Estimating smoking-attributable mortality in the United States: Geographic variation

Andrew Fenelon

Samuel H. Preston

Population Studies Center, University of Pennsylvania, 239 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19143 afenelon@sas.upenn.edu

Introduction

Tobacco is the largest single cause of premature death in the developed world and is growing in importance throughout the developing world. Recent estimates indicate that as many as 400,000 deaths annually in the United States (Mokdad et al., 2004) are caused by cigarette smoking. At the individual level, cigarette smoking is strongly linked to lung cancer; but smoking also confers increased risk of death from other cancers, heart diseases, stroke, and chronic respiratory conditions (Doll et al., 2004). CDC estimates of smoking attributable-risk indicate that only thirty percent of smoking-related deaths in the United States are caused by lung cancers.

Studies demonstrating the link between cigarette smoking and individual mortality typically involve detailed cohort smoking histories. The prospective study of British doctors beginning in 1951 (Doll et al., 2004) and the American Cancer Society's Cancer Prevention Studies Cohorts I and II (CPS-I and CPS-II) beginning in 1959 and 1982, respectively, provide rich data on the excess risks associated with cigarette smoking from a number of causes of death. While cohort studies often provide the most persuasive evidence of the increased mortality risk related to cigarette smoking behavior, they suffer from a few significant drawbacks. First, study cohorts may not be representative of the population-at-large and their experience may not reflect that of the general population of smokers (Thun et al., 1997). Second, smoking behavior categories may not reflect temporal changes in cohort smoking patterns, and many studies assume that baseline smoking status remains constant throughout the study. Finally, large-scale cohort studies require long periods of observation and detailed demographic information that may be unavailable for many relevant populations.

While cohort studies have been used to estimate the total number of deaths caused by cigarette smoking in a population (Rogers et al., 2005), others have attempted to deal with the problematic issues of cohort studies by using indirect methods. These studies, such as Peto, Lopez and colleagues (1992), use the death rate from lung cancer as an indicator of the accumulated damage from smoking in the population. Based on estimates of the relative risk of smokers compared with non-smokers of mortality from various causes of death, one can calculate the total mortality that would not have occurred in the absence of smoking. Preston, Glei, and Wilmoth (2009, 2010) developed an alternative method which relies on the statistical relationship between lung cancer and other causes of death across countries and time periods. Their method makes fewer assumptions than Peto-Lopez and produces results that are highly similar, validating the robustness of both approaches. In this paper, we use a method similar to Preston, Glei, and Wilmoth to calculate smoking-attributable mortality in the United States between 1990 and 2004. Rather than using the coefficients that they estimate from international data, however, we re-estimate the relationship using state-level data in the U.S. We then examine the extent to which cigarette smoking explains variation in adult mortality across U.S. divisions. Finally, we compare the results of various methods for estimating smoking attributable risk.

Background

Studies calculating the number of excess deaths due to cigarette smoking typically use an attributable-risk approach. They estimate the number of deaths that would not occur if smokers experienced the same death rates as non-smokers (Peto et al., 1994). This requires some estimate of the increased risk conferred to smokers by their behavior; researchers have developed many

different methods to calculate this excess risk and have applied them in a number of different settings.

The first set of methods could be termed *direct* methods, because the mortality differential between smokers and non-smokers is actually observed. These studies follow cohorts of smokers and non-smokers over periods of time to track the mortality experience of each respective population. In the United States, the most commonly cited study is the American Cancer Society Cancer Prevention Study Cohort II (CPS-II) which is composed of more than 1.2 million individuals followed since 1982. The Centers for Disease Control and Prevention (CDC) issues regular estimates of smoking-attributable mortality in the U.S. using relative risks from CPS-II (Adhikari et al., 2009) and separate estimates of smoking prevalence. Based on excess mortality among current smokers and former smokers relative to non-smokers, they calculate mortality attributable to cigarette smoking by applying relative risks to current smoking status data. They find that 440,000 annual deaths can be attributed to cigarette smoking in the late 1990s and early 2000s, more than one-fifth of all adult mortality.

The CDC approach has received criticism from those arguing that the relative mortality risks of smokers compared with non-smokers reflect more than just the effect of smoking. The assumption that smokers would have the same mortality experience as non-smokers in the absence of smoking ignores other behavioral and socioeconomic differences between smokers and non-smokers that may confer increased risk. Rogers et al. (2005) use the National Health Interview Survey 1990 supplement to control for many covariates of smoking behavior, which decreases the estimated number of deaths to 338,000.

While direct methods that control for smoking confounders offer relatively robust assessments of the mortality consequences of smoking, such extensive data on smoking status

and mortality is unavailable for most populations, and a more widely-applicable method is necessary. Furthermore, current-status smoking data used in these methods may not represent effective measures of exposure to smoking-related mortality. Peto, Lopez, and colleagues (1992) developed an *indirect* method for calculating smoking attributable mortality that relies exclusively on vital statistics data. Assuming that smoking behavior is the only factor which increases the risk of *lung cancer* death of smokers relative to non-smokers, they use CPS-II non-smoker lung cancer death rates to calculate age-specific 'proportion exposed' that reflects the prevalence of smoking-related damage. They then import relative risks for various disease categories from CPS-II and apply them to exposure composition of the population. In order to correct for confounding, they decrease the relative risks from causes of death other than lung cancer by half. They produced estimates of smoking-attributable mortality for the U.S. among other developed countries for the year 2000 (Peto et al., 2006). Although the lung cancer death rate does not directly measure the prevalence or intensity of smoking in a population, it may be a more reliable indicator of exposure to smoking-related damage than are self-reports and current-status surveys.

In practice the Peto-Lopez method is complex to implement and relies heavily on the relative risks from CPS-II. Moreover, it makes a rather arbitrary assumption regarding the confounding of smoking with other factors in producing cause-specific relative risks, simply halving them.[1] Preston, Glei and Wilmoth (2009) developed an alternative method to Peto-Lopez which makes fewer assumptions and does not rely heavily on the generalizability of CPS-II relative risks. Their method develops a statistical model to estimate the relationship between the lung cancer death rate and the death rate from other causes of death across developed

[1] Other studies using Peto-Lopez have relaxed this assumption, decreasing observed relative risks by different amounts or actually attempting to control for differences between smokers and non-smokers.

countries between 1950 and 2003. This relation is then used to estimate the mortality impact of smoking. They use lung cancer death rates of non-smokers in CPS-II between 1982 and 1988 (Thun et al., 1997) to produce an estimate of lung cancer attributable-risk and apply this each population of interest. The method produces results that are highly similar to those of Peto and Lopez without relying on strong assumptions regarding the relationship between smoking and cause-specific mortality apart from lung cancer itself.

A key advantage of Preston, Glei, and Wilmoth's is its applicability to varied populations. It can be implemented anywhere vital statistics data on causes of death exists for a number of populations or geographic units. In this paper, we apply the method using annual mortality data at the U.S. state level between 1990 and 2004. Based on the results of the model, we calculate smoking attributable mortality for the United States as a whole as well as for the nine Census divisions. Rates of cigarette smoking in the United States have been historically very high compared with Europe, decreasing only in recent years for males. But large regional differences in cigarette smoking behavior and related mortality remain. A few descriptive studies have examined the issue of geographic variation in smoking-attributable mortality in the United States (CDC, 2009) but none has fully explored the implications of smoking for geographic disparities in U.S. adult mortality. Following other recent studies demonstrating the potential for differences in smoking to explain mortality differentials (Preston and Wang, 2006, Jha et al., 2006), we demonstrate the impact of smoking-related mortality on regional patterns of mortality in the U.S. Finally, we compare attributable fraction estimates for the U.S. produced by a variety of different methods.

Data

We use vital statistics data for the fifty states annually between 1990 and 2004. Death data are available through the Multiple Cause-of-Death (MCD) public-use micro-data files released annually by the National Center for Health Statistics (NCHS). MCD files contain demographic, geographic, and cause-of-death information about all deaths occurring in the United States. Population denominators for death rate calculations come from bridged-race files available from the NCHS.[2] Mortality data by Census division are calculated by aggregating based on the scheme in Table A1. We restrict our analyses to ages 50 - 84 in order to capture the vast majority of smoking-related deaths.[3]

Method

Lung cancer is a unique condition in that it is so closely tied to one behavioral risk factor. While other causes of death have been shown to be linked to smoking behavior, none is related as strongly as is lung cancers. In CPS-II, smoking was responsible for more than 90% of lung cancer deaths among men and more than 70% among women (Thun et al., 1997). In places where reliable cohort smoking histories are unavailable for most of the population, the age-specific death rate from lung cancer has been used as an indirect indicator of the accumulated damage from cigarette smoking (Preston et al., 2009, Peto et al., 1992). If we then assume that variation in lung cancer death rates is almost exclusively the result of variation in cigarette smoking, the use of the lung cancer death rate as a population-level index of the impact of smoking is robust (see Preston et al. 2009 for a discussion of these issues).

Statistical Model

We use a variation of the method developed by Preston et al. (2009) to calculate smoking attributable mortality in the United States for the years 1990, 2000, and 2004. We estimate the

[2] Electronically from http://www.cdc.gov/nchs/nvss/bridged_race.htm, accessed May, 2009.
[3] The open-ended age interval produces irregular results for the attributable-risk estimation. More is discussed about this issue below.

relationship between the age-specific lung cancer death rate and the log of the death rate from other causes of death annually between 1990 and 2004. We use Poisson regression to predict the logarithm of death rate from causes other than lung cancer in five-year age groups from 50 – 54, 55 – 59, …, 80 – 84 as a function of the death rate from lung cancer

$$\ln(M_O) = \beta_L M_L + \beta_a X_a + \beta_S X_S + \beta_t T + \beta_{aL}(M_L \times X_a)$$

where M_L and M_O are the death rate for lung cancer and other causes respectively in each state, year, and five-year age group. X_a and X_S, are dummy variables for age-group and state respectively, while β_a and β_S are their corresponding coefficients. We include a linear time trend (*T*) as well as interactions between lung cancer mortality and age-group. We use age-specific population size as a statistical "offset" in the procedure to control for exposure to mortality. We estimate separate models by sex to allow for distinct relationships between smoking and mortality for men and women. The coefficients of interest are β_L and β_{aL}, denoting the age-specific relationship between lung cancer and other causes of death ($\beta_L^{'} = \beta_L + \beta_{aL}$); these coefficients are used to calculate the attributable fraction.

Attributable Risk Calculation

Lung cancer deaths attributable to cigarette smoking are estimated using values of lung cancer death rates among never-smokers, reported by Thun et al. (1997) from the CPS-II study between 1982 and 1988. The proportion of lung cancer deaths attributable to smoking is the ratio of smoking-related lung cancer death rate to the overall lung cancer death rate

$$A_L = \frac{M_L - M_L^*}{M_L}$$

where M_L^* is the lung cancer death rate among lifelong non-smokers, the expected death rate in the absence of smoking. While lung cancer mortality among never smokers does show some variation across populations (Thun et al., 2008), there is little evidence for long-term changes

across periods (Rosenbaum et al., 1998). The relationship between the prevalence of cigarette smoking and M_L^* is unclear.

Following Preston, Glei, and Wilmoth (2009), we calculate mortality attributable to smoking for causes of death other than lung cancer based on the relationship between lung cancer and other causes across states. First, we calculate the predicted number of deaths expected based on the observed lung cancer death rate. Next, we subtract the predicted number of deaths expected from the lung cancer death rates of non-smokers. We divide this difference by number of deaths expected from the observed rates to calculate the attributable fraction

$$A_O = \frac{e^{\beta_L^{'}(M_L)} - e^{\beta_L^{'}(M_L^*)}}{e^{\beta_L^{'}(M_L)}}$$

where $\beta_L^{'}$ is the model coefficient for lung cancer including age interactions ($\beta_L^{'} = \beta_L + \beta_{aL}$). The attributable fraction for total mortality is a weighted average of the attributable fractions for lung cancer and other causes

$$A = \frac{A_L D_L + A_O D_O}{D}$$

where D_L and D_O are deaths from lung cancer and other causes respectively and D is total deaths. In their application, Preston et al (2009) find that the estimated attributable fraction is generally robust to alternative specifications of age, time period, and interactions. Lung cancer also exhibits no statistically significant relationship with external causes of death which are assumed to be unrelated to smoking (Peto et al. 1992).

Two sets of parameters have been estimated using the Preston/Glei/Wilmoth (PGW) model. PGW (2010) dropped observations for ages 85+ from the data set used by PGW (2009) because they were subject to age misreporting and, as an open-ended interval, to extraneous influences resulting from differences in age distributions. These effects had produced a set of

parameters that were implausible at the oldest ages. Dropping these observations produced a smoother sequence of coefficients at older ages and reduced by a modest amount the fraction of deaths attributable to smoking. We use the results of PGW (2010) in comparisons reported below.

Variation in Mortality by U.S. Division

We estimate smoking-attributable mortality for ages 50+ for the United States as well as the nine U.S. Census Divisions. We calculate life expectancy at age 50 both including and excluding smoking-related deaths. Age-specific death rates in the absence of smoking (M_{abs}) include only those deaths not attributed smoking by our model

$$M_{abs} = \frac{D - D_A}{P}$$

where D_A the number of deaths attributed to smoking and P is the number of person-years of exposure. We then recalculate life tables for each Division with smoking-related deaths removed.[4]

Results

Table 1 shows estimated coefficients of the relationship between lung cancer and other causes by age-group and sex. If exponentiated, they can be interpreted as the proportional increase in death rate of other causes resulting from an increase in the lung cancer death rate of one per thousand, all else being equal. Coefficients are smaller at higher ages, reflecting both higher death rates overall and more varied factors influencing mortality at higher ages. Assumed lung cancer death rates among lifelong non-smokers from CPS-II are presented in Table 2. Given that we assume smoking to be the predominant factor determining population variation in lung cancer death rates, the rates in Table 2 are intended to approximate conditions in which

[4] We elect to simply 'remove' smoking deaths from the life table calculation as opposed to using 'cause-deleted' life tables in order to preserve the simplicity of interpretation. The results do not change substantively.

smoking was eliminated. The difference between these rates and observed lung cancer rates is used to calculate lung cancer attributable-risk. We find that more than 80% of lung cancer deaths among women and more than 90% among men can be attributed to smoking.

Smoking attributable fractions for the U.S. and by Division are displayed in Table 3 for 1990, 2000, and 2004. Between 1990 and 2004, women experienced increases in smoking-related mortality while men experienced decreases. In 1990, smoking was responsible for 28% of deaths among men and 15% among women in the age range 50 – 84. By 2004, it had decreased to 22.2% for men and increased to 19.0% for women. This finding is consistent with previous estimates of smoking-related mortality in the U.S. as well as studies demonstrating narrowing sex mortality differentials related to cohort changes in smoking behavior (Preston and Wang 2006). Historically high rates of smoking among American men have begun to decline, while the rates of American women have become more similar to men's.

There was little variation in attributable fraction by division among women in 1990. By 2004, the Mountain division had begun to emerge with a substantially lower fraction, actually declining from 17.9% in 2000 to 15.7% and 2004. Divisional differences in the impact of smoking have increased for women but are especially striking for men. In 1990, males in the East South Central and West South Central Divisions had attributable fractions greater than 31%; while in the Mountain and Pacific Divisions they were less than one-quarter. When considered alongside the high mortality already present the in the Southern states, the sizeable influence of smoking on these areas is even more salient. While smoking-attributable mortality has declined in each division in recent years, declines have been slower in southern divisions. In 2004, smoking accounted for nearly 30% of deaths among men in the East South Central Division compared with only 16% and 18% in the Mountain and Pacific Divisions respectively.

Table 4 presents the expected number of years lived between ages 50 and 84 both including and excluding smoking-related deaths. The difference between these two figures represents number of years lost to smoking-related conditions among older Americans. In 1990, females lived one year less between 50 and 84 as a result of smoking-related mortality; males lost 2.5 years to smoking. This value increased slightly to 1.1 years for women and decreased to 1.7 years for men by the year 2004. In 2004, the greatest impact of smoking was found in the West South Central and East South Central divisions. Women lost 1.4 and 1.2 years respectively in these divisions, while men lost 2.5 and 1.9 years respectively. Among men, the East South Central division is particularly strongly affected by the smoking epidemic. In both 1990 and 2004, this division experiences substantially higher smoking-related mortality than other divisions, and its survivorship has suffered accordingly. The Mountain and Pacific Divisions perform particularly well with respect to smoking, each losing less than 1.3 years in 2004.

Finally, we measure the extent to which smoking explains geographic variation in mortality across states and divisions in the United States. We compare the variance in life expectancy across states before and after removing smoking-related deaths. The proportional reduction in life expectancy represents the variance explained by mortality related to smoking. We find that smoking accounts for 39% of state variation in life expectancy among women in 1990 and 38% in 2004. Among men, it is more important, explaining 81% in 1990 and 65% in 2004. The sex difference reflects greater overall importance of smoking as well as a stronger correlation with state-specific mortality experience among men.

Comparison with Alternative Methods

Researchers have developed a number of direct and indirect methods for estimating the number of deaths in a population attributable to cigarette smoking. Different methods place

varying emphasis on assumptions regarding the relationship between smoking and mortality at the individual level, some assuming that all excess mortality among smokers is the result of smoking while others attempt to control for potential confounders. Table 5 shows estimated U.S. smoking-attributable fractions using many different methods. The first row shows estimates of smoking-attributable fraction for U.S. men and women in 1990 and 2004 using our method. Row 2 shows estimates obtained using coefficients found by Preston, Glei, and Wilmoth (2010) using a related method across a sample of 21 developed countries. Their attributable risk estimates are very similar to ours, especially for men. Their estimates for women are slightly higher (0.20 vs 0.186). Row 3 shows estimates using the Peto-Lopez method reported in Peto et al. (2006) for ages 35+, which are slightly higher than ours, but refer to the year 2000. Estimates from the CDC (Row 4) are generally consistent with ours although lower among women in 2004 (0.153). The estimates made by Rogers et al. (2005) using smoking-status data from the NHIS are substantially lower than our estimates for females (0.126) and quite similar to ours for males (0.212). The Rogers estimates would be downwardly biased if smoking status at baseline is misclassified or if any change in smoking status occurred during the seven-year follow-up period.

We can also compare our estimates to state-specific estimates made by the Center for Disease Control (2009). CDC used the Behavioral Risk Factor Surveillance Survey (BRFSS) to estimate the prevalence of smoking by state. This data source is based on telephone surveys and has a response rate that differs by state, in part because states have control over how the BRFSS is executed (i.e., questionnaire length, whether data collection is in-house or contracted out, sampling design) The national response rate in 2004 was 52.7% (Schneider and Latane 2007). CDC combined these estimated prevalences with estimates of the proportion of deaths from

various causes that are attributable to smoking, estimates that were drawn from deaths for 1982-88 (CDC 2009). Data used in the CDC estimates is thus somewhat dated and subject to reporting biases.

Despite considerable differences between our method and that of CDC, estimated attributable fractions are highly consistent across states and divisions. Figure 1 shows the comparability of CDC (2009) state-specific attributable risk estimates and those based on our method for 2004. For both men and women, the correlation is very high (0.87). Compared with CDC estimates, our attributable fractions are slightly higher for women and slightly lower for men, but the strong correlation suggests that both methods find the same geographic pattern of smoking-related mortality within the US.

Methodological Implications

The PGW model, unlike the approach used by the CDC and by Peto and Lopez (1992), does not borrow relative risk estimates from prospective studies of smokers and non-smokers. It uses lung cancer mortality as an indicator of the damage from smoking and assumes that such damage can be identified in other causes of death by modeling the relation between lung cancer mortality and mortality from other causes. Parameters of that model have been estimated using international and intertemporal data in PGW (2009, 2010).

The present paper applies the PGW model to recent cross-state data in the US. Table 6 compares coefficients from the present set of estimates to those of PGW (2010). The four sets of estimates are graphed in Figure 2. Several patterns are clearly evident:

1) The sets of coefficients in the contemporary US are quite similar for men and women, suggesting that lung cancer mortality is functioning in the US as stable indicator of the incremental mortality risk, presumably associated with smoking, for other causes

of death. On the other hand, female coefficients are much larger than male coefficients in the international data set investigated by PGW (2010).

2) Except at age 50-54, male coefficients estimated using US data are remarkably similar to those estimated from the international/intertemporal data. This provides an encouraging indication that the overall approach to estimating the impact of smoking is reliable for males.
3) The outlier series is the set of coefficients for females estimated from the international/intertemporal data. Coefficients for this series are substantially higher than those from the other three series at nearly every age interval.

We suspect that the high coefficients for women in the PGW (2010) series are a result of the recency of the smoking epidemic for women in that data set. The data set begins with observations from the early 1950's for all 21 countries. In most of these countries, few older women were smoking during that era. PGW's (2010) estimates of attributable risk from smoking for women in 1955 are above 0.01 in only two of 21 countries. In contrast, the median value for males was already 0.07 by 1955. The maturity of the smoking epidemic may, for example, affect the relation between lung cancer mortality and mortality from other causes of death by virtue of different lags in the relation between smoking and different causes of death. The US has a relatively mature smoking epidemic among both men and women (Forey et al. 2002). We suggest that the US coefficients for women in Table 6 may be more appropriate for countries such as the US with a long-standing tradition of women's smoking, whereas the PGW estimates for women may be more appropriate for relative newcomers. The choice is not as consequential for estimating the attributable fraction as it may appear from Table 6 because women's coefficients become closer at ages70-85 where deaths are heavily concentrated. And of course

lung cancer deaths are treated the same way in both methods. Indeed, Table 5 shows that the two sets of coefficients give relatively similar results for women. For men, the choice between the two series is basically immaterial because they are so similar to one another.

Discussion and Conclusion

Cigarette smoking is the single most important cause of premature death in the United States, accounting for more than one-fifth of all adult deaths. In the absence of deaths from smoking, U.S. life expectancy would be substantially increased. Given the importance of smoking as a modifiable risk factor, researchers have developed a number of methods attempting to calculate the excess mortality burden resulting from cigarette smoking. Cohort studies that track individuals with respect to smoking behavior and mortality provide informative evidence about the link between cigarette use and mortality, but suffer from imprecise classification of smoking status and ignore smoking status changes during follow-up. Moreover, detailed longitudinal studies of smoking that collect necessary smoking histories are unavailable for many relevant populations. To remedy some of these deficiencies, Peto et al. (1992) developed an indirect method which estimates the attributable fraction using the lung cancer death rate as a marker of smoking in developed-country populations. While some of their assumptions are rather arbitrary, the method has been widely applied in populations where cohort smoking histories do not exist.

In this paper, we applied an alternative indirect method to estimating mortality attributable to cigarette smoking in the United States. We calculated attributable fractions in 1990, 2000, and 2004 for the U.S. as a whole as well as the nine Census Divisions in order to examine geographic differences in smoking-related mortality. Our estimates indicate that

slightly more than 20% of all adult deaths in the United States in 2004 were caused by smoking, a figure which is highly consistent with previous estimates.

Geographic differences in smoking-attributable mortality should reflect historical differences in cigarette smoking. Very little smoking prevalence data is available at subnational levels of aggregation prior to 1985, but existing data match very closely to our patterns of estimated smoking-related mortality. Shopland et al. (1996) find the highest smoking prevalence for 1992-1993 in the East South Central, West South Central, and East North Central Divisions. The lowest prevalences are found in the Pacific, New England, and Mid-Atlantic Divisions. These regional patterns were also relatively constant between 1985 and 1992-1993 (Shopland et al., 1992). Other studies have connected state-specific cancer patterns to past prevalence of smoking (Jemal et al., 2006). Evidence from the past half century indicates that heavy smoking areas also experience high mortality from conditions commonly associated with smoking, especially lung cancer. As the smoking epidemic took hold among men in the southern states, lung cancer incidence and death rates increased rapidly there (Devesa et al., 1999, CDC, 2009). The results of our study further confirm the role of smoking in this long-term trend.

Since the 1980s, U.S. life expectancy has lagged substantially behind that of its European counterparts, and smoking has been a significant factor (Preston et al. 2009). In the coming decades, smoking-related mortality promises to increase in importance as quitting among women has been slower to occur than among men (Preston and Wang 2006). At the same time, countries experiencing more rapid declines in smoking among men may emerge as world longevity leaders. As these processes unfold, robust methods for estimating the mortality burden of cigarette smoking across a range of populations will be necessary. In countries with relatively mature smoking epidemics, where the vast majority of lung cancer cases are attributable to

smoking, our method provides reasonable and stable estimates of the impact of smoking on adult mortality. Within the United States, we find that our estimates of smoking-attributable mortality match very closely to historical smoking prevalence data, such as that used by CDC from the BRFSS. Our estimates of smoking attributable risk are also highly similar to those calculated by CDC (2009). We estimate slightly higher smoking-related mortality among women, and slightly lower among men, but preserve the regional pattern of the burden of smoking related mortality. The similarity of the two estimates despite considerable differences in the methods provides independent support for the results of each strategy and confirms the validity both approaches.

The coefficients estimated by Preston, Glei, and Wilmoth (PGW 2010) using international data since 1950 are very similar to those estimated using the current model for men, but substantially higher for women. While PGW coefficients may be appropriate for countries that are relatively new to the smoking epidemic, we believe our coefficients are more suitable in countries or regions with more mature and long-standing smoking epidemics. At the same time, the levels of smoking-attributable mortality implied by both sets of coefficients are similar.

Our study confirms the widespread impact of smoking on life expectancy in high-income countries (PGW 2010). In the United States, where declines in cigarette smoking in recent decades have been rapid, smoking-related mortality explains a substantial portion of geographic variation in life expectancy. Although mortality due to smoking has begun to decline among men, it remains the most important modifiable risk factor. The persistent geographic pattern of cigarette smoking within the US suggests that tobacco may preserve long-standing regional differences in mortality well into the future.

Table 1: Model Coefficients for lung cancer death rate by age and sex

Age	Male	Female
50-54	0.207	0.211
55-59	0.154	0.196
60-64	0.103	0.110
65-69	0.072	0.095
70-74	0.052	0.067
75-79	0.037	0.068
80-84	0.029	0.054

Estimated using Poisson regression in Equation (1). Includes controls and age interactions. The exponential of the above coefficients represents the proportional increase in the death rate for other causes associated with a one-per-thousand increase in the lung cancer death rate.

Table 2: Assumed lung cancer death rates of lifelong nonsmokers (per 1,000)

Age	Male	Female
50-54	0.06	0.06
55-59	0.05	0.07
60-64	0.12	0.12
65-69	0.22	0.17
70-74	0.35	0.31
75-79	0.52	0.33
80-84	0.89	0.58
85+	0.87	0.61

From Thun et al. (1997) for death rates of never smokers in the Cancer Prevention Study, Cohort II 1982 - 1988.

Table 3: Fraction of Deaths attributable to Cigarette Smoking by Division: 1990 - 2004

	Female			Male		
Division	1990	2000	2004	1990	2000	2004
New England	0.156	0.197	0.199	0.265	0.228	0.217
Mid-Atlantic	0.154	0.174	0.178	0.266	0.217	0.200
East North Central	0.154	0.185	0.195	0.284	0.243	0.237
West North Central	0.134	0.172	0.187	0.267	0.234	0.236
South Atlantic	0.160	0.188	0.190	0.305	0.251	0.237
East South Central	0.146	0.187	0.203	0.341	0.310	0.291
West South Central	0.157	0.181	0.187	0.315	0.256	0.239
Mountain	0.156	0.179	0.157	0.240	0.202	0.163
Pacific	0.175	0.182	0.180	0.244	0.189	0.180
Total US	**0.156**	**0.182**	**0.186**	**0.282**	**0.235**	**0.222**

Refers ages 50 - 84. Estimated using equations (2 - 4). Divisions categorized according to U.S. Census definitions.

Table 4: Expected years lived in ages 50-84 before and after the removal of smoking deaths: 1990 - 2004

	Females						Males					
	1990			2004			1990			2004		
	Without Smoking	With Smoking	Diff.	Without Smoking	With Smoking	Diff.	Without Smoking	With Smoking	Diff.	Without Smoking	With Smoking	Diff.
New England	26.81	25.77	1.0	27.81	26.70	1.1	24.49	22.26	2.2	25.95	24.38	1.6
Mid-Atlantic	26.44	25.40	1.0	27.45	26.45	1.0	24.20	21.91	2.3	25.51	24.01	1.5
East North Central	26.44	25.37	1.1	27.21	26.01	1.2	24.32	21.83	2.5	25.41	23.58	1.8
West North Central	26.95	26.04	0.9	27.65	26.54	1.1	24.60	22.27	2.3	25.90	24.13	1.8
South Atlantic	26.56	25.47	1.1	27.32	26.20	1.1	24.34	21.58	2.8	25.38	23.54	1.8
East South Central	26.06	24.97	1.1	26.65	25.28	1.4	23.93	20.61	3.3	24.89	22.39	2.5
West South Central	26.39	25.27	1.1	26.94	25.78	1.2	24.31	21.42	2.9	25.08	23.17	1.9
Mountain	26.52	25.48	1.0	27.57	26.72	0.8	24.02	21.98	2.0	25.72	24.56	1.2
Pacific	27.02	25.93	1.1	27.67	26.69	1.0	24.87	22.91	2.0	25.77	24.49	1.3
Total U.S.	26.57	25.51	1.1	27.35	26.25	1.1	24.37	21.89	2.5	25.48	23.79	1.7

Table 5: Mortality attributable to cigarette smoking in the U.S.: A comparison of estimates

	Females		Males	
	1990	2004	1990	2004
Current Model[1]	0.156	0.186	0.282	0.222
Preston, Glei, Wilmoth (2010)[2]	0.162	0.200	0.274	0.220
Peto-Lopez[3]	0.176	0.205[†]	0.277	0.243[†]
CDC Method[4]	0.140	0.153[*]	0.278	0.235[*]
Rogers[5]	—	0.126[†]	—	0.212[†]

[1]Coefficient estimates across 50 U.S. states, ages 50-84

[2] Estimates pertains to ages 50+ in 2003

[3]Ages 35+. Peto-Lopez estimates from (http://www.ctsu.ox.ac.uk/deathsfromsmoking)

[4]Estimates reported by the Centers for Disease Control and Prevention (2008), ages 35+

[5]Figures reported in Rogers et al. (2005) for the year 2000, ages 35+

[†] Estimates pertains to ages 35+ in 2000

[*] Estimates based on data for the period 2000-2004

Table 6: Model coefficients compared to Preston, Glei, and Wilmoth (2009)

	Fenelon and Preston[1]		PGW[2]	
Age	Male	Female	Male	Female
50-54	0.207	0.211	0.320	0.745
55-59	0.154	0.196	0.170	0.482
60-64	0.103	0.110	0.104	0.297
65-69	0.072	0.095	0.069	0.162
70-74	0.052	0.067	0.048	0.087
75-79	0.037	0.068	0.038	0.057
80-84	0.029	0.054	0.040	0.094

[1]Estimated using poisson regression in Equation (1).

[2]Estimated using negative binomial regression in Preston et al (2009)
Both models include controls and age interactions. The exponential of the above coefficients represents the proportional increase in the death rate for other causes associated with a one-per-thousand increase in the lung cancer death rate.

Figure 1: Comparability of attributable fractions based on our method and CDC across 50 states

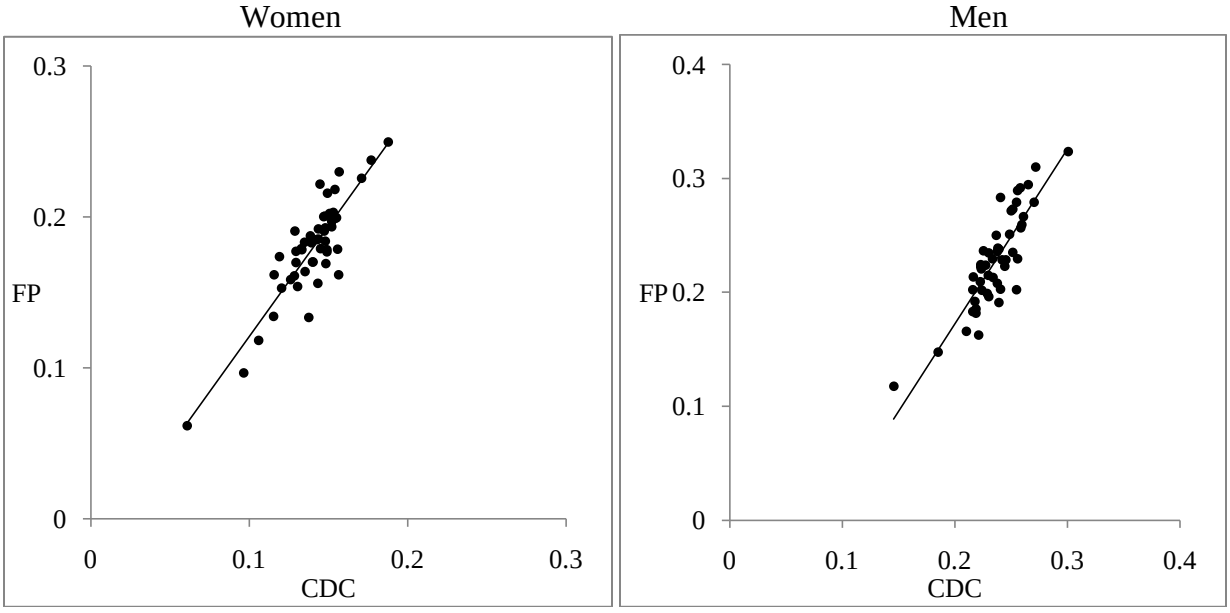

Figure 2: Coefficients for Lung Cancer Death Rates by Sex Based on the Current Model and Preston, Glei, and Wilmoth (2010)

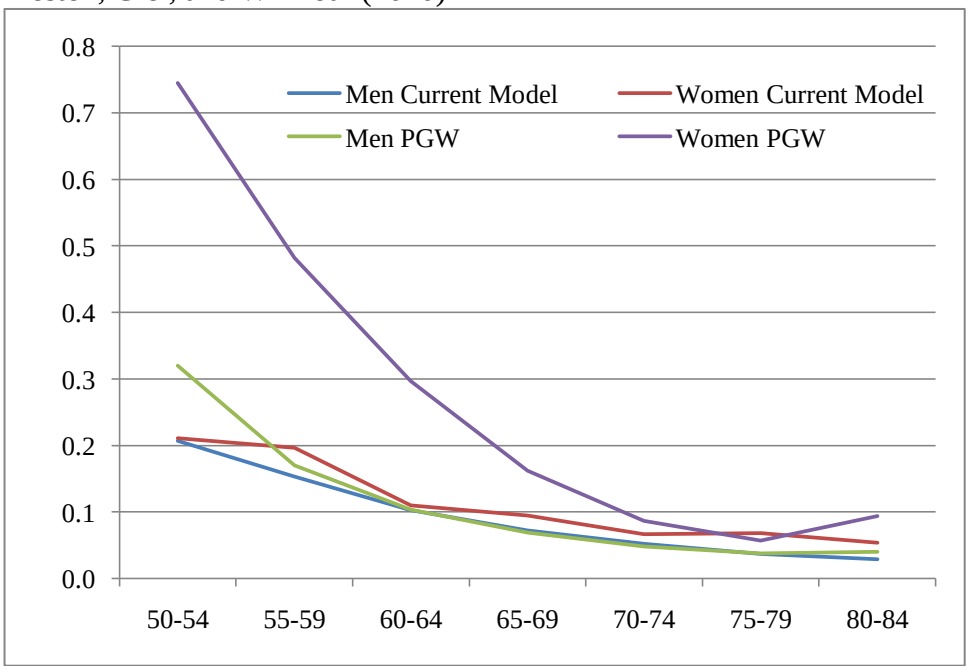

Table A1: Divisions and states

New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Mid-Atlantic	New Jersey, New York, Pennsylvania
East North Central	Illinois, Indiana, Michigan, Ohio, Wisconsin
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
South Atlantic	Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
East South Central	Alabama, Kentucky, Mississippi, Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, Texas
Mountain	Arizona, Colorado, Idaho, Montana, New Mexico, Utah, Wyoming
Pacific	Alaska, California, Hawaii, Nevada[1], Oregon, Washington

[1] Included with Pacific division as opposed to Mountain division since it shares more in common culturally and socially with the Pacific than with the Mountain states. This change does not alter the conclusions substantially.

References

ADHIKARI, B., KAHENDE, J., MALARCHER, A., HUSTEN, C. & ASMAN, K. (2009) State-Specific Smoking-Attributable Mortality and Years of Potential Life Lost-United States, 2000-2004 (Reprinted from MMWR, vol 58, pg 29-33, 2009). *Jama-Journal of the American Medical Association,* 301**,** 928-929.

CDC (2009) State-specific smoking-attributable mortality and years of potential life lost--United States, 2000-2004. *MMWR Morb Mortal Wkly Rep,* 58**,** 29-33.

DEVESA, S. S., GRAUMAN, D. J., BLOT, W. J. & FRAUMENI, J. F. (1999) Cancer surveillance series: Changing geographic patterns of lung cancer mortality in the United States, 1950 through 1994. *Journal of the National Cancer Institute,* 91**,** 1040-1050.

DOLL, R., PETO, R., BOREHAM, J. & SUTHERLAND, I. (2004) Mortality in relation to smoking: 50 years' observations on male British doctors. *British Medical Journal,* 328**,** 1519-1528.

FOREY, B., HAMLING, J., LEE, P., & WALD, N. (Eds.). (2002). *International smokingstatistics: A collection of historical data from 30 economically developed countries*. Oxford: Oxford University Press.

HILL, M. E., PRESTON, S. H. & ROSENWAIKE, I. (2000) Age reporting among white Americans aged 85+: Results of a record linkage study. *Demography*, 37**,** 175-186.

JEMAL, A., SIEGEL, R., WARD, E., MURRAY, T., XU, J. Q., SMIGAL, C. & THUN, M. J. (2006) Cancer statistics, 2006. *Ca-a Cancer Journal for Clinicians,* 56**,** 106-130.

JHA, P., PETO, R., ZATONSKI, W., BOREHAM, J., JARVIS, M. J. & LOPEZ, A. D. (2006) Social inequalities in male mortality, and in male mortality from smoking: indirect

estimation from national death rates in England and Wales, Poland, and North America. *Lancet,* 368**,** 367-370.

MOKDAD, A. H., MARKS, J. S., STROUP, D. F. & GERBERDING, J. L. (2004) Actual causes of death in the United States, 2000. *Jama-Journal of the American Medical Association,* 291**,** 1238-1245.

PETO, R., LOPEZ, A. D., BOREHAM, J., THUN, M. & HEATH, C. (1992) MORTALITY FROM TOBACCO IN DEVELOPED-COUNTRIES - INDIRECT ESTIMATION FROM NATIONAL VITAL-STATISTICS. *Lancet,* 339**,** 1268-1278.

PETO, R., LOPEZ, A. D., BOREHAM, J., THUN, M. & HEATH, C. (1994) *Mortality from Smoking in Developed Countries 1950-2000,* Oxford, Oxford University Press.

PETO, R., LOPEZ, A. D., BOREHAM, J., THUN, M. & HEATH, C. (2006) *Mortality from smoking in developed countries 1950-2000 (2nd edition),* Oxford, Oxford University Press.

PRESTON, S. H., GLEI, D. A. & WILMOTH, J. R. (2009) A new method for estimating smoking-attributable mortality in high-income countries. *International Journal of Epidemiology***,** 1-9.

PRESTON, S.. GLEI, D.A., AND WILMOTH, J.R. (2010) Contribution of Smoking to international differences in life expectancy. Forthcoming in Cohen, B., Crimmins, E., and Preston, S., editors. Divergent Trends in Life Expectancy. National Academies Press. Washington, D.C.

PRESTON, S. H. & WANG, H. D. (2006) Sex mortality differences in the United States: The role of cohort smoking patterns. *Demography,* 43**,** 631-646.

ROGERS, R. G., HUMMER, R. A., KRUEGER, P. M. & PAMPEL, F. C. (2005) Mortality attributable to cigarette smoking in the United States. *Population and Development Review,* 31**,** 259-+.

ROSENBAUM, W. L., STERLING, T. D. & WEINKAM, J. J. (1998) Use of multiple surveys to estimate mortality among never, current, and former smokers: Changes over a 20-year interval. *American Journal of Public Health,* 88**,** 1664-1668.

SCHNEIDER, K.L. and KATE L. LAPANE, (2007) Adult Health Behaviors *Public Health Rep.* 122(4): 432–434.

SHOPLAND, D. R., HARTMAN, A. M., GIBSON, J. T., MUELLER, M. D., KESSLER, L. G. & LYNN, W. R. (1996) Cigarette smoking among US adults by state and region: Estimates from the current population survey. *Journal of the National Cancer Institute,* 88**,** 1748-1758.

SHOPLAND, D. R., NIEMCRYK, S. J. & MARCONI, K. M. (1992) GEOGRAPHIC AND GENDER VARIATIONS IN TOTAL TOBACCO USE. *American Journal of Public Health,* 82**,** 103-106.

THUN, M., DAY-LALLY, C., MYERS, D. G., CALLE, E. E., FLANDERS, W. D. & ZHU, B. P. (1997) Trends in tobacco smoking and mortality from cigarette use in cancer prevention studies I (1959 through 1965) and II (1982 through 1988). IN BURNS, D., GARFINKEL, L. & SAMET, J. (Eds.) *Changes in cigarette-related disease risks and their implications for prevention and control, smoking and tobacco control. Smoking and tobacco control monograph no. 8, NIH publication no. 97-4213 (pp. 305-382).* Bethesda,

MD, Cancer Control and Population Sciences, National Cancer Institute, U.S. National Institutes of Health.

THUN, M. J., HANNAN, L. M., ADAMS-CAMPBELL, L. L., BOFFETTA, P., BURING, J. E., FESKANICH, D., FLANDERS, W. D., JEE, S. H., KATANODA, K., KOLONEL, L. N., LEE, I. M., MARUGAME, T., PALMER, J. R., RIBOLI, E., SOBUE, T., AVILA-TANG, E., WILKENS, L. R. & SAMET, J. M. (2008) Lung Cancer Occurrence in Never-Smokers: An Analysis of 13 Cohorts and 22 Cancer Registry Studies. *Plos Medicine,* 5**,** 1357-1371.