Cigarette Smoking and Cancer: Intensity Patterns in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study in Finnish Men

Jay H. Lubin¹, Jarmo Virtamo², Stephanie J. Weinstein³, and Demetrius Albanes³

¹ Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD.
² Department of Health Promotion and Chronic Disease Prevention, National Public Health Institute, Helsinki, Finland.
³ Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD.

Received for publication March 6, 2007; accepted for publication December 12, 2007.

Relative risks for lung and bladder cancers by smoking intensity level off at more than 15–20 cigarettes per day. A three-parameter excess relative risk model in pack-years and intensity quantified this leveling (Lubin et al., Am J Epidemiol 2007;166:479–89). Above 15–20 cigarettes per day was an “inverse exposure rate” effect whereby, for equal pack-years, the excess relative risk/pack-year decreased with increasing intensity; that is, smoking at a lower intensity for a longer duration was more deleterious than smoking at a higher intensity for a shorter duration. After adjustment for pack-years, intensity effects were quantitatively homogeneous across multiple case-control studies of lung, bladder, oral cavity, pancreas, and esophagus cancers. The authors extended those analyses to examine intensity patterns for incident bladder, esophagus, kidney, larynx, liver, lung, oropharynx, and pancreas cancers by using data from a single prospective cohort in Finland, the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study, with follow-up from enrollment, which occurred between 1985 and 1988, through April 2004. At more than 10 cigarettes per day, they found an inverse exposure rate pattern for each cancer site. After adjustment for pack-years, intensity effects were quantitatively homogeneous across the diverse cancer sites and homogeneous with intensity effects from the prior analysis of multiple studies. Consistency of intensity patterns suggested a general phenomenon and may provide clues to the molecular basis of smoking-related cancer risk.

cohort studies; Finland; models, statistical; smoking

Abbreviations: ATBC, Alpha-Tocopherol, Beta-Carotene Cancer Prevention; RR, relative risk.
alternative approach that estimates the effect of total exposure and quantifies the differential consequences associated with the manner of its delivery, namely, total exposure delivered at a low exposure rate (and long duration) or high exposure rate (and short duration). This approach uses either total exposure and exposure rate or total exposure and exposure duration. With a RR model that includes pack-years and cigarettes smoked per day, RRs for 20 and 30 cigarettes per day represent the differential effects of a lower exposure rate (and longer duration) or of a higher exposure rate (and shorter duration) on the risk with total exposure. The exposure rate modeling thus reflects an “exposure rate effectiveness factor” that describes modification of the risk with total exposure.

Investigators applied a three-parameter model for the excess RR by pack-years and smoking intensity that quantifies effects of intensity for fixed total pack-years (3). At less than 15 cigarettes per day, there was a “direct exposure rate” effect whereby, for equal total exposure, the excess RR/pack-year increased with increasing intensity. At higher intensities, there was an “inverse exposure rate” or “reduced potency” effect whereby, for equal total exposure, the excess RR/pack-year decreased with increasing intensity; that is, for fixed pack-years, smoking at a lower intensity for a longer duration was more deleterious than smoking at a higher intensity for a shorter duration. Additional analyses of cancers of the lung, bladder, oral cavity, pancreas, and esophagus found qualitatively and quantitatively comparable intensity effects across the diverse smoking-related cancers, suggesting that the intensity pattern may have reflected a general phenomenon (4). However, a reviewer raised concerns about potential confounding and case-control selection bias, since data came from multiple case-control studies from diverse source populations. To address those concerns and to extend the earlier results, the current analysis evaluates smoking patterns for incident cancers of the bladder, esophagus, kidney, larynx, liver, lung, oropharynx, and pancreas by using data from a single prospective cohort study.

Investigators administered questionnaires at baseline and at 4-month intervals during the active intervention period. Collection of covariate information ceased with the active intervention. We therefore assumed that smoking characteristics at the last visit remained unchanged during the post-intervention period.

The ATBC Study was approved by institutional review boards of participating institutions.

Models

We assumed that survival time follows a piecewise exponential distribution and applied Poisson regression methodology. For each cancer, we summarized person-years of follow-up and events in a multiway contingency table defined by attained age (seven levels; 50–54, . . . , 75–79, ≥80), year (nine levels; <1988, 1988–1989, . . . , 2000–2001, 2002–2004), pack-years (46 levels; <10, 10–11, . . . , 88–89, 90–94, 95–99, 100–109, 110–119, ≥120), cigarettes smoked per day (21 levels; <6, 6–7, . . . , 38–39, 40–44, 45–49, ≥50), time since cessation of smoking (four levels; 0, 1–2, 3–4, ≥5 years), and age at which smoking started (four levels; <15, 15–19, 20–24, ≥25 years). We also calculated person-years weighted means for the classification variables for each cell of the cross-tabulation.

We defined indicator variables 

\[ n_i, i = 1, \ldots, I, \text{ for } I \text{ intensity categories, where } n_i = 1 \text{ for intensities within category } i \text{ and zero otherwise, and modeled disease rate } r(x, d, n) \text{ as} \]

\[
r(x, d, n) = \exp(\gamma_1 x) RR(d, n) = \exp(\gamma_1 x)[1 + \sum \gamma_i n_i d],
\]

where \( d \) was total pack-years and \( n \) was number of cigarettes smoked per day. The vectors of variables \( x \) and parameters \( \gamma \), where “\( T \)” denoted vector transpose, characterized disease rates for never smokers. Within intensity category \( i \), RRs were linear in \( d \) (i.e., \( RR(n, d) = 1 + \gamma_i d \)), where \( \gamma_i \) was the excess RR/pack-year. To calculate estimates and 95 percent confidence intervals, \( \exp(\gamma_i^* \gamma) \) replaced \( \gamma_i \) to avoid range restrictions. Factoring out \( \gamma_1 \), we rewrote \( RR(d, n) \) as

\[
RR(d, n) = 1 + \gamma_1 d \sum (\gamma_i / \gamma_1) n_i.
\]

A natural extension for continuous intensity was

\[
RR(d, n) = 1 + \beta d g(n),
\]

where \( g(.) \) represented the intensity effects and \( \beta \) represented the excess RR/pack-year at \( g(n) = 1 \). We set \( g(n) = \exp(\phi_1 \ln(n) + \phi_2 \ln(n)^2) \). We also considered \( g(n) = \exp(\phi_1 \ln(n) + \phi_2 n) \) and \( g(n) = \exp(\phi_1 n + \phi_2 n^2) \); however, for most cancers, the deviance was smallest for the selected form, although deviations were generally similar, signifying comparable model fits. Fit was not improved with further inclusion of \( \ln(n) \times n, \ln(n)^2 \), or \( n^3 \).

We extended model 2 to evaluate departure from linearity by using

\[
RR(d, n) = 1 + \beta d \exp(\delta d) g(n),
\]

MATERIALS AND METHODS

Data

The Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study was a double-blind, placebo-controlled prevention trial. Between 1985 and 1988, investigators enrolled 29,133 males aged 50–69 years from 14 areas in southwestern Finland who smoked five or more cigarettes per day (5, 6). The trial ended April 30, 1993, but postintervention follow-up continued by linking the study roster to the Finnish Cancer Registry (7). For the current analysis, follow-up included time from randomization through April 30, 2004, incidence from a cancer of interest, or death, whichever occurred earliest. For pancreatic cancer, follow-up was conducted through April 30, 2003.

For cases diagnosed through April 1999, medical records were reviewed centrally by one or two study oncologists for diagnostic confirmation and staging. For cases diagnosed after April 1999, the Finnish Cancer Registry information was used for site and diagnosis date.
TABLE 1. Parameter estimates and tests of hypothesis from fitting model 2 to data on continuing cigarette smokers and on those smoking more than 10 cigarettes per day.

| Cancer site | RRsmk \* | y | \* | z | b/10 | p | z | p | \* | c | cigarettes per day. | Relative risk (RR) for male ever smokers compared with never smokers. Let y be the age- and year-specific Finnish cancer incidence rate, and I{ever smoker}. Then, I = (1 \* P\*I + \*I \* I) + P\*I \* I. RRsmk is based on information from the US Department of Health and Human Services (8) and Schottenfeld and Fraumeni (9). RRsmk values ranged from 2.0 for kidney, liver, and pancreas cancers to 12.0 for lung cancer (table 1). For males born between 1916 and 1940, proportions of ever smokers in 1978 through 2001 ranged from 0.65 to 0.80 (10). For the adjustment, we assumed the proportion 0.70 of males who ever smoked. All models included the logarithm of the age- and year-specific disease rate for never smokers as a fixed offset variable (x) to represent rates for never smokers.

RESULTS

Since administration of questionnaires ended in 1993, we had no information on smoking cessation or on former smokers who relapsed. Therefore restricted our analysis to smokers who never attempted to quit during the active follow-up period to minimize the possibility of changes in smoking behavior. We also excluded cigar or pipe smokers. Among continuing smokers, there were 317,154 person-years of follow-up, although totals varied slightly with outcome, and 403, 80, 241, 116, 170, 2,248, 168, and 244 cancers of the bladder, esophagus, kidney, larynx, liver, biliary tract, lung, oropharynx, and pancreas, respectively. Person-years weighted mean number of cigarettes per day, years of smoking, and pack-years were 20.8, 42.8, and 44.7, respectively, with higher means for cases (table 1). For each cancer and with never smoked as the referent, RRs increased with pack-years within each category of intensity (not shown), as observed previously (3, 4). For each cancer, we fitted model 1, which resulted in 71 finite estimates of excess RR/pack-year. The null hypothesis \( \delta = 0 \) reflected no departure from linearity. A similar extension within each category was used for model 1.

The ATBC Study included only smokers. To derive disease rates for male never smokers, we obtained male age-specific cancer incidence rates for 1980–1984, 1985–1989, 1990–1994, and 1995–1999 from the NORDCAN version 2.2 computer program from the Association of Nordic Cancer Registries (available at http://www.anncr.nordcan.org) and for 2000–2004 from the Finnish Cancer Registry (available at http://www.cancerregistry.fi/stats/eng/veng00090.html). We obtained approximate estimates of rates for never smokers by multiplying each age and year-specific rate by 1.0/(1 \* P\*ever smoked) + P\*ever smoked) \* RRsmk, where RRsmk is the RR by ever smoked compared with never smoked (table 1), which we summarized from information from the US Department of Health and Human Services (8) and Schottenfeld and Fraumeni (9). RRsmk values ranged from 2.0 for kidney, liver, and pancreas cancers to 12.0 for lung cancer (table 1). For males born between 1916 and 1940, proportions of ever smokers in 1978 through 2001 ranged from 0.65 to 0.80 (10). For the adjustment, we assumed the proportion 0.70 of males who ever smoked. All models included the logarithm of the age- and year-specific disease rate for never smokers as a fixed offset variable (x) to represent rates for never smokers.
homogeneity of excess RR/pack-year estimates with intensity was rejected at the 0.10 level; that is, the excess RR/pack-year varied with intensity (table 1). Excess RR/pack-year estimates increased with intensity at low intensities for the larynx, esophagus, and kidney but not for the other sites. The 95 percent confidence intervals suggested low power for estimating intensity effects below 10 cigarettes per day, because of few cases, which numbered 27, 4, 11, 1, 9, 107, 10, and 21 for cancers of the bladder, esophagus, kidney, larynx, liver/biliary tract, lung, oropharynx, and pancreas, respectively.

For data above 10 cigarettes per day, we computed summary estimates of the intensity parameters as a weighted mean of cancer site-specific estimates by using inverse covariance matrices as weights to evaluate consistency of intensity patterns. The summary estimates were $\phi_1 = 0.569$ and $\phi_2 = -0.150$. With $\phi_1$ and $\phi_2$ fixed at these values, we reestimated $\beta$ and calculated a 2 degrees of freedom likelihood ratio test of homogeneity of site-specific intensity estimates with the summary values (table 1, column 12). After adjustment for pack-years, intensity effects for each cancer were homogeneous with the summary values. A previous analysis of multiple case-control studies observed homogeneity of intensity effects with summary estimates of $\phi_1 = 2.72$ and $\phi_2 = -0.479$ (4). Model 2 with intensity parameters fixed at these values was also consistent with the current data for each cancer site, except oropharynx (table 1, column 13); the fitted models corresponded closely to the excess RR/pack-year patterns from the ATBC Study (figure 1, dashed line).

**DISCUSSION**

Previous analyses of lung cancer identified a linear RR relation in pack-years with effect modification by smoking intensity (3, 11). There was a direct exposure rate effect below 15–20 cigarettes per day and an inverse exposure rate effect at higher intensities. Analysis of five lung cancer studies, two bladder cancer studies, and studies of cancers of the oral cavity, pancreas, and esophagus identified qualitatively and quantitatively similar intensity patterns (4). All were case-control studies, and therefore results were potentially affected by intensity-related, differential “willingness to participate” (12). Using data from the ATBC Study, which was not subject to participation bias, we found similar
patterns of a decreasing excess RR/pack-year with increasing intensity above 10–15 cigarettes per day. Moreover, the inverse exposure rate pattern was homogeneous across the ATBC Study cancer sites and was qualitatively and quantitatively consistent with previous smoking intensity patterns (4).

Below 10–15 cigarettes per day, we found no consistent intensity pattern for all sites (figure 1), in particular, the direct exposure rate effect observed previously for lung cancer (3). Ranges for total pack-years were limited at low smoking intensities. The respective 75th percentile and maximum pack-years were 45.5 and 162 in the full cohort and 13.1 and 23.9 for smokers of fewer than 10 cigarettes per day. The limited range for pack-years resulted in substantial uncertainty in estimating smoking effects at low intensities. For the ATBC Study, this issue was exacerbated by the enrollment criterion that included only smokers of five or more cigarettes per day. Toxicologically, it was unclear whether the limiting value for the estimated excess RR/pack-year with decreasing mean smoking intensity, that is, one cigarette per day, one per week, one per month, and so forth, should also tend toward zero. Estimation at low intensities was potentially further burdened by misclassification of smoking intensity.

Because of space limitations in this paper, we did not present analyses for effect modification of model 2 by attained age and age at which smoking started. Since information on smoking status did not extend beyond the active follow-up period in 1993, evaluation of cessation of smoking was not feasible. A previous analysis of lung cancer showed a significant interaction of attained age and smoking intensity, but not an interaction of age and total pack-years, while variations in the excess RR with age at which smoking started were the result of significant interactions for both age at first smoking and intensity and age at first smoking and pack-years (11). For only bladder, lung, and oropharynx cancers were number of cases and smoking risk sufficient to evaluate age and age at which smoking first started. As in the previous lung cancer analysis, the inverse intensity effects exhibited less variation for ages less than 50 years; differences were statistically significant for lung and oropharynx cancers but not for bladder cancer. In addition, intensity showed less variation for ages of less than 20 years at pharynx cancers but not for bladder cancer. In addition, in the inverse intensity pattern for all sites (figure 1), in particular, the direct exposure rate effect observed previously for lung cancer (3). Ranges for total pack-years were limited at low smoking intensities. The respective 75th percentile and maximum pack-years were 45.5 and 162 in the full cohort and 13.1 and 23.9 for smokers of fewer than 10 cigarettes per day. The limited range for pack-years resulted in substantial uncertainty in estimating smoking effects at low intensities. For the ATBC Study, this issue was exacerbated by the enrollment criterion that included only smokers of five or more cigarettes per day. Toxicologically, it was unclear whether the limiting value for the estimated excess RR/pack-year with decreasing mean smoking intensity, that is, one cigarette per day, one per week, one per month, and so forth, should also tend toward zero. Estimation at low intensities was potentially further burdened by misclassification of smoking intensity.

Because of space limitations in this paper, we did not present analyses for effect modification of model 2 by attained age and age at which smoking started. Since information on smoking status did not extend beyond the active follow-up period in 1993, evaluation of cessation of smoking was not feasible. A previous analysis of lung cancer showed a significant interaction of attained age and smoking intensity, but not an interaction of age and total pack-years, while variations in the excess RR with age at which smoking started were the result of significant interactions for both age at first smoking and intensity and age at first smoking and pack-years (11). For only bladder, lung, and oropharynx cancers were number of cases and smoking risk sufficient to evaluate age and age at which smoking first started. As in the previous lung cancer analysis, the inverse intensity effects exhibited less variation for ages less than 50 years; differences were statistically significant for lung and oropharynx cancers but not for bladder cancer. In addition, intensity showed less variation for ages of less than 20 years at which smoking first started.

The inverse exposure rate patterns may have reflected intensity-dependent molecular mechanisms, such as enhanced DNA repair (13–16), saturation of activation pathways (17–19), or increased induction of detoxification enzymes (20). Previous application of the current models to bladder cancer case-control data showed that variations in smoking risk with N-acetyltransferase 2 (the NAT2 gene) status resulted from interactions with smoking intensity and not total pack-years of exposure, and that the relative increase in smoking risk for NAT2 slow acetylators, compared with rapid/intermediate acetylators, increased with smoking intensity (21).

Our results may have also derived from behavioral factors, such as high-intensity smokers inhaling less deeply while maintaining addiction-sufficient nicotine levels, resulting in decreasing carcinogenic yield per cigarette and decreasing risks at higher intensities. A previous analysis found no evidence that frequency or depth of inhalation was related to smoking intensity after adjusting for total pack-years (3). However, a study of 190 smokers observed increased cotinine and nicotine levels with increased intensity and a marginally significant \( p = 0.08 \) decline in “nicotine boost,” that is, an increase in blood plasma nicotine per cigarette (22). Using lung cancer case-control data, a sensitivity analysis based on the relation between cotinine, a measure of “internal” smoking intensity, and cigarettes per day evaluated the potential impact of declining carcinogenic yield per cigarette (4). Investigators concluded that overestimation of internal exposure rate by cigarettes per day may have contributed to the inverse exposure rate pattern but was unlikely to fully explain results.

Because the ATBC Study included smokers only, we estimated cancer rates for never smokers by adjusting Finnish male population cancer rates using the proportion of ever smokers (0.70) and RRsmk, the RR of ever smokers compared with never smokers for each cancer site. We evaluated the impact of these choices with a sensitivity analysis by using a variety of values for RRsmk and the proportion of ever smokers. Increasing or decreasing RRsmk induced smaller or larger disease rates for never smokers, respectively, which resulted in an increase or decrease in the site-specific excess RR/pack-year curves. With RRsmk doubled or halved (e.g., for lung cancer RRsmk set to 24.0 or 6.0, respectively), the general patterns of smoking intensity effects as defined by \( \phi_1 \) and \( \phi_2 \) were not markedly altered, although estimates of \( \beta \) changed. We also allowed RRsmk to vary smoothly with age (e.g., for lung cancer RRsmk set to 12.0 at age 50 years, and increasing or decreasing smoothly to 24.0 or 6.0 at age 70 years); estimates of \( \phi_1 \) and \( \phi_2 \) changed only slightly and did not significantly differ from the estimates shown in table 1. This minimal change was likely due to the relatively small correlation between age and pack-years (0.18) and age and number of cigarettes smoked per day (−0.15). We also varied the proportion of smokers from 0.60 to 0.80, with only minimal changes in parameter estimates.

Using data from the ATBC Study and adjusting for total pack-years, we found an inverse exposure rate pattern for intensities above 10–15 cigarettes per day that was qualitatively and quantitatively consistent across eight smoking-related cancers and consistent with previous results from diverse study populations. These patterns likely derived from both intensity-dependent molecular mechanisms and nicotine-related smoking behavioral factors.

ACKNOWLEDGMENTS

This research was supported by the Intramural Research Program of the National Institutes of Health and the National Cancer Institute and by US Public Health Service contracts N01-CN-45165, N01-RC-45035, and N01-RC-37004 from the National Cancer Institute, Department of Health and Human Services.

Conflict of interest: none declared.
REFERENCES


