training grant from the US National Cancer Institute (T32-CA09142). He would also like to thank Mr. Daniel Nikzad for his help compiling the FTC reports and Dr. Alan Fu for his statistical analysis advice.

References:

1. International Agency for Research on Cancer (IARC). Tobacco Smoking. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Personal Habits and Indoor Combustions; volume 100E. Lyon, France. 2012.

2. Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. Lung Cancer 2001; 31:139-48.

3. Pesch B, Kendzia B, Gustavsson P, Jockel KH, Johnen G, Pohlabeln H, Olsson A, Ahrens W, Gross IM, Bruske I, Wichmann HE, Merletti F, Richiardi L, Simonato L, Fortes C, Siemiatycki J,

Parent ME, Consonni D, Landi MT, Caporaso N, Zaridze D, Cassidy A, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, Stucker I, Fabianova E, Dumitru RS, Bencko V, Foretova L, Janout V, Rudin CM, Brennan P, Boffetta P, Straif K, Bruning T. Cigarette smoking and lung cancer--relative risk estimates for the major histological types from a pooled analysis of case-control studies. Int J Cancer 2012; 131:1210-9.

4. International Agency for Research on Cancer (IARC). Tobacco Smoke and Involuntary Smoking. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans; volume 83. Lyon, France. 2004.

5. Lam TH and Ho SY. Tobacco. In: Detels R, Gulliford M, Karim QA, Tan CC, editors. Oxford Textbook of Public Health, Sixth Edition. New York: Oxford University Press; 2015. p. 1217-1232.

6. Rodgman A and Perfetti TA. The chemical components of tobacco and tobacco smoke. Boca Raton: CRC Press; 2009.

7. Smith CJ, Perfetti TA, Garg R, Hansch C. IARC carcinogens reported in cigarette mainstream smoke and their calculated log P values. Food Chem Toxicol 2003; 41:807-17.

8. Zhang ZF, Boffetta P, Neugut AI, La Vecchia C. Cancer epidemiology and public health. In: Detels R, Gulliford M, Karim QA, Tan CC, editors. Oxford Textbook of Public Health, Sixth Edition. New York: Oxford University Press; 2015. p. 923-944.

9. World Health Organization. WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER Package. Geneva: World Health Organization; 2008.

10. Burns DM, Dybing E, Gray N, Hecht S, Anderson C, Sanner T, O'Connor R, Djordjevic M, Dresler C, Hainaut P, Jarvis M, Opperhuizen A, Straif K. Mandated lowering of toxicants in

cigarette smoke: a description of the World Health Organization TobReg proposal. Tob Control 2008; 17:132-41.

11. US Department of Health and Human Services. The Health Consequences of Smoking: 50 Years of Progress. A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.

12. National Cancer Institute. Risks Associated with Smoking Cigarettes with Low Machine-Measured Yields of Tar and Nicotine. Smoking and Tobacco Control Monograph No. 13. Bethesda, Maryland: NIH Publication No. 02-5047, October 2001.

13. Zang EA, Wynder EL. Cumulative tar exposure. A new index for estimating lung cancer risk among cigarette smokers. Cancer 1992; 70:69-76.

14. Garabrant DH, Held J, Langholz B, Peters JM, Mack TM. DDT and related compounds and risk of pancreatic cancer. J Natl Cancer Inst 1992; 84:764-71.

15. Harris RE, Zang EA, Anderson JI, Wynder EL. Race and sex differences in lung cancer risk associated with cigarette smoking. Int J Epidemiol 1993; 22:592-9.

16. Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. J Natl Cancer Inst 1996; 88:183-92.

17. Muscat JE, Richie JP, Jr., Thompson S, Wynder EL. Gender differences in smoking and risk for oral cancer. Cancer Res 1996; 56:5192-7.

18. Castelao JE, Yuan JM, Skipper PL, Tannenbaum SR, Gago-Dominguez M, Crowder JS, Ross RK, Yu MC. Gender- and smoking-related bladder cancer risk. J Natl Cancer Inst 2001; 93:538-45.

19. Blizzard L, Dwyer T. Case-control study of lung cancer during 1994-1997 in the birth cohort in Tasmania, Australia, with an excess of female cases during 1983-1992. Cancer Causes Control 2003; 14:123-9.

20. Cui Y, Morgenstern H, Greenland S, Tashkin DP, Mao J, Cao W, Cozen W, Mack TM, Zhang ZF. Polymorphism of Xeroderma Pigmentosum group G and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. Int J Cancer 2006; 118:714-20.

21. Hashibe M, Morgenstern H, Cui Y, Tashkin DP, Zhang ZF, Cozen W, Mack TM, Greenland S. Marijuana use and the risk of lung and upper aerodigestive tract cancers: results of a population-based case-control study. Cancer Epidemiol Biomarkers Prev 2006; 15:1829-34.

22. Truth Tobacco Industry Documents. University of California, San Francisco Library and Center for Knowledge Management. Available from: https://industrydocuments.library.ucsf.edu/tobacco/. Document ID's: jgxf0177; rtwl0027; ftwl0027; tznb0014; jhdl0154; pxyl0082; trgh0100; kswl0027; fsgh0100; mrwl0027; rqyv0041; jrdl0154; xrdl0154; fxlp0034; fqlk0154; hrdl0154; ltnh0067; krky0014; grdl0154; frdl0154; xspj0139; lsjh0111; kgwl0154; xgdn0047; rkdl0154; sypk0154; typk0154; yypk0154; hhwl0154; qqdl0154; pqdl0154; qyvl0154; lsdl0154; khwl0154; msdl0154; yfgh0094; pgwl0154; hgwl0154; yqpk0154.

23. Ogg CL. Determination of particulate matter and alkaloids (as nicotine) in cigarette smoke. J Assoc Off Agricultural Chem 1964; 47.

24. Moir D, Rickert WS, Levasseur G, Larose Y, Maertens R, White P, Desjardins S. A comparison of mainstream and sidestream marijuana and tobacco cigarette smoke produced under two machine smoking conditions. Chem Res Toxicol 2008; 21:494-502.

25. Dietrich T, Hoffmann K. A comprehensive index for the modeling of smoking history in periodontal research. J Dent Res 2004; 83:859-63.

26. Greenland S. A semi-Bayes approach to the analysis of correlated multiple associations, with an application to an occupational cancer-mortality study. Stat Med 1992; 11:219-30.

27. Greenland S. Bayesian perspectives for epidemiological research. II. Regression analysis. Int Journal Epidemiol 2007; 36:195-202.

28. Greenland S, Christensen R. Data augmentation priors for Bayesian and semi-Bayes analyses of conditional-logistic and proportional-hazards regression. Stat Med 2001; 20:2421-8.

29. Sullivan SG, Greenland S. Bayesian regression in SAS software. Int J Epidemiol 2013; 42:308-17.

30. International Agency for Research on Cancer (IARC). Consumption of Alcoholic Beverages. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Personal Habits and Indoor Combustions; volume 100E. Lyon, France. 2012.

31. Gandini S, Botteri E, Iodice S, Boniol M, Lowenfels AB, Maisonneuve P, Boyle P. Tobacco smoking and cancer: a meta-analysis. Int J Cancer 2008; 122:155-64.

32. Thun MJ, Carter BD, Feskanich D, Freedman ND, Prentice R, Lopez AD, Hartge P, Gapstur SM. 50-year trends in smoking-related mortality in the United States. N Engl J Med 2013; 368:351-64.

33. Jamal A, Homa DM, O'Connor E, Babb SD, Caraballo RS, Singh T, Hu SS, King BA. Current Cigarette Smoking Among Adults - United States, 2005-2014. MMWR Morb Mortal Wkly Rep 2015; 64:1233-40.

34. Thun MJ, Lally CA, Flannery JT, Calle EE, Flanders WD, Heath CW, Jr. Cigarette smoking and changes in the histopathology of lung cancer. J Natl Cancer Inst 1997; 89:1580-6.

35. Harris JE, Thun MJ, Mondul AM, Calle EE. Cigarette tar yields in relation to mortality from lung cancer in the cancer prevention study II prospective cohort, 1982-8. BMJ 2004; 328:72.

36. Burns DM, Anderson CM, Gray N. Has the lung cancer risk from smoking increased over the last fifty years? Cancer Causes Control 2011; 22:389-97.

37. Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. J Toxicol Environ Health 1997; 50:307-64.

38. Hecht SS. It is time to regulate carcinogenic tobacco-specific nitrosamines in cigarette tobacco. Cancer Prev Res (Phila) 2014; 7:639-47.

39. Bates DV, Fish BR, Hatch TF, Mercer TT, Morrow PE. Deposition and retention models for internal dosimetry of the human respiratory tract. Task group on lung dynamics. Health Phys 1966; 12:173-207.

40. National Research Council Committee on Health Risks of Exposure to Radon. Health Effects of Exposure to Radon: BEIR VI. Washington DC: National Academies Press, National Academy of Sciences. 1999.

41. Harris JE. Smoke yields of tobacco-specific nitrosamines in relation to FTC tar level and cigarette manufacturer: analysis of the Massachusetts Benchmark Study. Public Health Rep 2001; 116:336-43.

42. Furniss CS, Kraunz KS, Liu M, Peters E, Kelsey KT. Human papillomavirus 16 is associated with drinking and smoking behaviors and tumor suppressor gene hypermethylation in head and neck squamous cell carcinoma. Abstract from the Proceedings of the American Association for Cancer Research 2005; 65:120.

43. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, Viscidi R. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. J Natl Cancer Inst 2008; 100:407-20.

44. Maruyama H, Yasui T, Ishikawa-Fujiwara T, Morii E, Yamamoto Y, Yoshii T, Takenaka Y, Nakahara S, Todo T, Hongyo T, Inohara H. Human papillomavirus and p53 mutations in head and neck squamous cell carcinoma among Japanese population. Cancer Sci 2014; 105:409-17.

45. Minkoff H, Feldman JG, Strickler HD, Watts DH, Bacon MC, Levine A, Palefsky JM, Burk R, Cohen MH, Anastos K. Relationship between smoking and human papillomavirus infections in HIV-infected and -uninfected women. J Infect Dis 2004; 189:1821-8.

46. Rothman KJ GS, Lash TL. Modern Epidemiology, Third Edition. Philadelphia, PA: Lippincott Williams and Wilkins; 2008. p. 656

47. Janjigian YY, McDonnell K, Kris MG, Shen R, Sima CS, Bach PB, Rizvi NA, Riely GJ. Pack-years of cigarette smoking as a prognostic factor in patients with stage IIIB/IV nonsmall cell lung cancer. Cancer 2010;116:670-5.

48. Gillison ML, Zhang Q, Jordan R, Xiao W, Westra WH, Trotti A, Spencer S, Harris J, Chung CH, Ang KK. Tobacco smoking and increased risk of death and progression for patients with p16-positive and p16-negative oropharyngeal cancer. J Clin Oncol 2012;30:2102-11.

49. Reporting Harmful and Potentially Harmful Constituents in Tobacco Products and Tobacco Smoke Under Section 904(a)(3) of the Federal Food, Drug, and Cosmetic Act, 2015. FDA Center for Tobacco Products. Available from: http://www.fda.gov/TobaccoProducts/GuidanceComplianceRegulatoryInformation/ucm297752.htm

50. Yuan JM, Butler LM, Stepanov I, Hecht SS. Urinary tobacco smoke-constituent biomarkers for assessing risk of lung cancer. Cancer Res 2014;74:401-11.

Table 1. Distributions of sociodemographic and consumption characteristics among lung cancer cases (n=611) and UADT cases (n=601) compared to cancer-free controls (n=1,040)

Variable	Controls N (%)	Lung N (%)	Adjusted OR (95% CL's)	UADT N (%)	Adjusted OR (95% CL's)
Age[a]					
17-34	51 (4.9)	4 (0.7)	–	32 (5.3)	–
35-44	170 (16.3)	57 (9.3)	–	77 (12.8)	–
45-54	500 (48.1)	301 (49.3)	–	267 (44.4)	–
>54	319 (30.7)	249 (40.8)	–	225 (37.4)	–
Sex[a]					
Male	623 (59.9)	303 (49.6)	–	454 (75.5)	–
Female	417 (40.1)	308 (50.4)	–	147 (24.5)	–
Race/Ethnicity[b]					
Caucasian	634 (61)	359 (58.9)	1 (Reference)	341 (56.9)	1 (Reference)
African-American	102 (9.8)	96 (15.7)	1.99 (1.38, 2.89)	69 (11.5)	1.05 (0.72, 1.53)
Hispanic	204 (19.6)	70 (11.5)	0.95 (0.62, 1.45)	109 (18.2)	0.70 (0.48, 1.00)
Asian/Pacific-Islander	62 (6.0)	70 (11.5)	4.70 (3.05, 7.22)	64 (10.7)	2.71 (1.80, 4.09)
Other	37 (3.6)	15 (2.5)	0.67 (0.32, 1.38)	16 (2.7)	0.63 (0.32, 1.23)
Education (years of schooling)[b]					
<12	116 (11.2)	107 (17.5)	1 (Reference)	126 (21.0)	1 (Reference)
12	184 (17.7)	158 (25.9)	0.57 (0.36, 0.89)	147 (24.5)	0.60 (0.40, 0.90)
13-15	272 (26.2)	186 (30.4)	0.56 (0.36, 0.88)	156 (26.0)	0.49 (0.33, 0.73)
16	209 (20.1)	89 (14.6)	0.49 (0.30, 0.80)	103 (17.1)	0.43 (0.28, 0.67)
>16	258 (24.8)	71 (11.6)	0.37 (0.22, 0.61)	69 (11.5)	0.26 (0.17, 0.42)

			p-trend=0.0004		p-trend <0.0001
Cigarette Smoking Status[c]					
Never	491 (47.2)	110 (18.0)	1 (Reference)	182 (30.3)	1 (Reference)
Former	371 (35.7)	390 (63.8)	4.34 (3.27, 5.75)	338 (56.2)	1.71 (1.32, 2.21)
Current	177 (17.1)	111 (18.2)	2.30 (1.60, 3.30)	81 (13.5)	0.59 (0.41, 0.85)
Cigarette Pack-Years[c]					
Never Smokers	491 (47.2)	110 (18.0)	1 (Reference)	182 (30.3)	1 (Reference)
≤20	355 (34.1)	105 (17.2)	1.41 (1.02, 1.96)	150 (25.0)	0.96 (0.73, 1.27)
>20-40	137 (13.2)	213 (34.9)	8.36 (5.86, 11.92)	147 (24.5)	1.88 (1.34, 2.64)
>40	56 (5.4)	183 (30.0)	21.59 (13.85, 33.66)	122 (20.3)	3.37 (2.21, 5.14)
			p-trend < 0.0001		p-trend <0.0001
Alcohol Drink-Years[d]					
Never Drinkers	264 (25.4)	170 (27.8)	1 (Reference)	117 (19.5)	1 (Reference)
≤40	586 (56.3)	260 (42.6)	0.67 (0.49, 0.90)	232 (38.6)	0.94 (0.70, 1.27)
>40	189 (18.2)	180 (29.5)	0.77 (0.52, 1.14)	250 (41.6)	1.78 (1.26, 2.51)
			p-trend = 0.14		p-trend=0.0007

a. Age and sex are matching variables and their odds ratios are not valid

b. Models for race/ethnicity and education adjusted for each other, plus age, sex, pack-years, and drink-years

c. Models adjusted for age, race/ethnicity, sex, education, and drink-years

d. Models adjusted for age, race/ethnicity, sex, education, and smoking pack-years

Table 2. Distributions of cumulative tar and pack-years for cancer cases and cancer-free controls, stratified by sex

	Controls					Lung Cancer		UADT Cancer	
				Smokers					
Variable	N (%)	Mean (SD)	IQR	Tertile 1	Tertile 2	N (%)	Mean (SD)	N (%)	Mean (SD)
Cumulative Tar (kg)									
Overall	997 (95.9)	0.86 (1.58)	0.96	0.43	2.08	598 (97.9)	2.90 (2.86)	575 (95.7)	2.19 (2.78)
Men	589 (56.6)	1.06 (1.77)	1.55	0.55	2.48	297 (48.6)	3.73 (3.17)	435 (72.4)	2.50 (2.89)
Women	408 (39.2)	0.58 (1.21)	0.53	0.30	1.45	301 (49.3)	2.09 (2.25)	140 (23.3)	1.24 (2.17)
Missing	43 (4.1)					13 (2.1)		26 (4.3)	
Cigarette Pack-Years									
Overall	1039 (99.9)	9.09 (15.39)	12.81	5.25	21.00	611 (100)	30.50 (24.33)	601 (100)	21.88 (23.69)
Men	622 (59.8)	10.68 (16.46)	17.00	6.50	25.29	303 (49.6)	36.90 (25.27)	454 (75.5)	24.29 (23.66)
Women	417 (40.1)	6.73 (13.31)	6.47	3.75	17.10	308 (50.4)	24.20 (21.63)	147 (24.5)	14.43 (22.26)
Missing	1 (0.1)					0		0	

Table 3-1. Cumulative tar and pack-years and risk of overall, small cell, and squamous cell lung cancer

Variable	Controls N (%)	Lung N (%)	Adjusted OR (95% CL's)[a]	Small Cell N (%)	Adjusted OR (95% CL's)[a]	Squamous N (%)	Adjusted OR (95% CL's)[a]
Cumulative Tar							
Per IQR increase			1.61 (1.50, 1.73)		1.63 (1.44, 1.84)		1.71 (1.51, 1.92)
Never Smokers	491 (47.2)	110 (18)	1 (Reference)	4 (5.3)	1 (Reference)	8 (8.4)	1 (Reference)
Tertile 1	169 (16.3)	34 (5.6)	0.95 (0.60, 1.49)	2 (2.7)	1.70 (0.29, 9.80)	3 (3.2)	0.93 (0.24, 3.65)
Tertile 2	169 (16.3)	130 (21.3)	3.48 (2.47, 4.90)	18 (24)	17.63 (5.22, 59.61)	18 (18.9)	5.09 (2.08, 12.41)
Tertile 3	168 (16.2)	324 (53)	10.43 (7.34, 14.81)	49 (65.3)	45.97 (13.68, 154.48)	63 (66.3)	19.00 (8.02, 45.03)
			p-trend<0.0001		p-trend<0.0001		p-trend<0.0001
Cigarette Pack-Years							
Per IQR increase			2.16 (1.96, 2.39)		2.26 (1.89, 2.70)		2.41 (2.03, 2.88)
Never Smokers	491 (47.2)	110 (18)	1 (Reference)	4 (5.3)	1 (Reference)	8 (8.4)	1 (Reference)
Tertile 1	183 (17.6)	27 (4.4)	0.75 (0.47, 1.21)	1 (1.3)	0.86 (0.09, 8.06)	3 (3.2)	1.06 (0.27, 4.15)
Tertile 2	183 (17.6)	86 (14.1)	2.14 (1.49, 3.07)	12 (16)	10.58 (3.14, 35.58)	10 (10.5)	2.49 (0.93, 6.65)
Tertile 3	182 (17.5)	388 (63.5)	11.82 (8.42, 16.61)	58 (77.3)	54.30 (16.76, 175.96)	74 (77.9)	22.68 (9.79, 52.57)
			p-trend<0.0001		p-trend<0.0001		p-trend<0.0001

a. Models adjusted for age, race/ethnicity, sex, years of education, and drink-years

Table 3-2. Cumulative tar and pack-years and risk of lung adenocarcinoma and large-cell lung cancer

Variable	Controls N (%)	Adenocarcinoma N (%)	Adjusted OR (95% CL's)[a]	Large Cell N (%)	Adjusted OR (95% CL's)[a]
Cumulative Tar					
Per IQR increase			1.42 (1.30, 1.54)		1.62 (1.45, 1.81)
Never Smokers	491 (47.2)	77 (26.6)	1 (Reference)	14 (12.2)	1 (Reference)
Tertile 1	169 (16.3)	17 (5.9)	0.69 (0.38, 1.23)	5 (1.7)	1.12 (0.38, 3.27)
Tertile 2	169 (16.3)	70 (24.1)	2.64 (1.76, 3.96)	19 (6.6)	4.02 (1.88, 8.60)
Tertile 3	168 (16.2)	121 (41.7)	5.71 (3.75, 8.70)	75 (25.9)	17.25 (8.48, 35.10)
			p-trend<0.0001		p-trend<0.0001
Cigarette Pack-Years					
Per IQR increase			1.77 (1.58, 1.98)		2.18 (1.86, 2.55)
Never Smokers	491 (47.2)	77 (26.6)	1 (Reference)	14 (12.2)	1 (Reference)
Tertile 1	183 (17.6)	15 (5.2)	0.59 (0.32, 1.08)	3 (2.6)	0.68 (0.19, 2.43)
Tertile 2	183 (17.6)	46 (15.9)	1.66 (1.07, 2.58)	13 (11.3)	2.47 (1.09, 5.56)
Tertile 3	182 (17.5)	152 (52.4)	6.56 (4.39, 9.81)	85 (73.9)	18.16 (9.11, 36.22)
			p-trend<0.0001		p-trend<0.0001

a. Models adjusted for age, race/ethnicity, sex, years of education, and drink-years

Table 4. Cumulative tar and pack-years and risk of overall UADT cancer, UADT squamous cell carcinoma, and esophageal adenocarcinoma

Variable	Controls N (%)	UADT N (%)	Adjusted OR (95% CL's)[a]	Squamous N (%)	Adjusted OR (95% CL's)[a]	Esophageal Adenocarcinoma N (%)	Adjusted OR (95% CL's)[a]
Cumulative Tar							
Per IQR increase			1.21 (1.13, 1.29)		1.18 (1.10, 1.27)		1.27 (1.13, 1.43)
Never Smokers	491 (47.2)	182 (30.3)	1 (Reference)	149 (30)	1 (Reference)	18 (24.3)	1 (Reference)
Tertile 1	169 (16.3)	50 (8.3)	0.70 (0.48, 1.03)	41 (8.2)	0.69 (0.46, 1.05)	6 (8.1)	0.82 (0.31, 2.19)
Tertile 2	169 (16.3)	122 (20.3)	1.53 (1.11, 2.10)	98 (19.7)	1.43 (1.02, 2.01)	17 (23.0)	2.52 (1.21, 5.25)
Tertile 3	168 (16.2)	221 (36.8)	1.97 (1.42, 2.74)	185 (37.2)	1.79 (1.26, 2.54)	31 (41.9)	3.54 (1.69, 7.43)
			p-trend<0.0001		p-trend=0.0005		p-trend=0.0003
Cigarette Pack-Years							
Per IQR increase			1.36 (1.25, 1.49)		1.33 (1.21, 1.46)		1.46 (1.24, 1.73)
Never Smokers	491 (47.2)	182 (30.3)	1 (Reference)	149 (30)	1 (Reference)	18 (24.3)	1 (Reference)
Tertile 1	183 (17.6)	50 (8.3)	0.69 (0.48, 1.01)	41 (8.2)	0.69 (0.46, 1.04)	7 (9.5)	0.86 (0.34, 2.19)
Tertile 2	183 (17.6)	106 (17.6)	1.24 (0.90, 1.70)	85 (17.1)	1.18 (0.84, 1.67)	14 (18.9)	1.83 (0.85, 3.91)
Tertile 3	182 (17.5)	263 (43.8)	2.39 (1.75, 3.26)	222 (44.7)	2.22 (1.59, 3.09)	35 (47.3)	3.69 (1.82, 7.52)
			p-trend<0.0001		p-trend<0.0001		p-trend=0.0002

a. Models adjusted for age, race/ethnicity, sex, years of education, and drink-years

Table 5. Cumulative tar exposure and risk of lung and UADT cancer and subtypes, adjusted for pack-years in maximum-likelihood and semi-Bayes models

Cancer Type	Maximum-Likelihood Adjusted OR (95% CL's)[a]	Semi-Bayes Adjusted OR (95% Posterior Limits)[a]
Overall Lung Cancer		
Per IQR Increase	0.99 (0.86, 1.13)	0.96 (0.84, 1.10)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.73 (0.46, 1.17)	0.73 (0.47, 1.12)
Tertile 2	1.47 (0.98, 2.21)	1.55 (1.07, 2.24)
Tertile 3	1.56 (0.88, 2.76)	1.47 (0.88, 2.45)
	p-trend = 0.09	p-trend=0.05
Small Cell Lung Cancer		
Per IQR Increase	0.93 (0.73, 1.18)	0.93 (0.74, 1.18)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	1.48 (0.25, 8.61)	0.82 (0.30, 2.22)
Tertile 2	8.79 (2.47, 31.37)	2.73 (1.33, 5.61)
Tertile 3	7.61 (1.81, 32.08)	2.06 (0.88, 4.82)
	p-trend= 0.002	p-trend=0.004
Squamous Cell Lung Cancer		
Per IQR Increase	1.12 (0.92, 1.36)	1.07 (0.89, 1.29)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.52 (0.11, 2.44)	0.67 (0.26, 1.73)
Tertile 2	1.66 (0.62, 4.44)	1.59 (0.78, 3.21)
Tertile 3	1.53 (0.47, 5.03)	1.30 (0.57, 2.97)
	p-trend= 0.35	p-trend=0.26
Lung Adenocarcinoma		
Per IQR Increase	0.93 (0.78, 1.10)	0.91 (0.77, 1.07)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.59 (0.33, 1.07)	0.62 (0.37, 1.05)
Tertile 2	1.50 (0.94, 2.42)	1.56 (1.02, 2.37)
Tertile 3	1.56 (0.79, 3.08)	1.41 (0.78, 2.54)
	p-trend=0.15	p-trend=0.10
Large Cell Lung Cancer		
Per IQR Increase	1.07 (0.88, 1.29)	1.06 (0.88, 1.27)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.92 (0.31, 2.74)	0.80 (0.36, 1.79)
Tertile 2	2.07 (0.91, 4.74)	1.57 (0.83, 2.98)
Tertile 3	3.77 (1.43, 9.91)	2.51 (1.20, 5.25)

	p-trend=0.007	p-trend=0.003
Overall UADT Cancer		
Per IQR Increase	0.96 (0.84, 1.10)	0.98 (0.86, 1.12)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.61 (0.41, 0.90)	0.64 (0.44, 0.93)
Tertile 2	0.91 (0.63, 1.34)	0.95 (0.67, 1.35)
Tertile 3	0.59 (0.33, 1.05)	0.66 (0.40, 1.10)
	p-trend= 0.13	p-trend=0.14
UADT Sqamous Cell Carcinoma		
Per IQR Increase	0.94 (0.82, 1.09)	0.96 (0.83, 1.10)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.60 (0.39, 0.92)	0.64 (0.43, 0.94)
Tertile 2	0.86 (0.57, 1.30)	0.92 (0.64, 1.33)
Tertile 3	0.56 (0.30, 1.03)	0.64 (0.37, 1.10)
	p-trend=0.09	p-trend=0.10
Esophageal Adenocarcinoma		
Per IQR Increase	1.02 (0.80, 1.30)	1.07 (0.85, 1.34)
Never Smokers	1 (Reference)	1 (Reference)
Tertile 1	0.75 (0.28, 2.01)	0.83 (0.38, 1.78)
Tertile 2	1.69 (0.76, 3.78)	1.44 (0.76, 2.72)
Tertile 3	1.22 (0.39, 3.82)	1.12 (0.49, 2.56)
	p-trend=0.45	p-trend=0.43

a. Models adjusted for age, race/ethnicity, sex, years of education, drink-years, and pack-years

Supplementary Table 1. ROC-AUC % (95% confidence limits) for cancer risk models, by smoking exposure and disease type

Disease Type	Pack-Years[a]	Cumulative Tar	Cumulative Tar and Pack-Years	CSI[b]
Overall lung cancer	76.7 (74.2, 79.1)	75.0 (72.5, 77.5) p<0.0001	76.7 (74.3, 79.2) p=0.17	77.3 (74.9, 79.8) p=0.005
Lung small cell cancer	86.9 (82.7, 91.0)	85.1 (80.9, 89.4) p=0.017	87.1 (82.9, 91.3) p=0.11	87.6 (83.4, 91.8) p=0.091
Lung squamous cell	86.5 (82.1, 90.8)	84.0 (79.4, 88.6) p=0.014	86.4 (82.0, 90.8) p=0.47	87.1 (82.7, 91.5) p=0.10
Lung large cell cancer	81.8 (77.3, 86.2)	79.7 (75.1, 84.3) p=0.022	81.8 (77.3, 86.2) p=0.45	82.5 (78.0, 87.0) p=0.078
Lung adenocarcinoma	69.4 (65.8, 73.1)	68.2 (64.6, 71.8) p=0.001	69.7 (66.0, 73.3) p=0.11	70.1 (66.4, 73.7) p=0.025
Overall UADT cancer	66.1 (63.2, 68.9)	65.2 (62.4, 68.0) p=0.011	66.1 (63.2, 68.9) p=0.90	66.6 (63.7, 69.4) p=0.021
UADT squamous cell	66.4 (63.3, 69.4)	65.4 (62.4, 68.5) p=0.012	66.4 (63.3, 69.4) p=0.98	66.9 (63.8, 69.9) p=0.031
Esophageal adenocarcinoma	69.9 (63.2, 76.7)	69.5 (62.8, 76.2) p=0.54	70.0 (63.3, 76.8) p=0.67	69.9 (63.1, 76.6) p=0.86

[a] Reference variable for the p-values

[b] Comprehensive smoking index

Supplementary Table 2. Interaction between cumulative tar exposure and race/ethnicity on cancer risk

	Adjusted OR (95% CL's) for Tar Tertile 3 vs Never Smokers[a]		
Cancer Type	**White**	**Non-White**	**Interaction Odds Ratio**
Lung	16.65 (10.22, 27.13)	5.89 (3.43, 10.10)	0.36 (0.19, 0.69) p=0.002
UADT	2.17 (1.44, 3.27)	1.95 (1.09, 3.47)	0.75 (0.40, 1.39) p=0.36

a. Models adjusted for age, sex, years of education, and drink-years

Supplementary Table 3. Interaction between cumulative tar exposure and smoking status on cancer risk

	Adjusted OR (95% CL's) for Tar Tertile 3 vs Tertile 1[a]		
Cancer Type	**Former Smokers**	**Current Smokers**	**Interaction Odds Ratio**
Lung	10.80 (6.31, 18.49)	36.72 (3.97, 339.96)	5.98 (0.66, 54.22) p=0.11
UADT	2.96 (1.76, 4.99)	6.65 (1.42, 31.02)	3.31 (0.82, 13.51) p=0.09

a. Models adjusted for age, race/ethnicity, sex, years of education, and drink-years

$$T = \sum_{i=1}^{B} (t_i * D_i * C_i) * 10^{-6}$$

Equation 1

Equation 1

50x15mm (96 x 96 DPI)

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$

Equation 2

45x5mm (96 x 96 DPI)