Practice of Epidemiology

Cigarette Smoking and Cancer Risk: Modeling Total Exposure and Intensity

Jay H. Lubin[1], Michael C. R. Alavanja[2], Neil Caporaso[3], Linda M. Brown[1], Ross C. Brownson[4], R. William Field[5,6], Montserrat Garcia-Closas[7], Patricia Hartge[8], Michael Hauptmann[1], Richard B. Hayes[2], Ruth Kleinerman[9], Manolis Kogevinas[10,11], Daniel Krewski[12], Bryan Langholz[13], Ernest G. Létourneau[14], Charles F. Lynch[5], Núria Malats[10], Dale P. Sandler[15], Angelika Schaffrath-Rosario[16], Janet B. Schoenberg[17], Debra T. Silverman[2], Zuoyuan Wang[18], H.-Erich Wichmann[19], Homer B. Wilcox[17], and Jan M. Zielinski[20,21]

[1] Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[2] Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[3] Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[4] Department of Community Health, School of Public Health, St. Louis University, St. Louis, MO.
[5] Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, IA.
[6] Department of Occupational and Environmental Health, College of Public Health, University of Iowa, Iowa City, IA.
[7] Hormonal and Reproductive Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[8] Epidemiology and Biostatistics Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[9] Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.
[10] Centre for Research in Environmental Epidemiology, Municipal Institute of Medical Research (IMIM), Barcelona, Spain.
[11] Department of Social Medicine, Medical School, University of Crete, Heraklion, Greece.
[12] McLaughlin Centre for Population Health Risk Assessment, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada.
[13] Department of Preventive Medicine, School of Medicine, University of Southern California, Los Angeles, CA.
[14] Radiation Protection Bureau, Health Protection Branch, Health Canada, Ottawa, Ontario, Canada.
[15] Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC.
[16] Robert Koch-Institut, Abteilung Epidemiologie und Gesundheitsberichterstattung, Berlin, Germany.
[17] Cancer Epidemiology Services, New Jersey Department of Health and Senior Services, Trenton, NJ.
[18] Laboratory of Industrial Hygiene, Ministry of Health, Beijing, China.
[19] Institute of Epidemiology, National Research Center for Environment and Health (GSF), Neuherberg, Germany.
[20] Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario, Canada.
[21] Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada.

A recent analysis showed that the excess odds ratio (EOR) for lung cancer due to smoking can be modeled by a function which is linear in total pack-years and exponential in the logarithm of smoking intensity and its square. Below 15–20 cigarettes per day, the EOR/pack-year increased with intensity (direct exposure rate or enhanced potency effect), suggesting greater risk for a total exposure delivered at higher intensity (for a shorter duration) than for an equivalent exposure delivered at lower intensity. Above 20 cigarettes per day, the EOR/pack-year decreased with increasing intensity (inverse exposure rate or reduced potency effect), suggesting greater risk for a total exposure

Correspondence to Dr. Jay H. Lubin, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd., Rockville, MD 20852 (e-mail: lubinj@mail.nih.gov).

delivered at lower intensity (for a longer duration) than for an equivalent exposure delivered at higher intensity. The authors applied this model to data from 10 case-control studies of cancer, including cancers of the lung, bladder, oral cavity, pancreas, and esophagus. At lower intensities, there was enhanced potency for several cancer sites, but narrow ranges for pack-years increased uncertainty, precluding definitive conclusions. At higher intensities, there was a consistent reduced potency effect across studies. The intensity effects were statistically homogeneous, indicating that after accounting for risk from total pack-years, intensity patterns were comparable across the diverse cancer sites.

case-control studies; epidemiologic methods; models, statistical; neoplasms; risk; smoking

In studies of lung cancer and bladder cancer, odds ratios increase with smoking intensity, but often level off and even decline at high intensities (1, 2). This pattern suggests a changing impact of intensity on disease risk, although its precise implication is uncertain. Behavioral factors, such as high-intensity smokers' inhaling less deeply, or exposure-dependent biases are possible explanations for the leveling off of risk, but the involvement of intensity-dependent molecular mechanisms in carcinogenesis is another plausible explanation. To better understand the biologic basis of host reaction to the many chemical compounds in tobacco smoke, it is important to clarify the effects of total cigarette smoke exposure, smoking intensity, and smoking duration. In particular, we are interested in evaluating how the rate of delivery of cigarette smoke influences disease risk at a fixed total exposure.

Standard approaches for exploring the leveling off of odds ratios are problematic, since they depend on models for odds ratios that incorporate smoking duration and intensity; this can lead to difficulty in interpreting intensity effects due to changing total exposures. For example, in a logistic model, the intensity parameter represents the change in the natural logarithm of the odds ratio (OR) per unit of intensity (i.e., ln(OR) per cigarette per day) at a fixed duration. The duration parameter has a similar representation but at a fixed intensity. Therefore, odds ratios at two different intensities reflect not only the different intensities but also different total pack-years, since duration is fixed. For example, at 30 years' duration, differences in odds ratios for 20 and 30 cigarettes per day embed differences in total exposure, that is, 30 and 45 pack-years, respectively. Thus, the intensity parameter does not represent a "pure" intensity effect but includes an additional effect of total pack-years.

In the European Smoking and Health Study (ESHS), Lubin et al. (3, 4) recently described a three-parameter model for the excess odds ratio (EOR) for lung cancer which was linear in pack-years and exponential in the logarithm of smoking intensity and its square. The model defined the impact of intensity on the EOR/pack-year, thereby isolating the contribution of intensity at a fixed total exposure. Under 15–20 cigarettes per day, the EOR/pack-year increased with intensity (direct exposure rate or an enhanced potency effect); that is, total exposure imparted at a higher intensity for a shorter duration induced greater risk than an equal total exposure imparted at a lower intensity for a longer duration.

At higher intensities, the EOR/pack-year decreased with increasing intensity (inverse exposure rate, or a reduced potency or "wasted" exposure effect); that is, total exposure imparted at a lower intensity for a longer duration induced greater risk than an equal exposure imparted at a higher intensity for a shorter duration.

In the current analysis, we applied the model to data from additional lung cancer studies and to studies of other smoking-related cancers, including cancers of the bladder, oral cavity, pancreas, and esophagus, and evaluated smoking intensity patterns across the diverse cancer sites.

MATERIALS AND METHODS

Models

For disease outcome C, where $C = 1$ denotes a case and $C = 0$ denotes a noncase, we fit a model for the odds of disease of the form

$$P[C=1\,|\,\mathbf{x},d,n]/P[C=0\,|\,\mathbf{x},d,n] = \exp(\alpha^T \mathbf{x})\mathrm{OR}(d,n),$$

where x is a vector of adjustment variables, T denotes vector transpose, and $\mathrm{OR}(d,n)$ is the odds ratio model for d total pack-years and n cigarettes per day. For I intensity categories, we define indicator variables n_i, $i = 1, \ldots, I$, where $n_i = 1$ for intensities within category i and zero otherwise. We fit

$$OR(d,n) = 1 + \Sigma\, \gamma_i\, n_i\, d. \qquad (1)$$

Within category i, odds ratios are linear in d (i.e., $\mathrm{OR} = 1 + \gamma_i\, d$), where γ_i is the EOR/pack-year. For calculating estimates and 95 percent confidence intervals, we replace γ_i with $\exp(\gamma_i^*)$ to avoid range restrictions. Factoring out γ_1, model 1 (equation 1) becomes

$$\mathrm{OR}(d,n) = 1 + \gamma_1\, d\, \Sigma(\gamma_i/\gamma_1)n_i.$$

Note that (γ_i/γ_1) represents the relative variation of the linear relation with categories of intensity. A natural extension for continuous intensity is

$$\mathrm{OR}(d,n) = 1 + \beta\, d\, g(n), \qquad (2)$$

where β represents the EOR/pack-year at $g(n) = 1$, while $g(\cdot)$ defines the variation of the linear EOR/pack-year parameter with intensity. Model 2 (equation 2) is a simple EOR model in which the effects of pack-years and intensity

multiply. We considered three forms for $g(n)$, including: $g(n) = \exp\{\varphi_1 \ln(n) + \varphi_2 \ln(n)^2\}$; $g(n) = \exp\{\varphi_1 \ln(n) + \varphi_2 n\}$; and $g(n) = \exp\{\varphi_1 n + \varphi_2 n^2\}$, with $g(n) = \exp\{\varphi_1 \ln(n) + \varphi_2 \ln(n)^2\}$ providing the best fit. Fit was not substantially improved with further inclusion of $\ln(n) \times n$, $\ln(n)^3$, or n^3.

We evaluate departure from linearity by including an exponential factor in model 2,

$$\mathrm{OR}(d,n) = 1 + \beta\, d \exp(\delta d)\, g(n), \qquad (3)$$

where δ measures concavity ($\delta < 0$) or convexity ($\delta > 0$) in the EORs for pack-years. A test of the null hypothesis $\delta = 0$ is a test of no departure from linearity.

Case-control studies of smoking and cancer

We conducted 10 individual analyses using data from nine case-control studies and one set of pooled data from multiple studies (table 1). Data sets included those from the ESHS (5, 6), the German Radon Study (7–9), the North American Radon Pooling Project (10), the Gansu Lung Cancer Study (11), the Colorado Plateau Uranium Miners Study (12, 13), the National Bladder Cancer Study (14), the Spanish Bladder Cancer Study (15), the Puerto Rico Oral Cancer Study (16), and the Investigation of Tumors that Occur Excessively among Blacks, which we designated the Population Health Study (17, 18). We included these studies because the cancers were smoking-related and data sets were readily available. For consistency, we established common inclusion criteria, limiting subjects to ages 50–74 years at enrollment and to never smokers, current smokers, or former smokers who had stopped smoking within 5 years of enrollment, except where noted. The age restriction reduced the impact of any genetic predisposition in younger cases or diagnostic ambiguity in elderly cases. The restriction to never, current, and recent former smokers eliminated any effect modification by time since cessation of smoking. We excluded subjects who had started smoking after age 40 years and subjects who smoked pipes or cigars, except where noted.

For the German Radon Study (7–9), we included 75-year-old subjects, since this was the upper age limit in that study, and we excluded subjects who paused their smoking for more than 5 years and males who smoked 1–8 cigarettes per day. The latter exclusion maintained consistency with prior analyses and did not affect overall inference (4). The North American Radon Pooling Project (10) was a pooled analysis of data from studies of lung cancer and exposure to radioactive radon gas in homes (19–24). The Missouri-I study (19) enrolled never smokers and former smokers who quit smoking 10 years or more prior to enrollment, and thus did not qualify for the current analysis. With standard restrictions, numbers of cases for the Iowa (20), Missouri-II (21), New Jersey (22), Connecticut (23), Utah/South Idaho (23), and Winnipeg (24) studies totaled 238, 386, 323, 420, 261, and 499, respectively, and numbers of controls totaled 362, 340, 281, 396, 391, and 362. The Missouri-II, Connecticut, and Utah/South Idaho studies employed randomized recruitment designs. Therefore, we included fixed offsets to enable unbiased estimation of smoking effects (25, 26).

For the Gansu Lung Cancer Study (11), we excluded females, since only 10 percent of them smoked and those who did smoke averaged only five cigarettes per day. Standard restrictions eliminated the 35 percent of men who smoked only Chinese long-stemmed pipes. In order to retain the 50 percent of smokers who smoked both cigarettes and long-stemmed pipes, we calculated a cigarette-equivalent variable using observed odds ratios. We did this by summing 1) cigarettes per day, 2) three times the liang (50 g) of tobacco used per month in hand-rolled cigarettes, and 3) 0.65 times the liang used per month in long-stemmed pipes. The Colorado Plateau Uranium Miners Study (12, 13) was a case-control study nested within an underground miners' cohort. Smoking information came from four questionnaires administered between 1950 and 1960 and at other times through 1969. We assumed that smoking status at the most recent assessment was unchanged to the end of follow-up. The Population Health Study (17, 18) enrolled persons with four different types of cancers (pancreas, esophagus, and prostate cancer and multiple myeloma); however, we analyzed only cases with pancreatic and esophageal cancer and their controls. The latter site included only males. For the Population Health Study data, we extended the smoking cessation restriction to 10 years in order to increase sample size.

In all models, results were stratified by sex, study center, or location and were adjusted for age category and important study-specific factors (table 1). Models were fitted using the binary outcome module in the Epicure set of programs (27).

RESULTS

Table 1 describes the studies analyzed and their selected smoking characteristics.

Pack-years and smoking intensity

Modeling lung cancer in the ESHS. We illustrate the approach by expanding on the previous ESHS analysis (3). Since we were primarily interested in patterns rather than specific odds ratios, we computed odds ratios for the cross-classification of pack-years and intensity relative to never smoking, combining categories when data were sparse. Odds ratios for lung cancer by pack-years of smoking, relative to never smoking, were consistent with linearity within 12 intensity categories (figure 1). Homogeneity of the γ_i estimates was rejected ($p < 0.001$). Since patterns were linear, the EOR/pack-year (γ_i) completely characterized odds ratios by pack-years within each intensity category.

Figure 2 (ESHS panel) shows the plot of the EOR/pack-year estimates by intensity. Estimates increased with intensities up to 15–20 cigarettes per day and then decreased, and closely conformed to the fitted model 2 based on continuous smoking variables (solid line).

Modeling cancer outcomes. For each data set, odds ratios by pack-years were generally consistent with linearity within intensity categories. Across data sets, 105 intensity categories had finite γ_i estimates. Eleven tests rejected

TABLE 1. Study design, summary statistics,* and factors used for adjustment in case-control studies of the relation between cigarette smoking and cancer

Study (ref. no.)	Cancer site	Study design	Location of study	Dates of study	No. of cases	No. of controls	Adjustment factors
European Smoking and Health Study (6)	Lung	Hospital-based	Seven European study centers	1976–1980	7,804	15,207	Age, sex, and study center
German Radon Study (7)	Lung	Population-based	23 regions in Germany	1990–1997	4,071	4,628	Age, sex, and region
North American Radon Pooling Project (10)	Lung	Pooled data from seven population-based radon studies	United States and Canada	1982–1996	3,662	4,966	Age, sex, study, and radon
Gansu Lung Cancer Study (11)	Lung	Population-based	Two prefectures in Gansu Province, China	1994–1998	563 males	1,232 males	Age, prefecture, radon, and socioeconomic factors
Colorado Plateau Uranium Miners Study (12)	Lung	Nested case-control study from a cohort of underground miners	Colorado Plateau, United States	Enrolled in 1950–1960, with follow-up through 1990	263	10,322	Age, year, and radon
National Bladder Cancer Study (14)	Bladder	Population-based	10 areas in the United States	Enrollment for 1 year, starting between December 1977 and March 1978 (depending on area)	2,982	5,782	Age, sex, study center, race, and family history of urinary tract cancer
Spanish Bladder Cancer Study (15)	Bladder	Hospital-based	Five areas in Spain	1998–2001	1,090	1,264	Age, sex, region, fruit and vegetable intake, and employment in high-risk occupations
Puerto Rico Oral Cancer Study (16)	Oral cavity	Population-based	Puerto Rico	1992–1995	341	521	Age, sex, region, and total alcohol intake
Population Health Study (18)	Pancreas	Population-based	Three areas in the United States	1986–1989	481	2,135	Age, sex, study center, race, and alcohol intake
Population Health Study (17)	Esophagus	Population-based	Three areas in the United States	1986–1989	581 males	1,357 males	Age, sex, study center, race, and alcohol intake

Table continues

linearity at the $\alpha = 0.05$ level, with seven departures being convex and four concave. Adjusting for continuous intensity, there were no significant departures from linearity (table 2). EOR/pack-year estimates varied substantially within and across studies (figures 2 and 3). At higher intensities, EOR/pack-year estimates generally decreased with increasing intensity, except for the North American Radon Pooling Project, the Gansu Lung Cancer Study, and the Population Health Study (pancreas), where estimates were constant with intensity. The solid lines in figures 2

and 3 show fitted data from model 2 (parameter estimates in table 2). Smoking intensity significantly modified the EOR/pack-year at the 0.10 level in the ESHS, the German Radon Study, the North American Radon Pooling Project, the Colorado Plateau Uranium Miners Study, the National Bladder Cancer Study, the Spanish Bladder Cancer Study, and the Population Health Study (esophagus).

We considered several studies in greater detail. For the North American Radon Pooling Project, the fitted EOR function increased at low and high intensities. However,

TABLE 1.   Continued

Included in analysis				Smoking characteristics					
No. of cases		No. of controls		Ever smoking (%)		Mean duration† (years)		Mean intensity† (cigarettes/day)	
Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
3,991	634	6,660	1,224	80.2	37.1	42.0	38.9	21.7	16.2
1,726	460	1,424	645	73.0	33.6	41.4	37.4	19.9	15.1
610	1,517	540	1,592	84.3	71.1	45.4	43.1	24.4	21.2
266		438		87.6		34.4		16.1	
190		7,420		77.0		41.5		18.6	
990	373	1,798	776	63.6	37.2	44.7	39.2	25.8	19.0
449	52	467	69	76.6	7.4	46.1	28.7	24.2	18.8
168	35	181	47	69.6	34.1	45.2	44.5	28.8	23.9
135	166	689	469	60.8	42.2	42.2	39.9	23.1	17.9
331		601		76.4		42.9		23.3	

* Except where noted, the analytic data included subjects aged 50–74 years (75 years for the German Radon Study) who had never smoked or who started smoking before age 40 years and either currently smoked or had stopped smoking 5 or fewer years before enrollment (10 years for the Population Health Study groups). The data excluded persons who smoked pipes only, cigars only, or a mixture of pipes and cigars.

† Among current smokers and recent former smokers.

the fitted model underpredicted EOR/pack-year between 15 and 30 cigarettes per day and overpredicted EOR/pack-year above 30 cigarettes per day, suggesting that lower-intensity data were distorting the model. At low intensities, there were relatively few smokers, but they were highly influential. For 289 smokers (125 cases and 164 controls) who smoked fewer than 10 cigarettes per day, the median number of pack-years was 14.7 and the interquartile range was 10.6–19.6 pack-years, in comparison with 45 pack-years and 32.8–61.1 pack-years, respectively, for all North American

Radon Pooling Project smokers. After omitting smokers who smoked fewer than 10 cigarettes per day, estimates were $\varphi_1 = 1.91$ and $\varphi_2 = -0.27$, which were similar to estimates in other data sets (table 2). The EOR function decreased at higher intensities and conformed closely to EOR/pack-year estimates (figure 2, NARnP panel, dashed-dotted-dotted line).

Fitted EOR curves for the Population Health Study (pancreas) and Population Health Study (esophagus) increased sharply at low intensities. After omitting smokers who
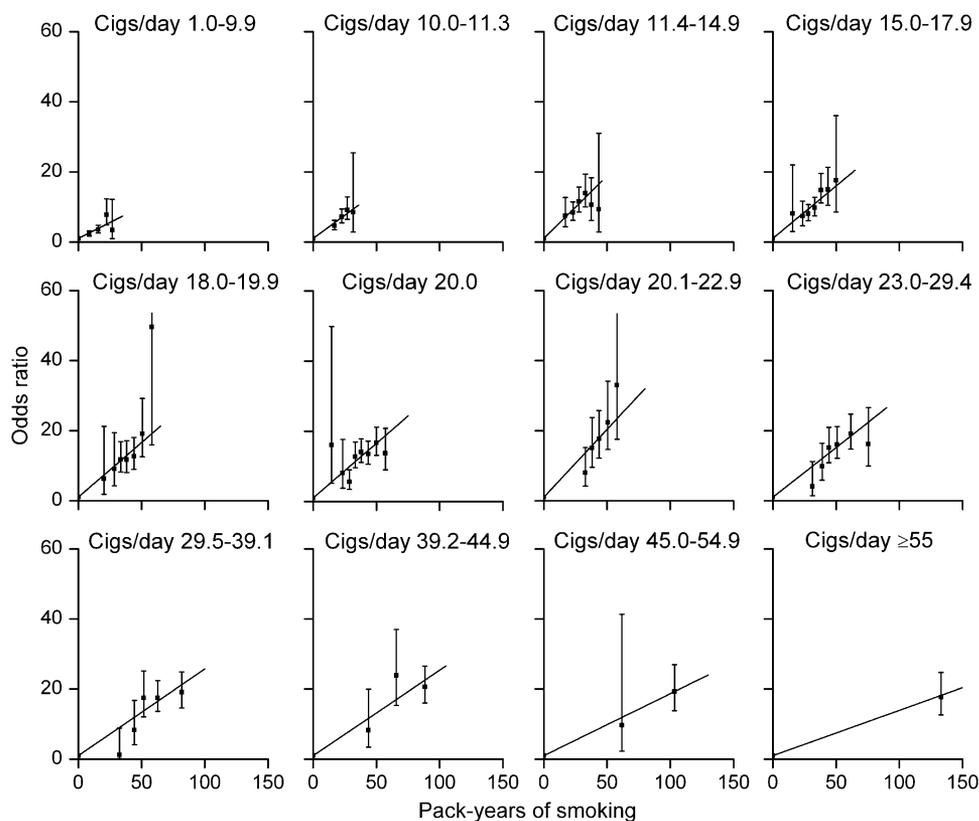
FIGURE 1. Odds ratios for lung cancer according to pack-years of cigarette smoking (black squares) and fitted linear excess odds ratio (solid line) within categories of number of cigarettes smoked per day (Cigs/day), European Smoking and Health Study, 1976–1980. All odds ratios were calculated relative to never smokers and are plotted at the mean pack-years within each category. Bars, 95% confidence interval.

smoked fewer than five cigarettes per day (four cases and 23 controls in the Population Health Study (pancreas) and seven cases and 14 controls in the Population Health Study (esophagus)), the EOR functions were similar to those of other studies (table 2; figure 3, dotted-dashed lines).

Estimates of β varied widely, while estimates of φ$_1$ and φ$_2$ appeared similar across data sets (table 2). This suggests a general commonality of intensity patterns, with differences in the fitted curves in figures 2 and 3 being due primarily to differences in β. To examine this more formally, we computed summary estimates across studies of φ$_1$ = 2.72 and φ$_2$ = −0.48 using a random-effects approach (28, 29). Using model 2, we estimated β for each data set with φ$_1$ and φ$_2$ fixed at the summary values (figures 2 and 3, dashed lines). Table 2 displays *p* values for the 2-df test of the null hypothesis that study-specific intensity parameters equaled the summary estimates. Results indicated statistical consistency of intensity parameters for all studies except the North American Radon Pooling Project, the National Bladder Cancer Study, the Spanish Bladder Cancer Study, and the Population Health Study (esophagus). For the North American Radon Pooling Project, homogeneity was rejected after omission of smokers who smoked fewer than 10 cigarettes per day, although intensity esti-

mates were similar to summary estimates. For the National Bladder Cancer Study and the Spanish Bladder Cancer Study, differences between study-specific and summary intensity parameters were due to initiation of smoking at age 12 years or younger (*p* = 0.19 and *p* = 0.29, respectively, in restricted data), and for the Population Health Study (esophagus), differences were due to smoking of very low intensity (*p* = 0.29, respectively).

Variations in smoking intensity effects by study center

Except for the Colorado Plateau Uranium Miners Study, each data set included multiple centers or independent studies. Assuming that centers or component studies represented independent replicates, there were significant variations in β estimates only for ESHS centers (*p* < 0.001) and for the contributing studies of the North American Radon Pooling Project (*p* = 0.02). In contrast, intensity patterns (i.e., φ$_1$ and φ$_2$ estimates) were homogeneous across centers within each study, except for the ESHS. However, intensity effects in the ESHS were consistent with homogeneity (*p* = 0.12) after additional adjustment for type of cigarette smoked (filter, nonfilter, or mixed types).
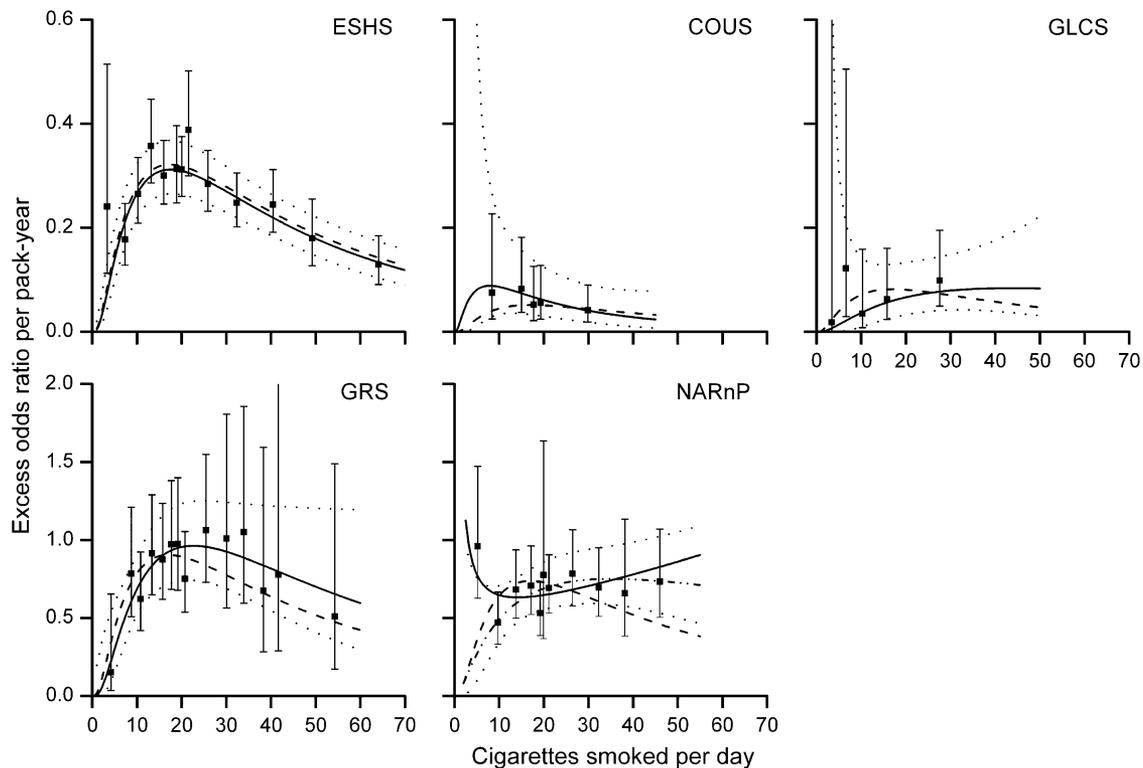
FIGURE 2.   Estimates of the excess odds ratio for lung cancer per pack-year of cigarette smoking in the European Smoking and Health Study (ESHS) (1976–1980), the Colorado Plateau Uranium Miners Study (COUS) (1960–1990), the Gansu Lung Cancer Study (GLCS) (1994–1998), the German Radon Study (GRS) (1990–1997), and the North American Radon Pooling Project (NARnP), a combined data set from six studies (1982–1996). The figure shows results based on a linear odds ratio model within categories of cigarettes smoked per day (black squares; bars, 95% confidence interval) and the estimated model 2 (equation 2 in text) fitted to data from each study (solid line; dotted lines, pointwise 95% confidence interval). The figure also shows results based on model 2 with the pack-years parameter (β) estimated and the intensity parameters fixed at summary values for all studies combined ($\varphi_1 = 2.72$ and $\varphi_2 = -0.479$) (dashed line). Model 2 was also fitted to NARnP data that omitted smokers who smoked fewer than 10 cigarettes per day (dashed-dotted-dotted line).

DISCUSSION

Previous analysis of the ESHS identified a linear relation for pack-years of smoking and lung cancer risk, with a distinct pattern for the effects of smoking intensity (3). Our analysis of additional studies of lung cancer and cancers of the bladder, oral cavity, pancreas, and esophagus suggests that this relation is broadly consistent across diverse cancer sites. At lower intensities, variations in the EOR/pack-year were consistent with an increasing intensity effect (direct exposure rate or an enhanced potency effect); however, results are subject to uncertainty because of the relatively few case subjects and the increased variability due to limited ranges for pack-years. At higher intensities, variations in the EOR/pack-year with intensity were consistent with a decreasing intensity effect (inverse exposure rate, or a reduced potency or wasted exposure effect), with smoking at a lower intensity for a longer duration being more harmful than smoking at a higher intensity for a shorter duration.

We also found that intensity effects were generally consistent in magnitude across diverse smoking-related cancer sites. It may at first seem counterintuitive that effects of smoking intensity could be similar in pattern and magnitude across diverse cancer sites. However, the overall levels of risk from smoking are reflected in the diverse EOR/pack-year estimates, while intensity effects reflect only the modulation of site-specific disease risk.

Cigarette smoking is most strongly associated with lung cancer, and EOR curves for the ESHS, the German Radon Study, and the North American Radon Pooling Project generally lay above curves for other cancer sites (figure 2). The two- to threefold greater EORs/pack-year in the German Radon Study and the North American Radon Pooling Project, as compared with the ESHS, were unexpected. Reasons for the greater EOR/pack-year estimates are unknown. The ESHS was hospital-based, and smokers could have been overrepresented among control patients, although restricting control diseases did not change the results. Risk estimates for the Gansu Lung Cancer Study data were low and reflect the lower odds ratios for lung cancer and smoking found in many studies carried out in China (30–34), as compared with studies in Western countries. The lower smoking risk in the Colorado Plateau Uranium Miners Study has been noted previously and may be due to different smoking
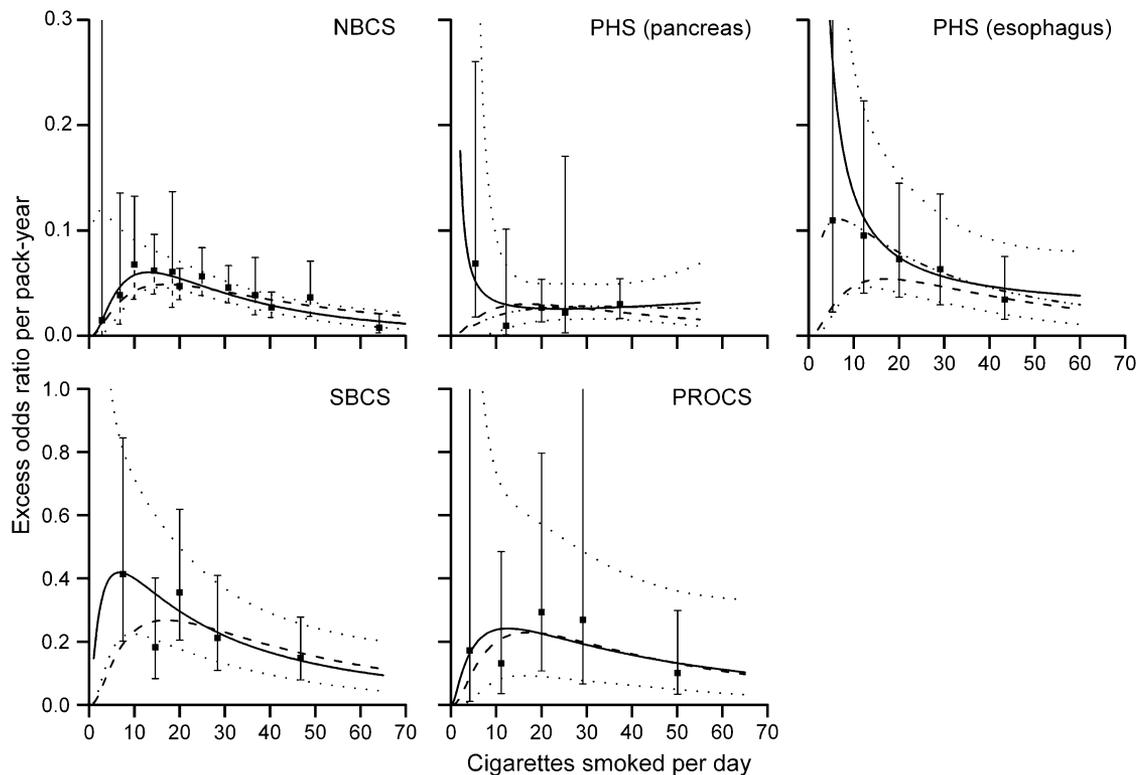
FIGURE 3. Estimates of excess odds ratio for cancer per pack-year of cigarette smoking in the National Bladder Cancer Study (NBCS) (1977–1978), the Investigation of Tumors that Occur Excessively among Blacks, designated the Population Health Study (PHS) (pancreas and esophagus cases and controls) (1986–1989), the Spanish Bladder Cancer Study (SBCS) (1998–2001), and the Puerto Rico Oral Cancer Study (PROCS) (1992–1995). The figure shows results based on a linear odds ratio model within categories of cigarettes smoked per day (black squares; bars, 95% confidence interval) and the estimated model 2 (equation 2 in text) fitted to data from each study (solid line; dotted lines, pointwise 95% confidence interval). The figure also shows results based on model 2 with the pack-years parameter (β) estimated and the intensity parameters fixed at summary values for all studies combined ($\varphi_1 = 2.72$ and $\varphi_2 = -0.479$) (dashed line). Model 2 was also fitted to PHS data that omitted smokers who smoked fewer than five cigarettes per day (dashed-dotted-dotted line).

practices among Native American miners, the low smoking rate in the study (smoking underground was banned in the 1950s), or misclassification of amount of smoking (35–39).

Smoking intensity effects were consistent within and across cancer sites, particularly after mild restrictions were imposed. We imposed restrictions rather than using a non-parametric generalized additive model (40) or a semiparametric spline model (13), which would have complicated comparisons across data sets. For completeness, we fitted cubic B-splines for intensity effects with one interior knot, and results were similar to those of the parametric models in figures 2 and 3. Confidence bounds were wide, particularly at low smoking intensities, highlighting the uncertainty of the direct intensity effect.

There is uncertainty about whether and to what extent intensity effects reflect smoking inhalation patterns and/or nicotine-related dependencies or fundamental biologic processes, such as modifications of activation pathways for carcinogens in cigarette smoke or DNA repair capacity. Patterns of variation in the EOR/pack-year by intensity could reflect inhalation practices if lower-intensity and higher-intensity smokers inhale fewer carcinogens per cig-

arette relative to moderate-intensity smokers. This would result in reduced risks at lower intensities and declining risks at higher intensities (1). Reduced frequency or reduced depth of inhalation is a plausible explanation; however, analysis of the ESHS data provided no indication that inhalation was related to smoking intensity after adjustment for total exposure (3).

The intensity effects may have been influenced by misclassification of the number of cigarettes smoked per day. However, consistency of effects for multiple diseases and for studies conducted in diverse locations suggests that an unusual pattern of misclassification would be required in order to fully explain the results. To induce an inverse intensity effect, misclassification would have to increase with increasing intensity above 15–20 cigarettes smoked per day, resulting in greater bias towards no effect of intensity and a decreasing exposure-response relation. To induce a direct intensity effect, misclassification would have to decrease with intensity up to 15–20 cigarettes smoked per day, resulting in less bias towards no effect and an increasing exposure-response relation. If only low and high intensities were differentially misclassified, then differential

TABLE 2.   Parameter estimates obtained from fitting models of the cigarette smoking–cancer relation to data from several studies*

Cancer site	Study (ref. no.)	β	φ_1	φ_2	p value for test of hypothesis		
					Linearity†	Intensity‡	Consistency§
Analyses of unrestricted data							
Lung	European Smoking and Health Study (6)	0.005	2.89	−0.51	0.65	<0.01	0.90
Lung	German Radon Study (7)	0.007	3.14	−0.50	0.87	<0.01	0.10
Lung	North American Radon Pooling Project (10)	2.45	−1.02	0.19	0.52	0.08	<0.01
Lung	Gansu Lung Cancer Study (11)	0.0004	2.84	−0.38	0.57	0.25	0.12
Lung	Colorado Plateau Uranium Miners Study (12)	0.014	1.79	−0.43	0.50	0.06	0.20
Bladder	National Bladder Cancer Study (14)	0.001	3.06	−0.59	0.52	<0.01	0.05
Bladder	Spanish Bladder Cancer Study (15)	0.147	1.10	−0.29	0.08	<0.01	0.02
Oral cavity	Puerto Rico Oral Cancer Study (16)	0.032	1.59	−0.31	0.07	0.28	0.94
Pancreas	Population Health Study (18)	0.593	−1.97	0.31	0.38	0.78	0.30
Esophagus	Population Health Study (17)	2.46	−1.58	0.14	0.31	0.01	0.01
Analyses of restricted data							
Lung	North American Radon Pooling Project¶	0.017	1.91	−0.27	0.15	0.06	<0.01
Bladder	National Bladder Cancer Study#	0.001	2.95	−0.57	0.90	<0.01	0.19
Bladder	Spanish Bladder Cancer Study#	0.061	1.68	−0.36	0.80	0.03	0.29
Pancreas	Population Health Study**	0.001	2.16	−0.32	0.49	0.97	0.48
Esophagus	Population Health Study**	0.043	1.01	−0.27	0.11	0.05	0.29

* Odds ratio $= 1 + \beta \times d \times \exp(\varphi_1 \ln(n) + \varphi_2 \ln(n)^2)$, where d is total pack-years of exposure, n is the number of cigarettes smoked per day, and β, φ_1, and φ_2 are parameters. Adjustment factors and descriptions of the data are given in table 1.

† p value for departure from linearity of the excess odds ratio ($\delta = 0$ in model 3 (equation 3 in text)).

‡ p value from a 2-df chi-squared test of no intensity effect ($\varphi_1 = 0$, $\varphi_2 = 0$).

§ p value from a 2-df chi-squared test of homogeneity of intensity parameters with weighted mean of intensity parameters for all data sets ($\varphi_1 = 2.72$ and $\varphi_2 = -0.48$).

¶ Excludes smokers who smoked fewer than 10 cigarettes per day.

\# Excludes smokers who started smoking at age 12 years or under.

** Excludes smokers who smoked fewer than five cigarettes per day.

misclassification would occur with pack-years of smoking, since duration is measured with relative accuracy. This would induce concavity in the relation between odds ratios and pack-years at low and high intensities. However, in the current study, relations between odds ratios and pack-years remained linear over the full range of intensities, suggesting that misclassification does not explain the observed patterns.

We also conducted a sensitivity analysis to explore the possibility that intensity patterns were due to nicotine satiation effects, whereby for high-intensity smokers, numbers of cigarettes smoked per day increasingly overestimate the "internal" exposure rate (4). Those analyses indicated that while some overestimation of internal exposure rate by the number of cigarettes smoked per day may have occurred, it is unlikely that nicotine satiation factors fully explain the inverse exposure rate pattern.

Intensity patterns, particularly those above 15–20 cigarettes per day, could reflect a differential "willingness to participate" for cases and controls—that is, intensity-related differential selection bias. This is unlikely, since cases' diagnoses were incident, cases were typically interviewed within a few months of diagnosis, and participation proportions were generally high. In 13 of the 15 studies

included in these analyses, participation proportions were similar in cases and controls and exceeded 70 percent, with participation in two studies (the Gansu Lung Cancer Study and the Spanish Bladder Cancer Study) exceeding 86 percent. There were two exceptions: the German Radon Study, with 76 percent and 44 percent participation among cases and controls, respectively, and the Population Health Study (pancreas), with 46 percent and 76 percent participation among cases and controls. In the latter study, the lower case proportion was due to the use of subject-only interviews. Intensity patterns were no different in these two studies. In addition, for differential enrollment to have influenced results, one would have to postulate that, conditional on equal total pack-years, the proportion of nonparticipating controls relative to nonparticipating cases was greater for intensities in the range of 15–25 cigarettes per day (resulting in too few controls and an increased EOR/pack-year) and then declined with increasing intensity (resulting in a decreased EOR/pack-year). In addition, the pattern of varying proportions of nonparticipation would have had to be present across the diverse cancer sites.

The intensity function has the form $g(n) = \exp\{\varphi_1 \ln(n) + \varphi_2 \ln(n)^2\}$. Since pack-years is the product of duration of smoking, y, and intensity, n (i.e., $d = y \times n$) and

OR $= 1 + \beta \, d \, g(n)$, equation 2 can be written in terms of duration and intensity as

$$OR = 1 + \beta \, y \exp\{\kappa \ln(n) + \varphi_2 \ln(n)^2\},$$

where $\kappa = \varphi_1 + 1$. Thus, model 2 is comparable to a model that is linear in duration for a fixed intensity. However, interpretation of β and intensity effects κ and φ_2 is problematic when either y or n is fixed, because of the changing total pack-years of exposure.

The relative simplicity of model 2 derives from the approximate linearity of the odds ratios for pack-years within levels of intensity and from the fact that a single parameter, γ_i, fully characterizes linearity. Variations in γ_i (EOR/pack-year) then represent the modulating effects of intensity. In principle, we could develop a similar model for the EOR by pack-years with allowance for duration of exposure. However, odds ratios for pack-years within duration categories are not linear; that is, there is no comparable model 1 within duration categories. A simple characterization of changes in the EOR by pack-years with allowance for duration is not readily apparent.

Our analysis included never smokers, current smokers, and recent former smokers. The extent to which age, sex, time since stopping smoking, and other factors influenced estimates of EOR/pack-year and the shape of the intensity function is the subject of another analysis (4).

After adjustment for total smoking exposure, we found a general and specific consistency in the effects of smoking intensity across diverse smoking-related cancer sites, particularly decreasing effects of intensity above a level of 15–20 cigarettes smoked per day. Because of uncertainty at low intensities, it remains unclear whether enhanced intensity effects occur at smoking levels below 15 cigarettes per day, and results for low intensities should be interpreted cautiously. Further analyses are needed to verify these observations.

REFERENCES

1. Vineis P, Kogevinas M, Simonato L, et al. Levelling-off of the risk of lung and bladder cancer in heavy smokers: an analysis based on multicentric case-control studies and a metabolic interpretation. Mutat Res 2000;463:103–10.
2. Vineis P, Alavanja M, Garte S. Dose-response relationship in tobacco-related cancers of bladder and lung: a biochemical interpretation. Int J Cancer 2004;108:2–7.
3. Lubin JH, Caporaso N. Cigarette smoking and lung cancer: modeling total exposure and intensity. Cancer Epidemiol Biomarkers Prev 2006;15:517–23.
4. Lubin JH, Caporaso N, Wichmann HE, et al. Cigarette smoking and lung cancer: modeling effect modification of total exposure and intensity. Epidemiology (in press).
5. Lubin JH, Blot WJ. Lung cancer and smoking cessation—patterns of risk. J Natl Cancer Inst 1993;85:422–3.
6. Lubin JH, Blot WJ, Berrino F, et al. Patterns of lung cancer risk according to type of cigarette smoked. Int J Cancer 1984;33:569–76.
7. Wichmann HE, Rosario AS, Heid IM, et al. Increased lung cancer risk due to residential radon in a pooled and extended analysis of studies in Germany. Health Phys 2005;88:71–9.
8. Kreuzer M, Heinrich J, Wolke G, et al. Residential radon and risk of lung cancer in Eastern Germany. Epidemiology 2003;14:559–68.
9. Kreienbrock L, Kreuzer M, Gerken M, et al. Case-control study on lung cancer and residential radon in Western Germany. Am J Epidemiol 2001;153:42–52.
10. Krewski D, Lubin JH, Zielinski JM, et al. Residential radon and risk of lung cancer—a combined analysis of 7 North American case-control studies. Epidemiology 2005;16:137–45.
11. Wang Z, Lubin JH, Wang L, et al. Residential radon and lung cancer risk in a high-exposure area of Gansu Province, China.Am J Epidemiol 2002;155:554–64.
12. Langholz B, Thomas D, Xiang A, et al. Latency analysis in epidemiologic studies of occupational exposures: application to the Colorado Plateau uranium miners cohort. Am J Ind Med 1999;35:246–56.
13. Hauptmann M, Berhane K, Langholz B, et al. Using splines to analyse latency in the Colorado Plateau uranium miners cohort. J Epidemiol Biostat 2001;6:417–24.
14. Hartge P, Silverman D, Hoover R, et al. Changing cigarette habits and bladder cancer risk—a case-control study. J Natl Cancer Inst 1987;78:1119–25.
15. Samanic C, Kogevinas M, Dosemeci M, et al. Smoking and bladder cancer in Spain: effects of tobacco type, timing, environmental tobacco smoke, and gender. Cancer Epidemiol Biomarkers Prev 2006;15:1348–54.
16. Hayes RB, Bravo-Otero E, Kleinman DV, et al. Tobacco and alcohol use and oral cancer in Puerto Rico. Cancer Causes Control 1999;10:27–33.
17. Brown LM, Hoover R, Gridley G, et al. Drinking practices and risk of squamous-cell esophageal cancer among Black and White men in the United States. Cancer Causes Control 1997;8:605–9.
18. Silverman DT, Dunn JA, Hoover RN, et al. Cigarette smoking and pancreas cancer—a case-control study based on direct interviews. J Natl Cancer Inst 1994;86:1510–16.
19. Alavanja MCR, Brownson RC, Lubin JH, et al. Residential radon exposure and lung cancer among nonsmoking women. J Natl Cancer Inst 1994;86:1829–37.
20. Field RW, Steck DJ, Smith BJ, et al. Residential radon gas exposure and lung cancer: The Iowa Radon Lung Cancer Study. Am J Epidemiol 2000;151:1091–102.
21. Alavanja MCR, Lubin JH, Mahaffey JA, et al. Residential radon exposure and risk of lung cancer in Missouri. Am J Public Health 1999;89:1042–8.
22. Schoenberg JB, Klotz JB, Wilcox HB, et al. Case-control study of residential radon and lung cancer among New Jersey women. Cancer Res 1990;50:6520–4.
23. Sandler DP, Weinberg CR, Shore DL, et al. Indoor radon and lung cancer risk in Connecticut and Utah. J Toxicol Environ Health A 2006;69:633–54 .

24. Letourneau EG, Krewski D, Choi NW, et al. Case-control study of residential radon and lung cancer in Winnipeg, Manitoba, Canada. Am J Epidemiol 1994;140:310–22.
25. Weinberg CR, Sandler DP. Randomized recruitment in case-control studies. Am J Epidemiol 1991;134:421–32.
26. Weinberg CR, Wacholder S. The design and analysis of case-control studies with biased sampling. Biometrics 1990;46:963–75.
27. Preston DL, Lubin JH, Pierce DA, et al. Epicure user's guide. Seattle, WA: HiroSoft International Corporation, 2006.
28. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
29. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10:1665–77.
30. Hu J, Galeone C, Lui R, et al. Smoking and lung cancer in Harbin, northeast China. Ann Oncol 2005;16:1605–8.
31. Yao SX, Lubin JH, Qiao YL, et al. Exposure to radon progeny, tobacco use and lung cancer in a case-control study in southern China. Radiat Res 1994;138:326–36.
32. Xu ZY, Blot WJ, Xiao HP, et al. Smoking, air pollution, and the high rates of lung cancer in Shenyang, China. J Natl Cancer Inst 1989;81:1800–6.
33. Yuan JM, Ross RK, Wang XL, et al. Morbidity and mortality in relation to cigarette smoking in Shanghai, China—a prospective male cohort study. JAMA 1996;275:1646–50.
34. Gao YT, Blot WJ, Zheng W, et al. Lung cancer and smoking in Shanghai. Int J Epidemiol 1988;17:277–80.
35. Hornung RW, Deddens JA, Roscoe RJ. Modifiers of lung cancer risk in uranium miners from the Colorado Plateau. Health Phys 1998;74:12–21.
36. Roscoe RJ, Deddens JA, Salvan A, et al. Mortality among Navajo uranium miners. Am J Public Health 1995;85:535–40.
37. Samet JM, Pathak DR, Morgan MV, et al. Radon progeny exposure and lung cancer risk in New Mexico uranium miners—a case-control study. Health Phys 1989;56:415–21.
38. National Research Council, National Academy of Sciences. Health effects of exposure to radon (BEIR VI). Washington, DC: National Academy Press, 1999.
39. Whittemore AS, McMillan A. Lung cancer mortality among United States uranium miners—a reappraisal. J Natl Cancer Inst 1983;71:489–99.
40. Rachet B, Siemiatycki J, Abrahamowicz M, et al. A flexible modeling approach to estimating the component effects of smoking behavior on lung cancer. J Clin Epidemiol 2004;57: 1076–85.