**eAppendix**

**A. The joint association of pack-years and cigarettes/day**

Relative risks (RR) by categories of pack-years, relative to never smokers, increased within each category of cigarettes/day, with the RR trend smaller at the higher intensities (eTable 1).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| eTable 1: Relative risks a and 95% confidence intervals by categories of pack-years of smoking and cigarettes/day. Data from the Atherosclerosis Risk in Communities (ARIC) Study. | | | | | | |
|  | Pack-years of smoking | | | | | |
| Cigs/day | 0 | 1-19 | 20-29 | 30-39 | 40-49 | ≥50 |
| 0 | 1.00 b |  |  |  |  |  |
| <10 |  | 1.29 | 1.63 |  |  |  |
|  |  | (1.1,1.5) | (0.7,3.9) |  |  |  |
| 10-19 |  | 1.15 | 1.80 | 2.11 | 2.98 |  |
|  |  | (1.0,1.3) | (1.5,2.1) | (1.7,2.7) | (2.0,4.5) |  |
| 20-29 |  | 1.30 | 1.58 | 1.80 | 1.97 | 2.39 |
|  |  | (1.1,1.6) | (1.3,1.9) | (1.5,2.1) | (1.7,2.3) | (2.0,2.8) |
| 30-39 |  | 1.06 | 1.36 | 1.20 | 1.94 | 2.16 |
|  |  | (0.7,1.7) | (0.8,2.2) | (0.8,1.7) | (1.5,2.5) | (1.8,2.6) |
| ≥40 |  | 1.81 | 0.70 | 1.70 | 1.25 | 1.93 |
|  |  | (1.0,3.1) | (0.4,1.3) | (1.1,2.6) | (0.8,1.9) | (1.7,2.2) |
| a Relative risks adjusted for center, race, birth year, age, sex, education, alcohol consumption, high blood pressure, diabetes mellitus, total cholesterol, body mass index and use of cigars or pipe exclusively.  b Referent category. | | | | | | |

**B. Effect modification by smoking-related factors**

The table provides parameter estimates for the fit of equation 3 to all data and within levels of several smoking-related effect modifiers, and P-values for tests of homogeneity. Results indicate no variation in the patterns of smoking risks by age started smoking (P=0.5), extent of inhalation (P=0.74) or additional use of cigars or pipes (O=0.79), but significant variation with years since cessation of cigarette smoking (P<0.01).

|  |  |  |  |
| --- | --- | --- | --- |
| eTable 2: Parameter estimates of relative risk (RR) regression model a for cardiovascular disease by pack-years and cigarettes/day and the P-value for homogeneity of estimates across smoking-related effect modifiers. Data from the Atherosclerosis Risk in Communities Study. | | | |
| Modifier | β | **γ** | P b |
| None | 0.184 | -0.697 |  |
| Age started smoking |  |  |  |
| <16 | 0.305 | -0.831 | 0.69 |
| 16-17 | 0.135 | -0.592 |  |
| 18-19 | 0.149 | -0.676 |  |
| ≥20 | 0.184 | -0.697 |  |
| Inhalation |  |  |  |
| Never/slightly | 0.171 | -0.699 | 0.68 |
| Moderately | 0.172 | -0.683 |  |
| Deeply | 0.305 | -0.827 |  |
| Cessation of smoking | |  |  |
| <1 | 0.194 | -0.719 | <0.01 |
| 1-9 | 0.259 | -0.627 |  |
| 10-19 | 0.092 | -0.585 |  |
| ≥20 | 0.135 | -0.684 |  |
| Use of cigars/pipes c |  |  |  |
| No | 0.189 | -0.705 | 0.97 |
| Yes | 0.162 | -0.661 |  |
| a All models adjusted for center, race, birth year, age, sex, education, alcohol consumption, high blood pressure, diabetes mellitus, total cholesterol, body mass index and use of cigars or pipe exclusively. For continuous pack-years (*d*) and cigarettes/day (*n*) with categorical modifying factor *x* with F levels, data fitted using: .  b P-value for test of homogeneity of smoking effects across factor *f* , i.e., *β1=…=βF* and *γ1=…= γF*,.  c Cigar/pipe use in addition to cigarettes. | | | |

The eFigure show that the strength of the pack-years and CVD association, i.e., the ERR/PKY estimates, declined smoothly with cigarettes/day within each level of age started smoking, inhalation and use of cigars/pipes (eFigures 1-S3).



eFigure 1: For age at started smoking, estimated excess relative risk/pack-year within categories of cigarettes/day (solid symbol) and fitted models for continuous pack-years (*d*) and cigarettes/day (*n*) (solid line, ; and dash line, ). Omitted dash line signifies that there was little visual improvement with extended model. Data from the Atherosclerosis Risk in Communities (ARIC) Study.



eFigure 2: For depth of inhalation, estimated excess relative risk/pack-year within categories of cigarettes/day (solid symbol) and fitted models for continuous pack-years (*d*) and cigarettes/day (*n*) (solid line, ; and dash line, ). Omitted dash line signifies that there was little visual improvement with extended model. Data from the Atherosclerosis Risk in Communities (ARIC) Study.



eFigure 3: For additional use of cigars/pipes, estimated excess relative risk/pack-year within categories of cigarettes/day (solid symbol) and fitted models for continuous pack-years (*d*) and cigarettes/day (*n*) (solid line, ; and dash line, ). Omitted dash line signifies that there was little visual improvement with extended model. Data from the Atherosclerosis Risk in Communities (ARIC) Study.

**C. CVD and smoking within a competing risks framework**

A Reviewer raised concern that competing risks, in particular lung cancer and other smoking-related diseases, may have preferentially removed heavy smokers from the population and influenced the CVD results in the main text. In response, we reanalyze smoking data accounting for multiple competing risks 1,2. For completeness, we jointly analyzed multiple outcomes, including CVD, lung cancer, other selected smoking-related cancers (esophagus, larynx, oropharynx, bladder, kidney, stomach, colon, rectum and pancreas) and all other deaths. We used a data augmentation approach which replicated the dataset for each outcome type and added an outcome-specific stratification variable as a covariate 1,3. For consistency with the main text, we used Poisson regression which necessitated replicating the appropriate person-time table for each outcome type 4. We used the Epicure software package, in particular DATAB for the generation of the person-time table, which we replicated and appended four times including a stratum indicator and specific event counts for each outcome, and AMFIT for the Poisson regression analyses 5.

*Data structure*

Analyses in the main text included 14,127 participants and 232,002 person-years with 3,391 CVD events, with follow-up defined through the earliest date of CVD event, death, lost to follow-up or 31 December 2008, the end of study period. Comprehensive information on cancer outcomes was readily accessible only through 31 December 2006. Therefore, the following analyses used that date as the end of the study period. We first repeated the analysis in the main text within the restricted dataset, including 215,116 person-years and 3,074 CVD events.

A second analysis incorporated competing risks as defined by the four outcome types. This analysis omitted 810 participants with a pre-enrollment doctor’s diagnosis of cancer, a total of 193 who would develop CVD during follow-up, leaving 13,317 participants. Through 2006, there were 202,182 person-years and 4,742 events, including 2,638 CVD cases and 350 lung, 684 other smoking-related (including 112 bladder, 280 colon and rectum and 114 stomach cancers) and 1,273 deaths from other causes (including 401 cancers, 363 diseases of the circulatory system and 147 diseases of the respiratory system). In addition, 232 (1.7%) were missing or had unknown final status and censored at last contact.

*Results for CVD and cigarette smoking*

Marginal RRs for pack-years and cigarettes/day for the full data (main text Table 1, column denoted “full FU”) were similar to the RRs for the restricted follow-up (denoted “Restricted FU”). For the joint analysis of pack-years and cigarettes/day using the restricted FU, the fitted equation 3 (red solid line) closely tracked the estimates of the ERR/pack-years within cigarettes/day categories (solid red symbol) (eFigure 4, upper left panel). This pattern was very similar to the fitted results from the standard analysis in the main text for the full data (red dash-dot line).

Incorporating competing risks, the fitted equation 3 (black solid line) closely tracked the estimated ERR/pack-years for CVD within cigarettes/day categories (solid black symbol) (eFigure 4, upper left panel). As in the main text, the fitted model used , since P=0.79 for test of in equation 3. There was a statistically significant variation in the strength of the pack-years association with cigarettes/day (P<0.01 for the test of ).

In equation 3, the γ parameter defined the (decreasing) strength of association for CVD with increasing cigarettes/day. Estimates with 95% CI of γ were -0.70 (-0.87, -0.52), -0.71 (-0.90, -0.53) and -0.75 (-0.96, -0.55) for the analysis from the main text with the full follow-up, the analysis in the restricted data and the competing risks analysis, respectively. Each analysis resulted in an approximately inverse root two rate of decline. The corresponding estimates of the RRs for 50 pack-years accrued at 20 cigarettes/day were 2.14, 2.20 and 2.14, respectively.

*Results for lung and other cancers and cigarette smoking*

Although numbers of events were limited, we evaluated the patterns of ERR/pack-year and their variation with cigarettes/day within the competing risk framework for CVD, lung cancer, other selected smoking-related cancers and other mortality. This enabled a comparison of the smoking intensity patterns across outcome types and in particular with previous analyses of smoking and cancer 6-8. We found increased marginal RRs for pack-years for lung cancer and other smoking-related cancers, after adjustment for cigarettes/day, while homogeneity of the strength of association by categories of cigarettes/day was rejected for lung cancer (P<0.01), but not for other smoking-related cancer (P=0.76) or other cancers (P=0.91) (eTable 3).

Previous analyses suggested that patterns of the variations in the strength of the pack-years association with cigarettes/day, i.e., the delivery rate effects, were consistent across diverse cancer sites. This suggested that the substantial smoking-related differences in risks among various smoking-related cancer sites derived primarily from the differential strengths of association for pack-years (β) and not the relative consequences of smoking longer durations at lower cigarettes/day versus shorter durations at higher cigarettes/day. In the ARIC data, we found that the strength of association, i.e., the ERR/pack-years, decreased above about 15-20 cigarettes/day for lung cancer, in agreement with previous analyses 6-8. The variation in the strength of association, g(.), for lung cancer was significantly different than the pattern for CVD (P<0.01). In contrast, the test of homogeneity of the smoking rate effects for lung cancer and other smoking-related cancers was not rejected (P=0.62). The latter indicated that models with a common curvature function g(.) were consistent for smoking-related RRs of lung cancer and other smoking-related cancer outcomes (eFigure 4 solid green line). However, for other smoking-related cancers, the smoking rate effects did not reach statistical significance (P=0.11 for the test of ), indicating that RRs were consistent with a simple linear increase in pack-years, perhaps due to limited numbers of events. For mortality from other causes, RRs with pack-years increased linearly (P=0.53 for the test of no departure from linearity) and with no variation in the strength of association (P=0.12).

*Effect modification by years since cessation of smoking*

Table 3 revealed a significant modification of the ERR/pack-years for CVD only with years since cessation of smoking. We therefore also evaluated potential bias within the competing risks context for evaluation of effect modification of ERR/pack-year by years since last smoked. For CVD, the estimates of the fitted ERR/pack-year at 20 cigarettes/day for <1, 1-4, 5-9 and 10+ years since last smoked were 0.18 (0.09, 0.34), 0.31 (0.16, 0.58), 0.13 (0.06, 0.27) and 0.13 (0.06, 0.26) for the restricted follow-up through 2006 without adjustment for competing risks, which were very similar to 0.19 (0.09, 0.38), 0.35 (0.18, 0.70), 0.11 (0.05, 0.27) and 0.15 (0.07, 0.32) with the adjustment for competing risks. Adjustment for competing risks had minimal impact on inference for effect modification.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| eTable 3. Numbers of incident events, relative risks (RR) with 95% confidence intervals (CI) by pack-years of cigarette smoking and cigarettes smoked per day. Data from the Atherosclerosis Risk in Communities Study a. | | | | | | | | | | | | |
|  | Cardiovascular disease b | | | | Competing risks analysis (restricted FU) c | | | | | | | |
|  |  | | | |  | | | | | | | |
|  | Full FU | | Restricted FU | | Cardiovascular disease | | Lung cancer | | Selected smoking related cancers | | Other mortality | |
|  | RR b | 95% CI | RR b | 95% CI | RR | 95% CI | RR | 95% CI | RR | 95% CI | RR | 95% CI |
| Never smokers d | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  |
| Pack-years of smoking | |  |  |  |  |  |  |  |  |  |  |  |
| 1-19 | 1.29 | (1.1,1.5) | 1.35 | (1.2-1.6) | 1.35 | (1.2,1.6) | 3.18 | (1.6,6.4) | 1.48 | (0.9,2.3) | 1.20 | (1.0,1.5) |
| 20-29 | 1.75 | (1.4,2.1) | 1.86 | (1.5-2.3) | 1.83 | (1.5,2.3) | 9.60 | (3.9,23.9) | 1.85 | (0.9,3.7) | 1.63 | (1.2,2.2) |
| 30-39 | 1.98 | (1.6,2.4) | 2.13 | (1.7-2.7) | 2.11 | (1.7,2.7) | 16.77 | (6.8,41.5) | 2.42 | (1.2,4.9) | 2.29 | (1.7,3.2) |
| 40-49 | 2.21 | (1.8,2.7) | 2.47 | (2.0-3.1) | 2.37 | (1.9,3.0) | 29.65 | (11.8,74.6) | 3.46 | (1.7,7.1) | 2.32 | (1.6,3.3) |
| ≥50 | 2.73 | (2.2,3.4) | 3.05 | (2.4-3.9) | 2.79 | (2.2,3.6) | 63.59 | (25.1,161.1) | 3.82 | (1.8,8.0) | 4.32 | (3.1,6.1) |
| P f | <0.01 |  | <0.01 |  | <0.01 |  | <0.01 |  | <0.01 |  | <0.01 |  |
| Cigarettes/day | |  |  |  |  |  |  |  |  |  |  |  |
| 1-9 | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  |
| 10-19 | 0.98 | (0.8,1.2) | 0.93 | (0.8-1.1) | 0.96 | (0.8,1.2) | 0.99 | (0.5,2.1) | 0.70 | (0.4,1.2) | 1.21 | (0.9,1.6) |
| 20-29 | 0.91 | (0.8,1.1) | 0.85 | (0.7-1.0) | 0.88 | (0.7,1.1) | 0.78 | (0.4,1.7) | 0.75 | (0.4,1.4) | 0.86 | (0.6,1.2) |
| 30-39 | 0.78 | (0.6,1.0) | 0.72 | (0.6-0.9) | 0.74 | (0.6,1.0) | 0.45 | (0.2,1.1) | 0.53 | (0.3,1.1) | 0.78 | (0.5,1.1) |
| ≥40 | 0.70 | (0.6,0.9) | 0.64 | (0.5-0.8) | 0.68 | (0.5,0.9) | 0.38 | (0.2,0.9) | 0.49 | (0.2,1.1) | 0.83 | (0.6,1.2) |
| P f | <0.01 |  | <0.01 |  | <0.01 |  | <0.01 |  | 0.76 |  | 0.16 |  |
| Cases |  |  |  |  |  |  |  |  |  |  |  |  |
| Never-smokers | 1,098 |  | 971 |  | 831 |  | 25 |  | 91 |  | 473 |  |
| Smokers | 2,293 |  | 2,103 |  | 1,807 |  | 328 |  | 203 |  | 984 |  |
| a Compete data results from main text Table 1 with follow-up (FU) through 2008 and with restricted FU through 2006. Model included center, race, birth year, age, sex, education, alcohol consumption, high blood pressure, previous diabetes mellitus, total cholesterol, body mass index, use of cigars or pipe exclusively, pack-years and cigarettes/day, with the RR for 1-9 cigarettes/day set to one for identifiability.  b RRs from an analysis with censoring based on the earliest date of cardiovascular disease event, death or loss to FU.  c RRs from a competing risks analysis, with outcomes incident lung cancer and other smoking-related cancer (oropharynx, bladder, kidney and pancreatic) and deaths from all other causes.  d Referent category of never smokers.  f P-value for score test of no linear trend. | | | | | | | | | | | | |



eFigure 4: For competing risks analysis with follow-up (FU) through 2006, estimated excess relative risk/pack-year (ERR/pack-year) within cigarettes/day categories (solid black symbol), with continuous pack-years and cigarettes/day models (solid black line) for CVD, lung cancer, other selected smoking-related cancers and all other causes of mortality. For lung and other smoking-related cancers, models fitted with common curvature parameters (green line). For CVD, estimated ERR/pack-year using methods from main text (solid red symbol) and equation 3 (solid red line) fitted to restricted FU and full FU (dash-dot red line). All results adjusted for age, birth year, sex and other factors (see text). A common aspect ratio applied to all panels. Data from the Atherosclerosis Risk in Communities (ARIC) Study.

Reference List

(1) Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics.* 1995;51(2):524-532.

(2) Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Wiley; 2015.

(3) Larson MG. Covariate analysis of competing-risks data with log-linear models. *Biometrics.* 1984;40(2):459-469.

(4) Pierce DA, Preston DL. Joint analysis of site-specific cancer risks for the atomic-bomb survivors. *Radiat Res.* 1993;134(2):134-142.

(5) Preston DL, Lubin JH, Pierce DA, McConney ME. *Epicure User's Guide*. Seattle, Washington, USA: HiroSoft International Corporation; 2006.

(6) Lubin JH, Caporaso N. Cigarette smoking and lung cancer: modeling total exposure and intensity. *Cancer Epidemiol Biomarkers Prevention.* 2006;15(3):517-523.

(7) Lubin JH, Alavanja MCR, Caporaso N et al. Cigarette smoking and cancer: modeling total exposure and intensity. *Am J Epidemiol.* 2007;166(4):479-489.

(8) Lubin JH, Virtamo J, Weinstein SJ, Albanes D. Cigarette smoking and cancer: intensity patterns in the Alpha-Tocopherol Beta-Carotene Cancer Prevention Study in Finnish men. *Am J Epidemiol.* 2008;167(8):970-975.