Published in final edited form as:

Risk Anal. 2012 July ; 32(Suppl 1): S39–S50. doi:10.1111/j.1539-6924.2011.01759.x.

Development of the Counterfactual Smoking Histories Used to Assess the Effects of Tobacco Control

Theodore R. Holford^{1,*} and Lauren D. Clarke²

¹Yale School of Public Health, New Haven, CT, USA

²Cornerstone Systems Northwest, Inc., Lynden, WA, USA

Abstract

Publication of the Surgeon General's Report in 1964 marshaled evidence of the harm to public health caused by cigarette smoking, including lung cancer mortality, and provided an impetus for introducing control programs. The purpose of this paper is to develop estimates of their effect on basic smoking exposure input parameters related to introduction of the Report. Fundamental inputs used to generate exposure to cigarettes are initiation and cessation rates for men and women, as well as the distribution of the number of cigarettes smoked per day. These fundamental quantities are presented for three scenarios: actual tobacco control in the US; no tobacco control in which the experience before 1955 was assumed to continue; and, complete tobacco control in which all smoking ceased following publication of the Report. These results were derived using data from National Health Interview Surveys, and they provide basic input parameters for the Smoking History Generator used by each of the lung cancer models developed by the Cancer Intervention and Surveillance Modeling Network (CISNET).

Keywords

lung cancer; smoking history; tobacco control

The Surgeon General's Report that appeared in 1964⁽¹⁾ presented a comprehensive summary of evidence on the harmful effects of cigarette smoking on human health, and its publication marks a watershed in forming a consensus in the public health community on the effects of a significant health hazard that needed to be controlled. However, cigarette smoking remains to this day as a significant risk factor affecting US public health.

Estimation of health effects from cigarette smoking requires detail on the history of exposure in a subject. This is summarized using initiation and cessation rates by birth cohorts, which are used in the Smoking History Generator, described in Chapter 5.⁽²⁾ It is more common to estimate the prevalence of smoking,^(3, 4) which depends on initiation and cessation. However, prevalence by itself does not provide sufficient detail to allow accurate estimates of risk for a disease with a long incubation period, such as lung cancer. The smoking history generator provides not only the prevalence of smokers, but estimates of the distribution of initiation, duration of smoking, duration of cessation among former smokers and dose.

We model two alternative extreme responses that might have resulted from reaction to scientific evidence on the causal link between cigarette smoking and lung cancer mortality. The first is no tobacco controls (NTC), in which we assume that smoking rates would have

*Address correspondence to Theodore R. Holford, 60 College St., New Haven, CT 06520, USA; theodore.holford@yale.edu.

remained as they were before the Report, subject to age trends within cohorts. The opposite extreme is complete tobacco control (CTC) in which conclusions drawn in the Report were widely embraced and cigarette smoking effectively ceased in 1965. Each of these scenarios requires accurate smoking history data for the population as a starting point, which, in turn, provide basic input parameters for the Smoking History Generator used by each of the lung cancer models developed by members of the Cancer Intervention and Surveillance Modeling Network (CISNET).

The first goal of these models is to characterize the impact of actual tobacco control (ATC) experienced in the US, which requires data giving what actually transpired. Once those trends have been characterized, we can consider alternative scenarios that might have resulted from different responses of the US population to scientific knowledge on the relationship between cigarette smoking and health. Actual smoking patterns have been developed using data from the National Health Interview Surveys (NHIS), which began collecting data on cigarette smoking in 1965, as described in Chapter 2. Therefore, the manner in which the population used cigarettes in the first two-thirds of the twentieth century, including detail like smoking initiation, cessation and dose, could only be constructed retrospectively. Challenges posed by the available NHIS data were (a) detail for ethnic subgroups; (b) reliable recall of events that occurred many years before; and (c) available data on only those who had survived at time of NHIS surveys. The strength of NHIS lies in its provision of a probability sample of the US population, but minorities tended to make up relatively small numbers in that sample. While it would be useful to separately assess the impact of smoking among different ethnic groups, sample size limitations make it difficult to obtain smoking history data in sufficient precision to allow this. We began the process of quantifying smoking history by first considering the largest ethnic group, whites, which is described in Chapter 2.⁽⁵⁾ After becoming satisfied that this group could be reasonably characterized, we extended the approach to the entire population in order to obtain estimates for the nation.

Smoking history, as defined by distributions of initiation, intensity and cessation was constructed for five-year birth cohorts. Models for lung cancer mortality that are discussed in this monograph seek to quantify the effect of smoking on individuals 30 to 84 years of age from 1975 through 2000. Figure 1 represents this block of time as a rectangular area with birth cohorts represented by diagonal lines with a slope of one on the coordinate grid with age on the vertical axis and calendar time on the horizontal. As can be seen in Figure 1, the earliest birth year represented in our block of time was 1890, i.e., individuals who would have turned 85 in 1975. Recall accuracy of events that may have occurred more than 60 years earlier, when these individuals were in their teens, must be a concern. Also affecting precision of NHIS survey results on smoking cessation were small sample sizes brought on by much lower cigarette smoking rates at the turn of the twentieth century, along with poorer survival among smokers. These limitations are particularly acute for females. Therefore, we decided to only consider smoking history data from the 1900 birth cohort onward. This left an upper left triangular corner of the rectangle (see Figure 1) without complete smoking history data which introduced an additional challenge in formulating smoking input parameters for these models.

In this chapter, we describe the method used to develop temporal models within the age-period-cohort modeling framework that provided accurate quantitative characterizations of the smoking history data. These temporal models are described for initiation and cessation rates. For the number of cigarettes smoked per day, we present and discuss the mean number of cigarettes by quintile. We also describe the manner in which these rates would be modified by the NTC and CTC scenarios.

1. METHODS

Development of smoking history scenarios corresponding to the population as a whole, as well as the alternative that might have been anticipated under alternative scenarios, required that we first develop estimates of functions describing elements of cigarette smoking exposure: initiation, cessation and dose.

1.1. Data Sources

Data on temporal trends in cigarette smoking behavior were obtained from NHIS, which are described in further detail by Anderson et al in Chapter 2.⁽⁵⁾ Individual exposure at a given time was considered to have been driven by smoking initiation and cessation and these are the outcomes that were included as inputs for the Smoking History Generator (SHG) discussed in Chapter 5.⁽²⁾

Smoking initiation rates were derived from the NHIS surveys which obtain data on age when an individual started smoking. Data used in this chapter were provided by Anderson et al and details on the approach used are presented further in Chapter 2,⁽⁵⁾ but we provide a synopsis here. The proportion of individuals who started smoking at a given age for a particular cohort provided an initial crude estimate of initiation. However, this would clearly be biased downward because of the effect of smoking on mortality as proportionately more smokers would have died before the survey. An adjustment factor was obtained using two estimates of the prevalence of ever smokers by age. The first calculates cumulative incidence rates using the crude proportion of initiators just described. The second used cross sectional estimates of prevalence, which decline with age as smokers die off more quickly than nonsmokers. A model is fitted to the cross sectional estimates and these may be extrapolated to age 30 for early cohorts which do not have younger subjects. The adjustment factor essentially inflates the first estimate of the initiation rates so that its cumulative incidence matches the second at age 30. The rationale for using age 30 is that the mortality difference for smokers would be very small up to this point.

The process begins with the age of smoking cessation, defined as abstinence for two or more years. For each year of age, the NHIS data were used to estimate the proportion that stopped smoking among those who were smokers at the onset of the year of age. No adjustment for mortality differences between current and former smokers was applied. However, there was good agreement between estimates of current smoking prevalence derived from these estimate and those obtained directly from the survey data which suggests that any adjustment would be small. Splines were used to smooth these estimates, as described in Chapter 2.⁽⁵⁾

Yearly initiation and cessation rates were used for five-year birth cohorts beginning with 1900–1904 and extending through 1980–1984. As already noted, cohorts represented in our range of ages 30–84 and years 1975–2000, include individuals who may have been born as early as 1890. Our decision to limit smoking history data from NHIS to data on cohorts born since 1900 required that we develop a reasonable assumption regarding the characteristics of those in the upper left triangle in Figure 1. It is especially important to understand changes in smoking cessation rates with age for these early cohorts because they would have had the longest span of years with limited knowledge from the scientific community about the harmful effects of cigarette smoking. Hence, their cessation behaviors are least likely to be affected by the publicity on the harms of smoking. Because cigarettes were initially used primarily by males, white males provided a sample size in NHIS for cohorts born before 1900 that could not be matched in terms of sample size by another ethnic group. Because so few females smoked in early cohorts, it was not also feasible to use these data. Thus, the

experience of white males offered an opportunity to investigate how changes in cessation rates vary with age for cohorts born around the turn of the twentieth century.

1.2. Initiation model

Age, period and cohort factors were considered for the analysis of initiation rates, i.e., the probability an individual began smoking in a given year as defined in Chapter 2. Traditional APC models were explored, but these were found to not provide an adequate description of the data. As an alternative to a complex model requiring interactions, we considered models for age trend within cohorts, in an attempt to not only explore variation in these age patterns by cohort, but to also discover cohorts that appear to have achieved a steady state for an important health behavior prior to widespread knowledge of its harmful effects.

Age trends for the initiation rate by cohort, $P_c(a)$, were represented by a cubic spline in age, a , for cohort c following transformation to the log scale, i.e., $\log P_c(a)$. To control for instability at extreme ages, we used the method described by Durrleman and Simon⁽⁶⁾ in which the trend at the youngest and oldest ages were forced to be linear. Selection of knots first considered equally spaced intervals. A subset of these knots was selected by backward elimination of covariates associated with a particular knot.

Regression coefficients corresponding to spline variables were estimated by ordinary least squares using PROC REG in SAS[®]. Observed initiation rates for individuals ages 10–69 years that were >0.2 percent were included in the analysis. Very small initiation rates become large negative values when the log transformation is used and these were excluded because the proportions were (a) not accurately determined, (b) had little effect on the population proportion of current smokers, and (c) overly influenced the model fit.

Accurate initiation data for cohorts born in the 19th century were not available from NHIS in part due to the age of this population when the surveys were conducted beginning in the mid 1960s. We assumed that initiation rates for birth cohorts 1890–1894 and 1895–1899 were 50% and 25% lower than the 1900–1904 birth cohort, respectively. The impact of such an arbitrary choice was considered to be small because (a) cigarette smoking rates are very low in these cohorts especially in women, and (b) although these are the oldest individuals with the highest mortality rates, they only affect the earliest years of this analysis (ages 76–84 in 1975, 77–85 in 1976, ..., 84 in 1884), (c) tobacco control efforts were at an early stage, and (d) benefits of tobacco control on lung cancer were only beginning to be realized.

For the three smoking scenarios considered in these analyses, the following initiation rates were needed for the SHG:

- Actual Tobacco Control—observed rates for the entire period
- No Tobacco Control—observed rates before 1955 and the fitted rates based on data before 1955 in subsequent years
- Complete Tobacco Control—observed rates before 1965 and 0 in subsequent years.

1.3. Cessation model

Smoking cessation patterns have also changed considerably during the twentieth century for a variety of societal reasons including tobacco control. Models describing cessation rates at age a for cohort c , $Q_c(a)$, were obtained by fitting spline models for age by birth cohort to the log rates, $\log Q_c(a)$. The only birth cohorts unaffected by growing awareness of health effects of tobacco prior to age 85 during the years of interest in this analysis, 1975–2000, would have been born before 1890. However, the numbers of individuals in this cohort who participated in NHIS, and had ever smoked were small. Therefore, in developing a model,

we first explored fitting the data for white males, because this was the largest group of ever smokers in the earliest birth cohorts interviewed. These rates were also compared to those derived for cohorts born before 1910 which were based on larger numbers, especially for women. An understanding of the relationships among these birth cohorts was considered to be important for generating plausible scenarios for a steady state in the US population if there were no tobacco controls.

For cessation rates under the three smoking scenarios, the following parameters were employed in the SHG:

- Actual Tobacco Control—observed rates for the entire period
- No Tobacco Control—fitted rates for the 1900–1904 birth cohort were used for all birth cohorts
- Complete Tobacco Control—not relevant because no smokers would have remained at the starting point of interest, 1975.

2. RESULTS

Yearly smoking initiation rates among selected birth cohorts for males and females are shown in Figures 2(a and b) respectively. The rise to a peak shows the modal age of initiation in the late teens and early twenties for those who do smoke at some time in their lifetime.

For males, the first cohort born in the century showed a lower peak than the later cohorts. Consequently, the remaining cohorts that began smoking before 1955 were used to infer the experience that would have been expected to continue for males in the absence of tobacco control.

Trends for females are considerably more complex than those observed for males in that they were affected by more complex changes in society, including (a) growing social acceptance of smoking by women, (b) fundamental changes in gender roles including the struggle for gender equality, and (c) effective incorporation of these societal changes into cigarette ad campaigns at just the time that sales to men began to sag in the 1960s. As seen in Figure 2(b), early twentieth century cohorts of women continued to take up smoking at considerably older ages than their male counterparts, so that the shape of the curve did not decline as quickly as seen in Figure 2(a). In subsequent cohorts, the shape of the smoking initiation curve is more similar to that of men, although the overall peak continued to grow up into the 1950s.

The comparison of males to females suggests that that for men, the initiation pattern appears to have stabilized for several birth cohorts just prior to the increased concern about the health effects of smoking, and these could be used to describe a steady state for the population in the absence of tobacco control. For women, the general age patterns appear to have stabilized, but the rates continued to grow overall up to that point, so that it became necessary to develop a different model from that used for men.

Age-specific smoking cessation rates for selected birth cohorts are shown in Figures 3(a and b). In both cases, the rates appear to increase steadily with age with no evidence of a downturn. Another aspect of this pattern is that the age at upturn occurs earlier in life, as for the more recent cohorts. This is likely to reflect awareness of harm from smoking occurred at younger ages with each advancing cohort, including those born at the turn of the century that would have been in their 50s when tobacco control began.

In order to develop models for smoking initiation and cessation rates that might have been expected in a population without a tobacco control program, it is important to understand the US experience for the generations leading into the mid 1950s. This is especially true for cessation rates, which tend to increase with age. However, if only cohorts born since 1900 are considered, there is only information about cessation rates up to about age 55, when they would then have begun to be exposed to tobacco control. Therefore, it was useful to explore cohorts born as early as 1880, although the size of the population was too small to yield accurate estimates, especially for women smokers. To develop some confidence in the shape of the curves in these early cohorts, we first explore the data for white males.

2.1. Models for White Males

To begin the process of developing a model for temporal trends in smoking unaffected by tobacco control, we considered the data for white men which represents a population that is somewhat homogeneous ethnically and the number who ever smoked among cohorts born in the late nineteenth and early twentieth century was large enough to provide more stable estimates of smoking cessation rates.

Figure 4 shows the estimated initiation rates for white men by year for individual birth cohorts. As already noted for all males, the peak rates occur for those in their late teens and early twenties, and this peak grows considerably for cohorts born in the first half of the century, declining for the more recent cohorts. While the decline appears to accelerate and occur at younger ages after 1964, these data suggest that this pattern actually began even earlier. Thus, data from the decade prior to publication of the Surgeon General's report were excluded from the analysis of trends in smoking behavior in the absence of any tobacco control, and only the years 1954 or earlier were used to estimate smoking initiation rates for the NTC scenario.

Curves obtained by fitting cubic splines with linear extremes⁽⁶⁾ (knots at ages 13, 17, 21, 25–29 and 33) to data for years before 1955, ages 10–39 are shown in Figure 4. The fit for the 1900–1904 birth cohorts is not as good as the subsequent cohorts which fit reasonable well over the years included in this analysis. These spline curves do not do as well at picking up the very sharp peaks in the observed data, which appear to shift somewhat among the cohorts. Nevertheless, they were felt to provide a plausible counterfactual to the pattern in 1955 or later. A comparison of the proportion who ever smoked using the observed and fitted initiation rates showed excellent agreement (not shown).

Figure 5 shows age-specific smoking cessation rates for white males by cohort. Of particular interest were patterns in cohorts born in the nineteenth century because these would not have been affected by scientific knowledge presented in the Surgeon General's Reports until they were 65 for the 1890 cohort and 75 for the 1880 cohort. As one sees in the figure, these earlier cohorts show a steady rise at progressively older ages, beginning around age 45 for the 1880–1884 and 1900–1904 cohorts, and 55 for the 1890–1894 cohort. The steepness of the increase is similar for the 1890–1894 and 1900–1904 cohorts, but it is somewhat shallower for 1880–1884. Estimated cessation rates for these early cohorts become increasingly imprecise. Therefore, the pattern of cessation rates for the 1900–1904 cohort appeared to provide a plausible pattern that may have resulted had there been no tobacco control.

NHIS includes information on the number of cigarettes smoked, which can in principle be used to construct estimates of dose for each cohort of smokers by year. However, these estimates become increasingly imprecise for smokers early in the twentieth century. Figure 6 presents estimates of mean cigarettes smoked per day for the third quintile of dose. These values are shown to be flat before the middle 1960s because the values derived directly from

the survey data were found to be too unreliable. More recently, the dose has declined, which may be a result of behavioral changes that took place since the Surgeon General's Report. To estimate dose under the assumption of no change since 1964, the constant values shown in the figure were used, and these are given for each quintile in Table I.

2.2. Models for All Males

Figure 2(a) shows smoking initiation rates for all males by year and alternate five year birth cohorts. The first birth cohort of the century, 1900–1904, is somewhat lower than others leading to 1954, so a separate line was fitted for this group. Because the trends are similar to those for whites, a cubic spline (knots at ages 13, 17, 21, 25–29 and 33) for the log initiation rate was fitted to the remaining data for cohorts with experience before 1955. These estimates were carried forward after 1954 as estimates of the likely experience of the population of US males if there had been no tobacco control (NTC).

Age-specific smoking cessation rates for males are shown in Figure 3(a). A log-linear model was used to estimate rates for the NTC case for the 1900–1904 birth cohort which were fitted to a cubic spline (knots at ages 15, 20, 25, 30, 35, and 55), and shown by the solid line in the figure.

Number of cigarettes smoked per day displayed a similar pattern to that observed for white males (not shown). Table I provides estimated values of the mean number per day found for years prior to 1964, and these were used for the NTC scenario.

2.3. Models for All Females

As in the models for males, age was represented by cubic splines in a model for the log initiation rates which was allowed to vary by cohort, and the cohort parameter was considered to be linear for those leading up to the end of the period before 1955. To find an optimal set of knots, we began with an over parameterized model with knots at every year for ages 11–31, then employed backward elimination to arrive at a final set of knots. Good fit for the age patterns was achieved by using knots at ages 15, 16 and 17. For the first four cohorts (1900–1904 through 1915–1919) the age parameters were allowed to vary, yielding fundamentally different shapes of the curves, but in the subsequent cohorts age parameters that were identical provided excellent fit, indicating that the log rates had essentially identical shape with respect to age. Cohort effects were included as nominal categories for the first 7 cohorts (1900–1904 through 1930–1934) and linear thereafter until 1955 when it was held constant. The smooth lines in Figure 2(b) shows the resulting fit to the observed data during the initial phase of this analysis. For the NTC scenario the fit corresponding to the cohort born in 1945–1949 was carried forward, which is also displayed in the figure.

Figure 3(b) shows the smoking cessation rates by age and birth cohort for females. As in the males, the 1900–1904 birth cohort was used to derive estimates for the NTC case, which were determined by fitting a cubic spline model to the log rates (knots at ages 12, 15, 17, 20, 22, 25, 30, 35 and 55).

Table I provides estimated values of the mean number cigarettes smoked per day found for years prior to 1964, and these were used for the NTC scenario. Females showed a similar pattern of decreasing the number of cigarettes smoked per day within each cohort (not shown), but the data were not found to be sufficiently precise before 1964 to allow one to determine accurate estimates of dose trends.

3. DISCUSSION

In this analysis we develop models that characterize observed smoking initiation rates, cessation rates and mean number of cigarettes smoked per day for men and women in the US prior to the introduction of tobacco control. Tobacco control policies and programs were introduced gradually and they evolved over time as more effective strategies became available.

While the primary focus of this monograph is the effect of controls that began around the time of publication of the 1964 Surgeon General's Report, we used the data prior to 1955 to obtain estimates of smoking behavior in a setting with little or no tobacco control. This was necessary because the data suggest that rates in men were already beginning to decline even before publication of the report. By carrying these results forward, we obtain input parameters for the NTC scenario for the SHG. Likewise, the CTC scenario employed the observed rates until 1965, when initiation rates were assumed to be zero and all current smokers were assumed to have quit. Note that articles had begun appearing in JAMA in 1950 and the British medical Journal in 1952 about the harms from cigarettes affecting lung cancer, and Reader's Digest had published an article entitled "Cancer by the Carton" in December 1952.

Cigarettes introduced a major risk factor for lung cancer into the lifestyle of the US population, and societal changes during the twentieth century played a huge role in their acceptance. This analysis demonstrates the major difference in trend for initiation and cessation that occurred for men and women brought on by factors that were only indirectly related to cigarettes themselves. However, effective marketing of cigarettes capitalized on these changes by introducing different strategies for men and women at critical junctures, e.g., distribution of free cigarettes to men drafted into the military in the Second World War and the concern for gender equality in basic rights that occurred subsequently. At the same time, the causal link between cigarette smoking and major diseases including lung cancer led to the development of control programs designed to discourage initiation and to increase cessation rates.

Smoking initiation rates for men appear to have essentially reached a steady state during the first half of the twentieth century. We have ascribed the changes in the second half to be largely due to control programs in the US, but from observational data one cannot be certain that other changes in what is considered to be fashionable behavior may not have also played a role. One effect that we do see occurring after publication of the Surgeon General's Report is a reduction in the peak initiation rates as curves become lower for succeeding cohorts. However, a second effect can also be seen by comparing the observed initiation rates with the values projected under the NTC scenario. The observed initiation rates fall off more quickly following publication of the Surgeon General's Report. Hence the overall decline in ever smoking among men is due to a combination of lower peak initiation rates and a shortening of the period when they were likely to begin smoking. More recent patterns among women are similar, but tobacco control programs were introduced at just the time when smoking among women was on the rise, so that two competing phenomena were in play that were influencing trends in opposing directions.

The data from the earliest birth cohorts available for this analysis suggest that smoking cessation rates increase with age even in the absence of tobacco control. Whether some individuals experience symptoms that they attribute to cigarette smoking and thus quit to relieve them is not known. In recent cohorts, we see both a steeper increase with age and an earlier age of cessation than was apparent among those in the earliest birth cohorts.

Available survey data on cigarette smoking in the US population has a significant limitation in not providing data for the upper left corners of our rectangle of interest, shown in Figure 1. To obtain model estimates for these rates, strong assumptions were required, and their rationale has been described. While other choices could have been made as to how to deal with the data limitations, these early cohorts clearly experienced much lower rates of cigarette smoking than those born more recently. Hence, far fewer estimated lung cancer deaths from smoking by these cohorts would still have been the result. In addition, the fact that this is a triangle shows that progressively fewer age groups are affected, so that the limitation is eliminated by 1985. Much of the impact of tobacco control would be expected to occur subsequently. Therefore, the impact of this limitation on the estimated effect of tobacco control is not likely to be large.

A second limitation in our underlying assumption for cigarette smoking exposure is the simplification in which individuals start smoking at a given point in time and then quit at a future date. The experience of individuals may be far more complex including individuals who are occasional smokers for a period of time, or they may go through several attempts at quitting before success. In addition, the level of smoking over a lifetime may fluctuate and not remain constant. Population surveys in which smoking history is retrospectively constructed do not yield this level of detail for individuals, so it was not possible to include this in our model. However, a similar criticism could be directed at the epidemiology studies used to estimate risk, in that exposure history is generally constructed from information gleaned from similar questionnaires with the corresponding inherent limitations.

The impact of changes in the distribution of exposure to cigarette smoking on lung cancer mortality is the subject of many of the later chapters in this monograph. Since increased cancer risk can remain for decades following exposure to a carcinogen, the effects of tobacco control may not be apparent until long after the program has been initiated. The results presented here indicate that initiation, cessation and dose have changed significantly since the period leading up to publication of the first Surgeon General's report on the health effects of smoking. But to quantify the health impact of these changes requires specific detail on progression of disease risk following exposure. Alternative carcinogenesis models for lung cancer have been considered, and many of these have been used by the groups contributing to this monograph. One expects that a benefit will have been realized from tobacco control, but not as great as might have occurred under more effective control programs. The impact of these exposure changes in the population will depend on the magnitude of effect of cigarette smoking and the lingering duration of that effect over time. Thus, quantification of effect is dependent on the carcinogenesis model and further details are developed in subsequent Chapters.^(7–12)

Acknowledgments

This work was conducted in collaboration with the Cancer Intervention and Surveillance Network (CISNET) and we are grateful for their insights and assistance with obtaining and analyzing data on smoking behavior over time. Funding was generously provided by a National Cancer Institute grant, CA97432. In addition, we thank David Levy, the editor for this manuscript, as well as referees for valuable comments on earlier drafts of this work.

References

1. United States Surgeon General's Advisory Committee on Smoking and Health. Smoking and health: Report of the advisory committee to the surgeon general of the public health service. Washington: U.S. Department of Health, Education, and Welfare, Public Health Service; U.S. Government Printing Office; 1964.
2. Jeon, J.; Meza, R.; Clarke, L.; Levy, D. Actual and counterfactual smoking prevalence rates in the US population via micro-simulation.

3. Harris JE. Cigarette smoking among successive birth cohorts of men and women in the United States during 1900–80. *Journal of the National Cancer Institute*. 1983; 71:473–9. [PubMed: 6577223]
4. Warner KE. Effects of the antismoking campaign: An update. *American Journal of Public Health*. 1989; 79 (2):144–51. [PubMed: 2913831]
5. Anderson, CM.; Burns, DM.; Dodd, KW.; Feuer, EJ. Risk Analysis. Birth cohort specific estimates of smoking behaviors for the U.S. Population.
6. Durrleman S, Simon R. Flexible regression models with cubic splines. *Statistics in Medicine*. 1989; 8:551–61. [PubMed: 2657958]
7. Schultz, F.; Boer, R.; de Koning, HJ. Erasmus mc lung cancer model.
8. Hazelton, W.; Jeon, J.; Meza, R.; Moolgavkar, S. FHCRC lung cancer model.
9. McMahon, P.; Kong, CY.; Johnson, BE.; Weinstein, MC.; Weeks, JC.; Tramontano, AC.; Cipriano, LE.; Bouzan, C.; Gazelle, S. The MGH ITA lung cancer policy model: Tobacco control versus screening.
10. Levy, D. PIRE lung cancer model.
11. Deng, L.; Spitz, M.; Gorlova, P.; Kimmel, M. Rice-MD Anderson lung cancer model.
12. Holford, TR.; Ebisu, K.; Mc Kay, L.; Oh, C.; Zheng, T. Yale lung cancer model.

Figure 1. Temporal relationships between ages and periods of interest, cohorts with available smoking history data, appearance of Surgeon General's (SG) Report, and years before tobacco control.

Figure 2. Smoking initiation rates for the US: Observed (dots and broken line), fitted (light solid line) and counterfactual source data (shaded region).

Figure 3.
Smoking cessation rates by age and selected birth cohorts.

Figure 4. Smoking initiation rates for US white males: Observed (dots and broken line), fitted (light solid line) and counterfactual source data (shaded region).

Figure 5.
Smoking cessation rates for US white males.

Figure 6.
Mean cigarettes per day for third quintile by year and cohort.

Table I

Mean number of cigarettes by quintile of smoking, gender and ethnicity.

Ethnicity	White		All
	Males	Males	Females
Quintile 1	8.7	7.3	4.8
Quintile 2	19.7	17.6	11.3
Quintile 3	20.5	20.0	18.9
Quintile 4	30.3	25.7	20.0
Quintile 5	45.0	42.2	33.6