HealthPartners Institute for Education and Research

ModelHealth™: Tobacco

Technical documentation

May 2015

Contents

ModelHealth: Tobacco

Executive summary

The HealthPartners Institute for Education and Research ModelHealth™: Tobacco is a state transition Markov model that estimates the cost-effectiveness of smoking cessation initiatives as well as the behavioral impact, health outcomes, and medical utilization impact of smoking policy. ModelHealth: Tobacco initially was developed with the intention of estimating the health impact and cost-effectiveness of both community interventions recommended in the *Guide to Community Preventive Services* and the clinical tobacco screening recommendation of the United States Preventive Services Task Force (USPSTF). The model employs a flexible framework in which the impact of the intervention under analysis is evaluated at the individual level. These individual effects are aggregated up to the population or community level.

This document presents a description of the model, an overview of its modeling framework, the development of its inputs, and a detailed discussion of the modeling framework and embedded algorithms.

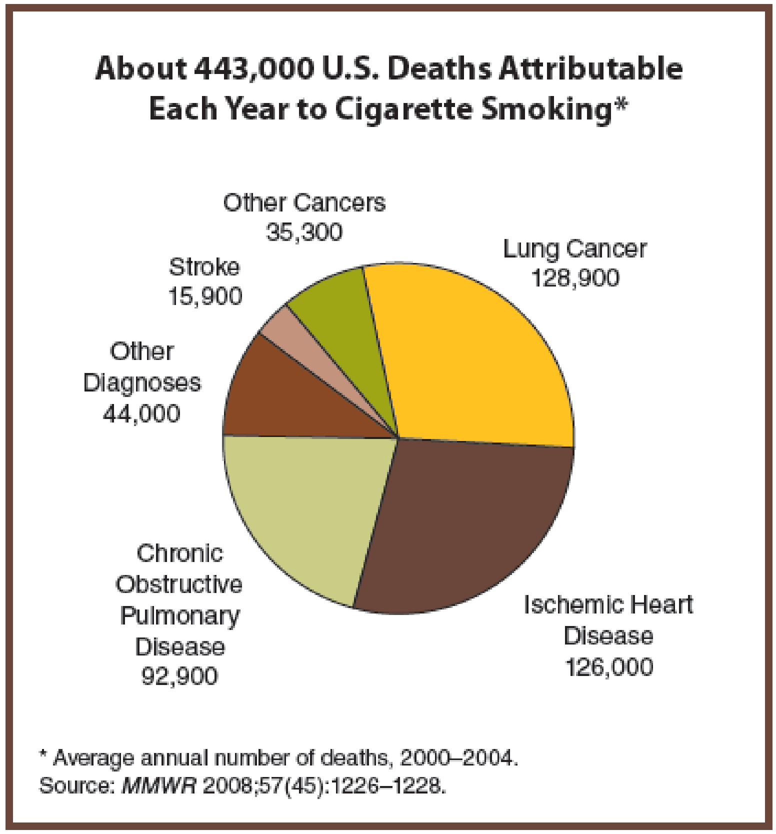
The work underlying ModelHealth: Tobacco consists of two parts: data and model. The model is based in TreeAge PRO 2015. The data underlying this model are the result of extensive literature searches, abstraction, and an adjudication process. An exhaustive discussion of that process is beyond the scope of this document. Here, we discuss key findings as they are incorporated as parameter values in the model. This approach determines key model items such as: disease risk, costs, and intervention effectiveness. Where adequate published estimates were unavailable, primary data analysis using large public-use datasets was used. These analyses determined smoking behavior such as initiation, cessation, quit attempts and relapse.

ModelHealth: Tobacco tracks smoking behavior of an individual and determines subsequent disease risks and health outcomes. By modeling at the individual level, the model is able to accommodate estimation of multiple "birth" cohorts with unique initial ages to provide the cross-sectional results presented in the Community Health Advisor website.

Introduction

An estimated 46 million U.S. adults use various forms of tobacco such as cigarettes, cigars, pipes and smokeless tobacco. Tobacco use continues to be the largest cause of preventable morbidity and mortality in the U.S. It is responsible for an estimated 450,000 deaths each year and nearly 8.5 million cases of serious illness. Tobacco-related diseases include cardiovascular disease, cancers of lung, larynx, esophagus and the oral cavity, as well as respiratory diseases. The effects of second-hand smoke are also serious, with more than 126 million non-smoking adults and children becoming exposed to toxins that may result in cancer, lung cancer, acute respiratory infections and asthma attacks. The economic burden of tobacco use is high, with $96 billion lost in medical expenses and an equal amount lost in productivity losses, according to CDC estimates.

Figure 1. U.S. deaths attributable to smoking annually



Both the United States Preventive Services Task (USPSTF) for and the *Guide to Community Preventive Services* (The Community Guide) have multiple recommendations regarding tobacco use with particular focus upon smoking. The current project has focused on analysis of related interventions.

ModelHealth: Tobacco was developed to evaluate the health impact and cost-effectiveness of implementing these recommendations targeting diverse populations, within different environments and jurisdictions, and over varying time-frames.

ModelHealth: Tobacco has been successfully used in support of multiple analyses, including state-level tax policy, USPSTF and Community Guide recommendations, cessation aid strategies, and mass media purchases.

USPSTF recommendations

The United State Preventive Services Task Force (USPSTF) recommendations include these three regarding tobacco use.

- The USPSTF strongly recommends clinicians screen all adults for tobacco use and provide tobacco cessation interventions for those who use tobacco products (*A Recommendation*).[1]
    - The USPSTF found good evidence that brief smoking cessation interventions, including screening, brief behavioral counseling (less than 3 minutes), and pharmacotherapy delivered in primary care settings, are effective in increasing the proportion of smokers who successfully quit smoking and remain abstinent after 1 year.
- The USPSTF strongly recommends that clinicians screen all pregnant women for tobacco use and provide augmented pregnancy-tailored counseling to those who smoke (*A Recommendation*).[1]
    - The USPSTF found good evidence that extended or augmented smoking cessation counseling (5-15 minutes) using messages and self-help materials tailored for pregnant smokers, compared with brief generic counseling interventions alone, substantially increases abstinence rates during pregnancy, and leads to increased birth weights. Although relapse rates are high in the post-partum period, the USPSTF concluded that reducing smoking during pregnancy is likely to have substantial health benefits for both the baby and the expectant mother. The USPSTF concluded that the benefits of smoking cessation counseling outweigh any potential harms.
- The USPSTF recommends that primary care clinicians provide interventions, including education or brief counseling, to prevent initiation of tobacco use among school-aged children and adolescents (*B Recommendation*).[2]

The USPSTF found adequate evidence that behavioral counseling interventions –such as face-to-face or phone interaction with a health care provider, print materials, and computer applications–can reduce the risk of smoking initiation in school-aged children and adolescents. The USPSTF found no evidence on the harms of behavioral interventions to prevent tobacco use; however, the magnitude of these potential harms is probably small to none. The USPSTF concludes with moderate certainty that primary care–relevant behavioral interventions to prevent tobacco use in school-aged children and adolescents have a moderate net benefit.

Following the established methods of the Prevention Priorities project and the National Commission on Prevention Priorities (NCPP), ModelHealth: Tobacco was used to estimate the cost-effectiveness of implementing the three USPSTF recommendations, which have an *A* or *B Recommendation.*

Community Guide recommendations

The *Guide to Community Preventive Services* (i.e. *The Community Guide*) performed systematic evidence reviews in the following areas: Reducing tobacco use initiation; increasing tobacco use cessation; reducing exposure to environmental tobacco smoke; restricting minors' access to tobacco products; and decreasing tobacco use among workers. *The Community Guide* encountered mixed evidence that is summarized by topic area below[3]:

- In the area of tobacco use initiation, *The Community Guide* found strong evidence of effectiveness and recommended the following interventions:
    - Policies that increase the unit price of tobacco products[4]
    - Mass media campaigns when combined with other interventions.[5]
- In the area of increasing tobacco cessation, *The Community Guide* found strong or sufficient evidence to recommend the following:
    - Increasing the unit price of tobacco products,[4]
    - Mass media campaigns when combined with other interventions[5]
    - Reduced out-of-pocket costs for cessation therapies[6]
    - Multicomponent interventions that include telephone support.[7,8]
- In the area of reducing exposure to environmental tobacco smoke, *The Community Guide* found strong evidence to recommend:
    - Smoking bans and restrictions.[9,10]
- In the area of restricting minors access to tobacco products, *The Community Guide* found strong evidence and recommended:
    - Community mobilization combined with additional interventions.[11]
- In the area of decreasing tobacco use among workers, *The Community Guide* found strong or sufficient evidence to recommend the following:
    - Smoke-free policies to reduce tobacco use[9]
    - Incentives and competitions when combined with additional interventions.[12]

Following the established methods of the Prevention Priorities project and the National Commission on Prevention Priorities (NCPP), ModelHealth: Tobacco was used to estimate the cost-effectiveness of those recommendation rated as having strong and/or sufficient evidence by *The Community Guide*.

State and local policies

Prior to development of the Community Health Advisor, ModelHealth: Tobacco had been used in two projects to assess the impact of several policies implemented at the state or local level. The two projects were:

- In partnership with ClearWay™ Minnesota, HPIER modeled an excise tax implemented statewide in Minnesota to increase the per-pack price of cigarettes. The analysis examined the impact of the policy on statewide tobacco smoking prevalence, as well as tobacco smoking prevalence in selected regions. Tobacco smoking prevalence, along with initiation and cessation, were forecasted over 10, 20, 30, and 50 year time-frames. Results were presented in terms of key demographic and employer groups.
- In partnership with *CDC's Office of Smoking and Health*, the impact of a per-pack excise tax, coupled with a mass media campaign funded by earmarked tax revenue. The analysis examined nationwide and local campaigns. Results illustrated the level of funding available, as well as the combined impact of pricing and revenue policies.

National policies

ModelHealth: Tobacco has been used to assess the impact of policies implemented at the national level. Examples of these policies are:

- An excise tax that resulted in a $.50/pack increase in the price of cigarettes. This "what-if" analysis was conducted to compare predictions of the model to other smoking tobacco models, particularly the model currently used by the Congressional Budget Office (CBO).
- Zero out-of-pocket costs for tobacco cessation therapy. This analysis was conducted to determine the effectiveness of a nationwide policy of zero out-of-pocket costs for smoking cessation therapy.
- CDC's *Tips From Former Smokers* (Tips) campaign. This analysis was conducted to explore the potential longer-term impact of the first nationwide mass media campaign.

Model description

Overview

ModelHealth: Tobacco is a Markovian individual-based model (i.e. Markov microsimulation). A Markov microsimulation is a model in which simulated individuals (sometimes called "agents") age over time, while facing period-specific 'risks' of changing health behaviors and/or health outcomes. In each cycle (currently, the equivalent of one year), individuals may remain in their current state or transition to a different one.

In the model, the state (age, smoking status, health, etc.) of each individual is tracked over time. Each state has associated costs and benefits. Each agent is uniquely defined by a set of heterogeneous characteristics. Three sources are used to determine how these characteristics are distributed across the simulated population: published estimates from peer-reviewed literature, survey datasets such as the National Health Interview Survey

(NHIS) and Behavioral Risk Factor Surveillance System (BRFSS), or publically available data such as the Campaign for Tobacco-Free Kids. The model is intended to conduct analysis at a variety of levels: national, state, and local. Thus, population parameters can be re-calibrated to sub-populations by changing the distributions of these characteristics.

When simulated agents are introduced into the model, they are assigned to a population strata, or cohort. Each cohort is defined by a unique combination of initial age, sex, ethnicity, and Census region. Each cohort is equally sized in the model. These demographic characteristics define the actions (e.g. the choice to begin smoking or not) of each individual. Possible actions are modeled in a framework that also identifies the consequences of these actions in terms of their resulting state.

Estimates from the model come from aggregating across agents the state-specific costs and benefits resulting from their individual actions. As noted in the prior paragraph, each agent is assigned to a population stratum at initiation. To create estimates specific to a location or region, these strata are weighted to specific areas of interest without requiring extensive reconfiguration of the model.

Agents may interact in the model, but allowing such interactions requires special configuration and adds considerable computational burden. Between-agent interactions may occur on two levels. First, at the macro level, the actions of each agent affect population-wide values. For instance, each smoker's cigarette consumption (cigarettes per day) contributes to total tobacco tax revenue. When aggregated across the entire population, this tax revenue determines the potential size of a mass-media campaign. This campaign, in turn, could alter smoking behavior and subsequent tax revenue in later years, and so on.

The second way agents are allowed to interact is at a micro-level. For instance, the smoking behavior of parents directly affects the likelihood a child will begin smoking. Similarly, peer groups could be defined where the behavior of one agent affects another.

The following section presents the general model structure. This is followed by a more detailed discussion of the two sub-models comprising the model. Then, some implementation details for ModelHealth: Tobacco are provided.

Model structure

ModelHealth: Tobacco starts by generating a population of simulated individuals, or agents. Members of the population are created by assigning individual characteristics (sex, race and ethnicity, lifetime educational attainment, census region, and age) proportional to that of the corresponding population being modeled. Once a simulated population is created, agents are aged through life in year increments, or cycles. Future births are introduced into the model at the start of each cycle. (How future births are introduced depends on the type of analysis being performed. When agent-level interactions are allowed, these are tied to existing simulated agents in the model. Otherwise birth-cohorts are introduced each cycle.)

The user is able to define the number and size of each population strata. For instance when assessing a USPSTF recommendation, a simulated birth cohort can be created by defining a set of population strata all with the initial age of 0. All agents in such an analysis will have the same initial age, but other demographics differ across strata. When assessing community or national recommendations, however, a cross-sectional population is needed. To create a dynamic cross-section, a range of initial ages that reflect the age range of the population being modeled is created. Some of these strata will have negative initial ages to indicate that stratum will be born sometime in the future. For instance, strata with an initial age of -5 represent a future birth cohort that will be born in year 5 of the simulation.

At initiation or birth, the basic demographics of age, sex, and ethnicity for each agent are determined by random draw with cohort-wide proportions equal to the population being modeled. Four broad racial/ethnic groups are currently included: black, Hispanic, white, and other.

An agent's lifetime educational achievement is determined based on the three basic demographic factors of age, sex, and ethnicity. Three broad levels of lifetime educational achievement are contained in the model: No high school diploma, high school, and bachelor's degree or higher. These broad categories were used because they are consistently defined and identifiable across the multiple data sources used to parameterize the model. The likelihood of agents attaining a certain level of education is set in proportion to published data from 2010 published data from the National Center for Educational Statistics (NCES) (Table 1). [13]

Table 1: Rates of lifetime educational attainment*

Age	Race	Sex	No high school or GED	Post-secondary	College Degree
15-18	White	Male	95.32%	4.54%	0.14%
		Female	93.02%	6.73%	0.25%
	Black	Male	39.71%	45.53%	14.77%
		Female	30.63%	49.96%	19.42%
	Hispanic	Male	36.66%	27.04%	36.30%
		Female	32.01%	30.46%	37.54%
	Other	Male	47.26%	21.22%	31.52%
		Female	57.32%	22.78%	19.90%
19-25	White	Male	96.54%	3.10%	0.36%
		Female	93.02%	6.69%	0.29%
	Black	Male	54.62%	39.32%	6.06%
		Female	44.50%	45.90%	9.60%
	Hispanic	Male	48.17%	31.13%	20.70%
		Female	41.83%	34.40%	23.77%
	Other	Male	63.11%	21.17%	15.73%
		Female	63.01%	21.79%	15.20%
26-64	White	Male	95.28%	4.55%	0.17%
		Female	94.35%	5.42%	0.23%
	Black	Male	63.86%	30.94%	5.20%
		Female	51.64%	39.82%	8.55%
	Hispanic	Male	68.29%	18.78%	12.93%
		Female	61.60%	22.19%	16.21%
	Other	Male	76.00%	11.38%	12.63%
		Female	79.94%	12.11%	7.95%
65+	White	Male	93.18%	6.17%	0.66%
		Female	91.19%	8.18%	0.64%

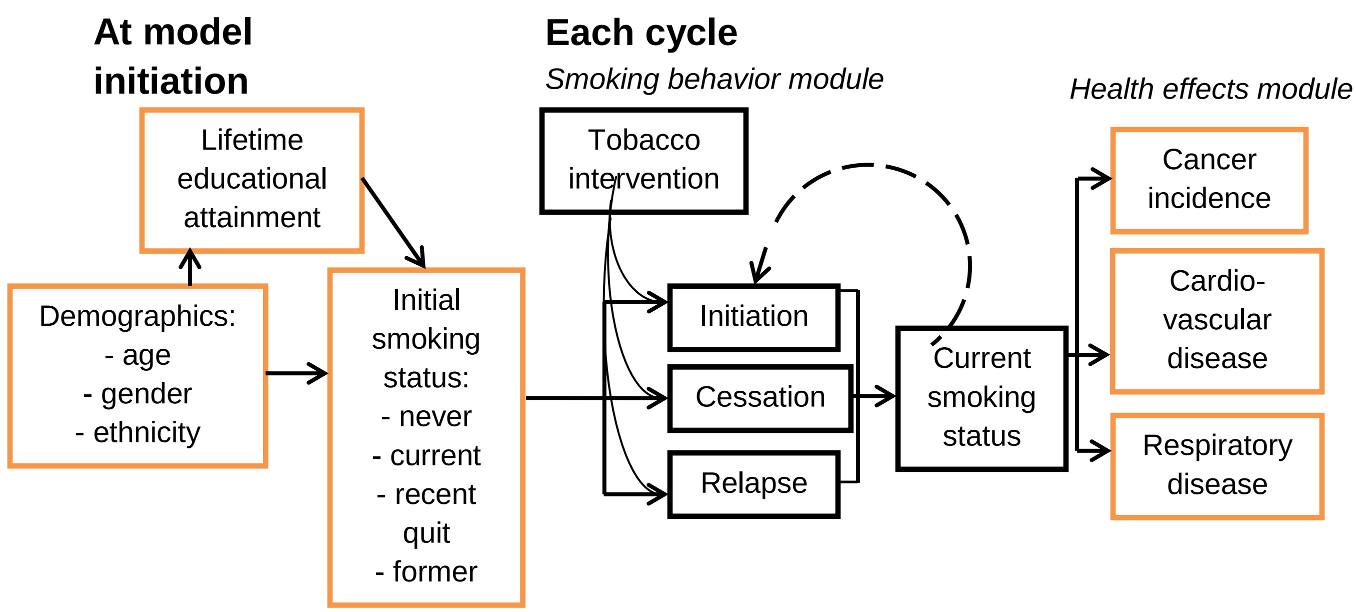
Black	Male	33.82%	45.79%	20.39%
	Female	27.77%	48.23%	24.00%
Hispanic	Male	30.80%	19.47%	49.73%
	Female	32.50%	20.04%	47.46%
Other	Male	45.52%	15.54%	38.94%
	Female	62.30%	14.14%	23.57%

**Source: National Center for Educational Statistics http://datainventory.ed.gov/*[13]

The model tracks smoking behavior over a lifetime and uses the smoking behavior module and the health effects module to determine the disease risk and health outcomes associated with that behavior.

During each cycle, tobacco and disease risk is evaluated. Figure 1 summarizes the flow of an individual agent within the model:

Figure 2. Structure of ModelHealth: Tobacco



The smoking behavior module

The smoking behavior module tracks smoking behavior over the lifetime of each simulated agent. The impact of the strategy, or intervention, under evaluation on smoking behavior is determined by comparing the smoking behavior of each agent under a “baseline” scenario (i.e. a world *without* the intervention/s) to that agent’s smoking behavior given a world with the intervention/s.

Population-wide effects (prevalence, tobacco consumption, initiation, and cessation) are determined by aggregating agent-level effects. For instance, *the Community Guide to Preventive Services* recommends increasing the unit price of tobacco products. The impact of that increase on rates of smoking initiation among 15- to 18-year-olds is determined by counting the number of agents who began smoking between ages 15-18 in the baseline scenario but *did not* begin smoking in that age range when the intervention was applied.

Individuals are in one of four smoking states: *never smoker*, *current smoker*, *recent quitter*, and *former smoker*. The probability that an individual is in a given smoking state is determined by two sets of multivariate risk equations that account for that individual's demographic profile. Specifically at model initiation, the likelihood that an agent is in any one of the four smoking states is conditioned on his/her age, gender, ethnicity, and – for those older than age 25 – the lifetime educational attainment at introduction into the model. Similarly, the likelihood that an agent who is currently in the never smoker state begins smoking within a given cycle is conditioned upon his/her age, gender, ethnicity, and – if older than age 25 – lifetime educational attainment.

Although the specific final multivariable risk equations vary in terms of covariates and dependent variables, there were several criteria that were consistent across analyses. The statistical relationships between each covariate and other predictors were screened prior to its inclusion in a final risk equation. If the inclusion of a covariate violated assumptions (e.g., co-linearity, normality, disproportionate cell size) appropriate adjustments (e.g., center around mean, transformation, re-categorization) were made or its inclusion reconsidered. Interaction terms (e.g. differential rates of initiation between young women and young men, differential rates of cessation between African-Americans with higher education and those without a high-school diploma, etc.) were considered based on the following criteria: representing at least 10% of the larger groups (e.g. at least 10% of women *and* at least 10% of those under the age of 18, at least 10% of African Americans within each educational category), and a coefficient significant at the 10% level.

Two different data sources were used to develop these risk equations. Estimation of risk equations corresponding to ages 18 and less used Youth Risk Behavior Survey (YRBS) data. Estimates of risk equations corresponding to ages 19 and older used data from the National Health Interview Survey (NHIS).

As noted, the first set of risk equations determine an agent's initial smoking status (i.e. smoking status at introduction into the model). This set of risk equations is used only at model initiation (i.e. year 0 in the model) and allows for analysis of birth cohorts and cross-sectional cohorts. The second set of risk equations determines smoking behavior (initiation, cessation, and relapse) over time and is used during every cycle in the model.

For both, two specifications for the risk equations were used to determine initial smoking status and smoking behavior. One, which excluded lifetime educational attainment, was used for ages younger than 25. The other, which included lifetime educational attainment, was used for ages older than 25. The cut point of age 25 for including lifetime educational attainment was chosen for two reasons. The first was conceptual; the second empirical. Our analysis intended no causal inference regarding the relationship between smoking behavior and educational attainment. Given our broad educational categories, most have achieved their final educational status by age 25. Prior to age 25, and particularly prior to age 18, there is arguably a reciprocal relationship between smoking status and educational attainment. Smoking status is likely a proxy measure for additional risky behaviors that, in turn, impact the likelihood of educational attainment. However, quantification of this simultaneous relationship was difficult and excluded from the current model.

After age 25, the relationship between educational attainment and tobacco use is better defined. For this age range, lifetime educational attainment can conceptually serve as a proxy measure of health literacy and engagement, which are factors that impact smoking behavior. We examined this conceptual conjecture by comparing equations with and without lifetime educational attainment for those older than age 25. When lifetime educational attainment was included in risk equations corresponding to ages greater than 25, sharper estimates of prevalence and behavior (i.e. initiation, cessation, and relapse) were obtained as measured by Mallow's C-statistic and the predictive value of the risk equations.

To support comparison to other widely reviewed models, an additional step was taken. The initiation and cessation rates in the model were calibrated to reflect those published by the Congressional Budget Office tobacco model. This process is discussed in a later section.

There is considerable variation in tobacco prevalence across the United States at the state and local level. For instance, according to the 2010 Behavioral Risk Factor Surveillance Survey (BRFSS), smoking prevalence at the county level ranged from a low of 0.8% to a high of 47%. To support analysis of specific regions, an additional analytic step is performed. Patterns of smoking initiation, or "initiation bands", were estimated. This analysis is discussed in a later section.

The final component of the smoking behavioral sub-module is determination of cigarette consumption among current smokers. The component of the model is only used for specific analyses and configurations that also allow agent-level interaction. Age, sex, ethnicity, and gender specific patterns of cigarette use were estimated using 2012 NHIS data.

Initial smoking status

A multinomial logistic regression with outcomes corresponding to the four smoking states was used to estimate the likelihood of an individual having an initial smoking status given his/her age, gender, ethnicity, and lifetime educational attainment. The estimated distribution across potential smoking states was then used to determine each agent's initial smoking status at introduction into the model.

During estimation of equations determining initial smoking state, the never smoker state was the comparison, or baseline, group. The multinomial logistic regression then modeled the likelihood of an individual initially being in another state. Neither the YRBS nor the NHIS directly ask respondents about their current smoking status. Instead, the following definitions were used to create our model states:

- **Never smoker:** Having smoked *fewer* than 100 cigarettes in their lifetime
- **Current smoker:** Having smoked *at least* 100 cigarettes in their lifetime and having smoked in the last week

- **Recent quitter:** Having smoked *at least* 100 cigarettes in their lifetime and having quit for *less than* 4 years
- **Former smoker:** Having smoked *at least* 100 cigarettes in their lifetime and having quit for *4 or more years*

The usual definitional prerequisite of having smoked at least 100 cigarettes in their lifetime was applied to screen out experimental smoking. Recent quitters were distinguished from former smokers for purposes of disease risk as discussed in the section: The Health Impact Module. Table 2 contains the results from estimation of the final multinomial model.

Table 2: Results of multinomial estimation predicting initial smoking status

	Current Smoker	*95% Conf Interval*	**Former Smoker**	*95% Conf Interval*
Ref. Category*	-0.798	*(-0.874 , -0.722)*	-1.922	*(-2.029 , -1.816)*
Female	-0.453	*(-0.495 , -0.411)*	-0.605	*(-0.646 , -0.564)*
24-44	0.559	*(0.482 , 0.635)*	1.151	*(1.039 , 1.263)*
45-64	0.541	*(0.462 , 0.621)*	1.813	*(1.702 , 1.925)*
65+	-0.538	*(-0.632 , -0.443)*	2.203	*(2.090 , 2.315)*
Black	-0.475	*(-0.535 , -0.416)*	-0.714	*(-0.779 , -0.648)*
Hispanic	-1.249	*(-1.322 , -1.176)*	-0.723	*(-0.788 , -0.659)*
Other	-0.702	*(-0.799 , -0.604)*	-0.793	*(-0.893 , -0.694)*
High School	0.688	*(0.634 , 0.741)*	0.112	*(0.054 , 0.169)*
Post-Secondary	-1.293	*(-1.356 , -1.230)*	-0.394	*(-0.442 , -0.346)*

**Reference Category is Young, White, and Male, with no HS education*

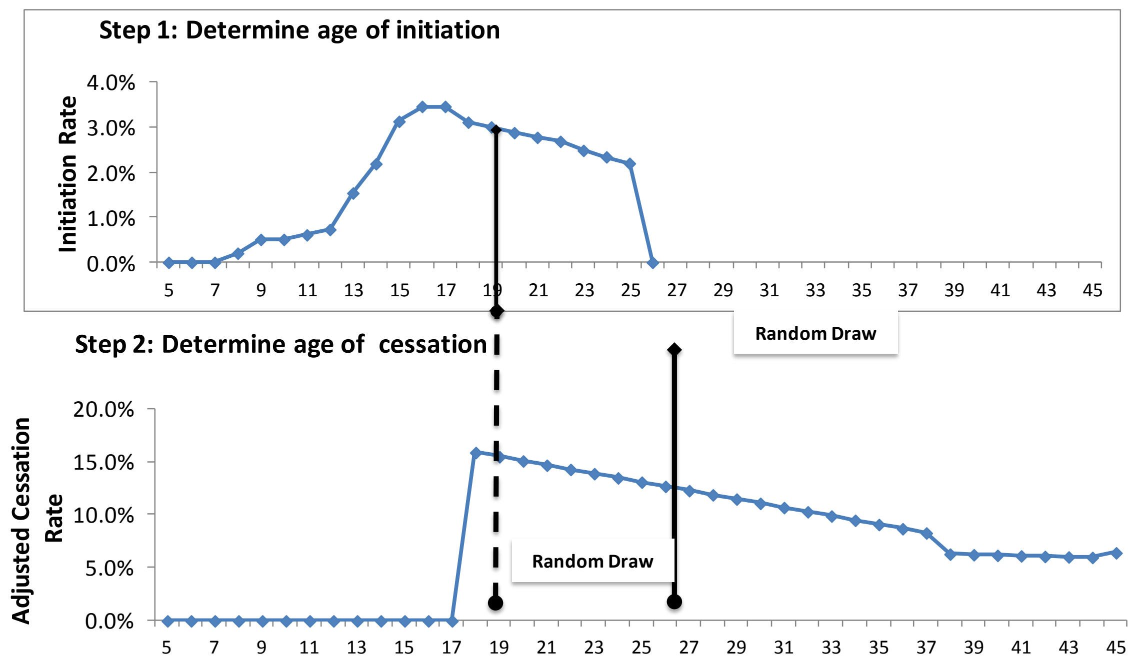
Initializing agents into current, former smoker state

Within the behavioral model, "time in state" (i.e. the number of model cycles spent within the smoking state and/or the number of cycles spent as a former smoker) partially determined the likelihood of quitting or relapsing. When performing a cross-sectional analysis, an age of smoking initiation is assigned to all agents initialized as either current smokers, recent quitters, or former smokers. In addition for those initialized in the recent quitter or former smoker state, an age of cessation is also assigned.

These ages of initiation and/or cessation are determined by random draws from distributions configured to estimated patterns of initiation and cessation. The estimation of these patterns is discussed in a later section. Here, the method of determining age of initiation and cessation for agents initialized into one of these three states is described.

The process is illustrated in Figure 2, which depicts the process for an agent initialized into the model as a 45-year old former smoker. First, in Step 1 for all agents initialized in any of these three states (current, recent quitter, former smoker), a random draw from a distribution drawn configured to initiation rates estimated from the NHIS determines the age at which the person first started smoking (age 19). Then, for those initialized as recent quitters and former smokers (Step 2), a random draw from a second distribution configured to cessation rates estimated from NHIS and truncated at the age of initiation determines the age of cessation (age 26). These two ages are then used to determine the time spent smoking and time since cessation, which are used in by the model when determining future smoking behavior.

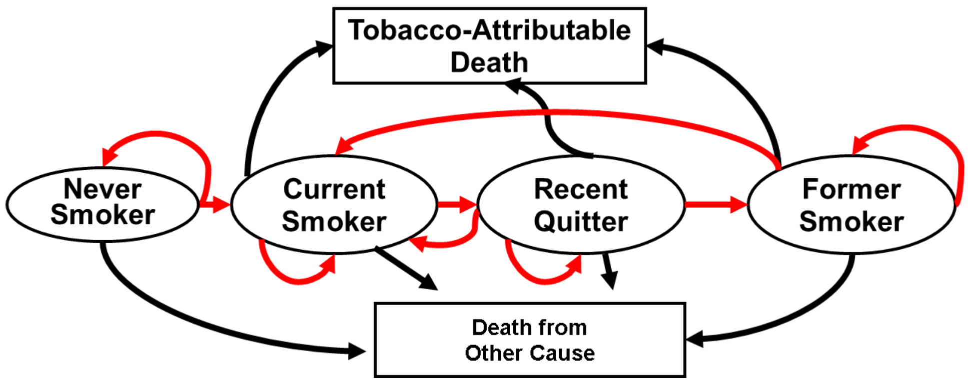
Figure 3: Determination of age of initiation and cessation



Lifetime smoking behavior

An individual's "risk" of changing smoking status (i.e. *transitioning to another smoking state*), is determined by current state, time in that state, and demographics. Individuals who have never smoked can either remain in the never smoker state or begin smoking and transition to the current smoker state. A current smoker who is in the current smoker state can remain or quit and transition to the recent quitter state. A recent quitter either remains in the recent quitter state, relapses into the current smoker state, or moves to the former smoker state once four years have passed. A former smoker either relapses into the current smoker state or remains in the former smoker state. In addition, all individuals are faced with a risk of dying of either a tobacco-related illness or some other cause. Figure 4 illustrates this conceptual framework of the natural history of smoking tobacco use.

Figure 4: Natural History of Smoking Tobacco Use



Using the aforementioned definitions of smoking status, we identified new smokers (Initiators) from among current smokers as those indicating that they had begun smoking within the last 12 months.

Three separate logistic regressions determined the risk of smoking initiation by comparing Initiators to Never Smokers. The first, which used YBRS data, applied to ages younger than 18. The second and third, which used NHIS data, applied to ages 18-24 and 25 and older, respectively. Similar to the initial smoking status risk equations, the 19-24 specification was distinguished by inclusion of lifetime educational achievement. Tables 3 and 4 contain the results of these estimations.

Table 3: Youth tobacco smoking initiation rates*

Age	Male	Female
8	0.002	0.006
9-10	0.005	0.006
11-12	0.010	0.013
13-14	0.022	0.021
15-16	0.027	0.027
17-18	0.010	0.013

**From YBRS data*

Table 4: Results of logistic regressions predicting adult smoking

	Tobacco Initiation	*95% Conf Interval*	**Tobacco Cessation**	*95% Conf Interval*
Ref. Category	-27.7099	*(-33.273 , -22.146)*	-1.772	*(-2.133 , -1.411)*
Female	3.5358	*(4.351 , 2.721)*	-0.046	*(-0.053 , -0.039)*
24-44	9.814	*(12.472 , 7.156)*	-0.1545	*(-0.179 , -0.130)*
xFemale	-10.0481	*(-12.656 , -7.440)*	-0.00165	*(-0.002 , -0.001)*
45-64	10.441	*(12.846 , 8.036)*	-0.1181	*(-0.139 , -0.098)*
xFemale	-5.817	*(-7.292 , -4.342)*	0.2346	*(0.294 , 0.175)*
White	-6.3501	*(-7.745 , -4.955)*	0.2966	*(0.369 , 0.224)*

xFemale	-3.8882	*(-4.893 , -2.884)*	*Not Significant*	
Black	3.4254	*(4.151 , 2.700)*	-0.0603	*(-0.073 , -0.048)*
xFemale	-3.4627	*(-4.426 , -2.499)*	*Not Significant*	
Hispanic	5.0037	*(6.435 , 3.572)*	0.0776	*(0.094 , 0.062)*
xFemale	-0.0798	*(-0.096 , -0.063)*	*Not Significant*	
No High School	6.5959	*(8.319 , 4.872)*	-0.00755	*(-0.009 , -0.006)*
xFemale	-3.8882	*(-4.791 , -2.986)*	*Not Significant*	
High School	9.2186	*(11.708 , 6.729)*	0.0191	*(0.022 , 0.016)*
xFemale	-3.4627	*(-4.365 , -2.561)*	*Not Significant*	
Post-Secondary	4.5348	*(5.593 , 3.477)*	0.3067	*(0.384 , 0.230)*
xFemale	-0.0798	*(-0.096 , -0.064)*	*Not Significant*	

**Reference Category is Young Mixed Race, Male*

We assumed no cessation among youth younger than age 18 and estimated two cessation risk equations for adults. From the NHIS data, we identified Quitters as those indicating they had ceased cigarette use within the last 12 months with no indication of relapse. Two logistic regressions (18-24 and 25 and older) compared Quitters to Current Smokers to determine the likelihood of smoking cessation. Once again, the 19-24 specification was distinguished by inclusion of lifetime educational achievement.

Relapse after quitting tobacco use is time-sensitive. The longer a person has successfully quit smoking, the less likely they are to relapse. The cross-sectional design of both the YBRS and NHIS surveys made estimation of relapse rates that account for time since cessation difficult. Instead, we used published estimates based upon longitudinal studies. These values were then adjusted during calibration provide reasonable values of age, sex, and ethnic specific tobacco use rates.

Table 5: Baseline Smoking Tobacco Relapse Rates

Years Since Successful Quit	***Probability of Relapse***	***Source***
1	0.37	Hughes 2003
2	0.08	Wetter 2004
3	0.08	Wetter 2004
4	0.08	Wetter 2004
5	0.08	Wetter 2004
6	0.038	Wetter 2004
7	0.038	Wetter 2004
8	0.021	Wetter 2004
9	0.021	Wetter 2004
10	0.021	Wetter 2004
11	0.005	Wetter 2004

Recalibration of tobacco model to CBO model

To facilitate comparison, ModelHealth: Tobacco was calibrated to reflect baseline tobacco use projections of the Congressional Budget Office (CBO). These calibrated initiation and cessation rates are used as the baseline in the current model.

The CBO does not provide detail on how its tobacco baseline is parameterized. We therefore worked with Figure 1-1 in the 2010 CBO report as our guide. CBO only reports its projection of smoking prevalence among all adults in Figure 1-1. Our model determines annual smoking prevalence based upon initiation, cessation and relapse, as mediated by sex, age, race-ethnicity and educational attainment. The average adult smoking prevalence reported in Figure 1-1 could be reproduced with infinite combinations of smoking initiation, cessation, and relapse rates among males and females of different ages, race-ethnicities, and educational attainment. In addition, predictions of smoking prevalence among adults depend heavily on recent, current and near-term teen smoking initiation rates. Therefore, with only Figure 1-1 and a general description of the CBO's approach as a guide, we tested a reasonable set of parameter modifications to adjust the smoking prevalence rates produced by our model over the next 10 years to better reflect CBO's baseline.

Summary of calibration steps:

The following summarizes attempts and modifications to the model to calibrate to the long-term trend incorporated into the CBO tobacco tax model. The final results are the new baseline results included in all CDC tobacco estimates.

- **Step 1:** Recalibrate relative Education and Ethnicity effects within the current model
    - **NOTE:** CBO provides only population-based temporal trends. To ease calibration, the following approach was used:
        - Create and Age/Sex initiation and cessation risk table that will be used to calibrate to CBO estimates
        - For each calibration attempt, re-apply demographic differences preserving relative differences across ethnic groups as originally estimated from the NHIS dataset.
        - Education and ethnicity effects were then re-applied to initiation and cessation rates post calibration.

- **Step 2:** Compare new ModelHealth: Tobacco baselines to CBO baselines and adjust accordingly. Here, the following additional model adjustments and reparameterizations were explored to produce overall baseline tobacco use that is more similar to that shown in Figure 1-1 of the CBO report. The following outlines the different changes to the HPIER model, identified **sources** of deviation from the CBO model, and adjustments implemented along with the effect.
    - **Baseline (Initial) attempt:**
        - Population Tobacco Prevalence was too low
        - **Adjustments**
            - Added 8 yr old initiation
            - Reduced cessation for ages>50 (prior models used TreeAge Truncation option in the cessation table, TBL_TobaccoRisk)
            - Increased 9-10 yr old initiation

        - Increased 15-18 initiation
    - **Impact**
        - Elevated prevalence in Youth and Young Adults
        - Elevated prevalence in Adults

- **Source1:** Initial Tobacco Prevalence table contained “jumps” in the prevalence because TreeAge was set to use truncation. This caused “bunching” by certain ages rather than a smooth change in prevalence.
    - **Initial Resolution Attempt:** TreeAge was set to use interpolation
        - **Impact:** Smoothed smoking prevalence across all ages 0-100
    - **Final Resolution:** Interpolation outside of the model was used to fill missing age ranges
        - **Impact:** After allow interpolation in TreeAge, more realistic initial prevalence was created. This better captures estimates of age-based relative smoking mortality
            - ***Note:*** *Initial prevalence was fixed after age 70 at 5%*

- **Source 2:** Tobacco relapse was elevated causing too much “churn” in initial populations due to only current and former smokers in initial population
    - **Initial Resolution Attempt:** Relapse smoothed over initial 10-yr range
        - **Impact:** improved relapse estimates, but there were too many relapses among older age groups (Aged 50 or older)
    - **Final Resolution:** The rates of relapse were reduced and smoothed over each entire 10-year age range

- **Source 3:** Inconsistent initiation patterns for eight- to 12-year-olds. The initial estimates were based on small samples and literature-based estimates that grouped estimates by two-year age groups (8-10, 11-12).
    - **Initial Resolution Attempt:** Baseline initiation rates were reduced.
        - **Impact:** The initiation pattern for the 8-12 year old age group improved but provided elevated prevalence at age 13.
    - **Final Resolution:** Initiation rates were calibrated to arrive at observed prevalence at age 13, which was estimated from a larger, more-representative YRBF sample. The calibration assumes increasing initiation rates using a relative rate (RR) increase of 30% per year. This RR increase was estimated assuming a linear growth path in initiation over the 8-12 year old age range.

- **Source 4:** Unique initiation rates for ages 15-17 were missing from abstracted articles (combined age range) and use of the truncation option within TreeAge created an “average initiation” across this entire age-range.
    - **Initial Resolution Attempt:** Interpolation in TreeAge implemented
        - **Impact:** Initiation across all age ranges was forced to 0.

    - **Final Resolution:** Initiation rates were calibrated in a manner similar to that done for the 8-12 year old age. A 30% increase in the initiation rate per age arrived at the observed 18 year old prevalence.
        - **Impact:** Initiation and prevalence for new birth cohorts reflected the observed prevalence of current 18 year old cohorts.

- **Source 5: (*Key source of deviation from CBO model*)** New birth cohorts had smoking prevalence at age 18 similar to that of current 18 year olds but 3-5% higher than the prevalence forecasted by the CB0 model. *This increased prevalence was consistent with (2010) NHIS data, but not consistent with the CBO model that shows decreasing prevalence over time*.
    - **Initial and Final Resolution Attempt:** Decrease initiation rates across ages ranges using 10 yr moving average (MA) process
        - **Impact:** Lowered prevalence among new birth cohorts that resulted in a new "steady-state" population prevalence of approximately 13-14%.
        - ***NOTE:*** *This 13-14% steady state prevalence assumes demographics (sex and ethnicity) approximate to the NHIS representative sample and that may differ when weighted to population under examination.*

- **Source 6: (*Key source of deviation from CBO model*):** Estimated initiation patterns from NHIS used age-based categories that created and stepped function and subsequent "jagged" patterns of initiation. This created elevated patterns of initiation for ages up to 24.
    - **Initial and Final Resolution Attempt:** Smoothed initiation rates using a moving average process across ages ranges holding implied prevalence at end date (age 24 within birth cohort) constant
        - **Impact:** Removed "jumps" in prevalence among birth cohorts, but initiation remained relatively high

- **Source 8: (KEY *Key source of deviation from CBO model*):** Adjusted initiation among 13-17 year olds to reflect baseline patterns of prevalence at model initiation
        - **Impact:** Prevalence among adolescents initially declines and then stabilizes in a manner similar to that implied by the CBO model.

The results of the recalibration are shown in the figures below.

Figure 5: Prevalence of Final Calibration by Age Group

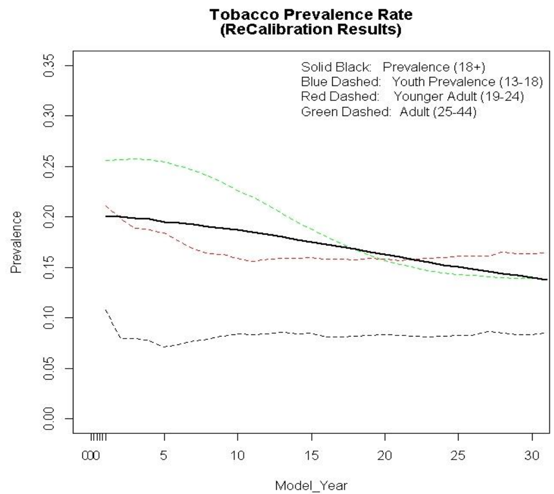


Figure 6: Initiation Rates by Age Group

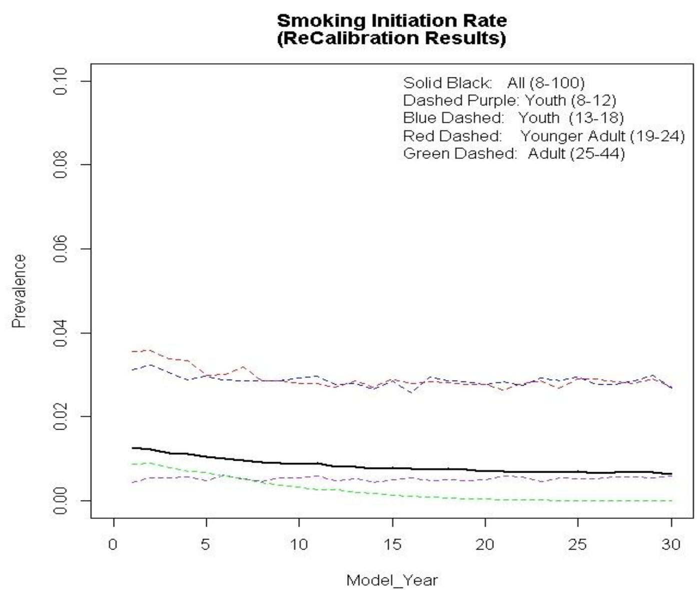
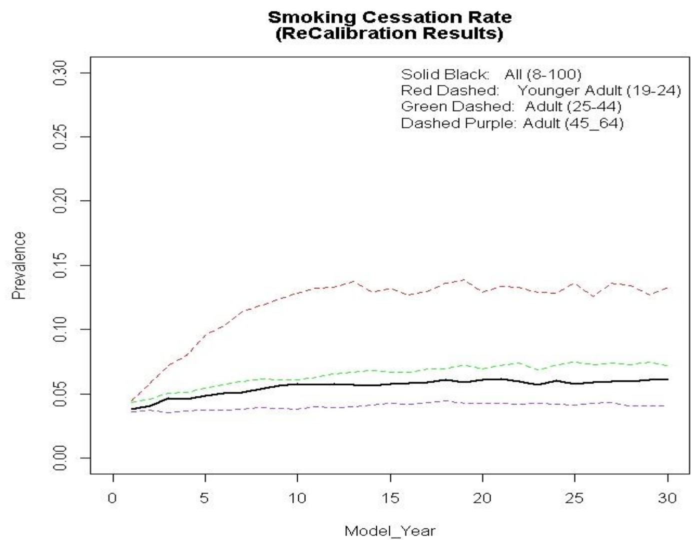


Figure 7: Cessation Rates by Age Group



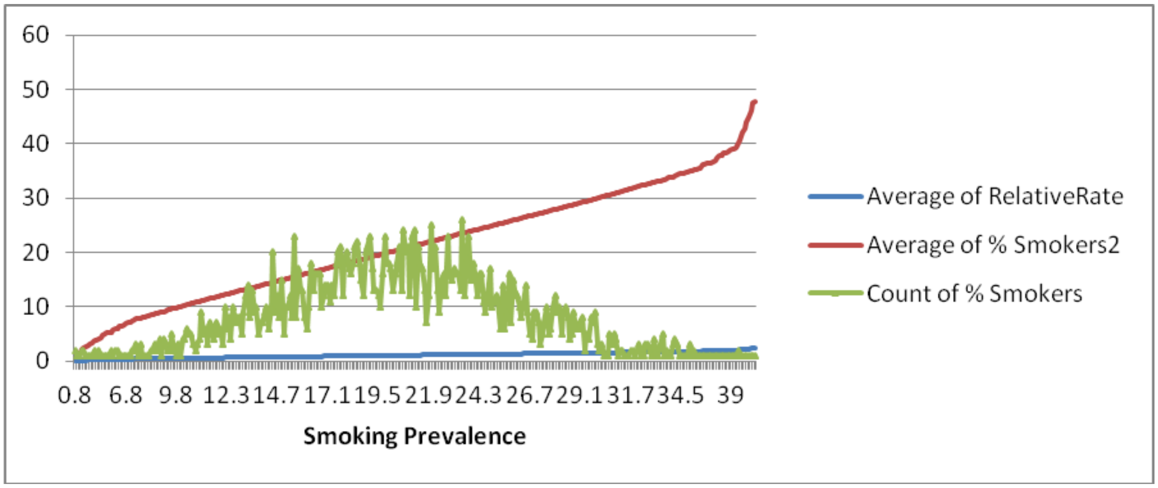
Identification and specification of smoking initiation bands

There is considerable variation in tobacco prevalence across U.S. regions. Using a classification and regression tree (CART) analysis, bands were identified using County Health Rankings data, which is based on the 2012 behavioral risk factor surveillance survey (BRFSS). These county-level data were used to create evenly spaced bands of smoking prevalence with different patterns of initiation identified to arrive at the observed prevalence within each band.

The approach to identification and definition of the initiation bands is as follows:

1. All counties with data within the BRFSS were grouped into categories of tobacco prevalence using a tree-classification algorithm[14,15]
2. Determine each grouping's mean and median prevalence relative to the overall U.S. prevalence:
    a. Relative Prevalence$_{\text{County i}}$ = Prevalance$_{\textit{County I}}$ / U.S. prevalence
    b. Calculate descriptive statistics of county prevalence
    c. ***NOTE 1:*** *Groupings were determined according to ranges of prevalence not deciles of the county distribution (See Figure 8).*

Figure 8: Initiation rates by age group



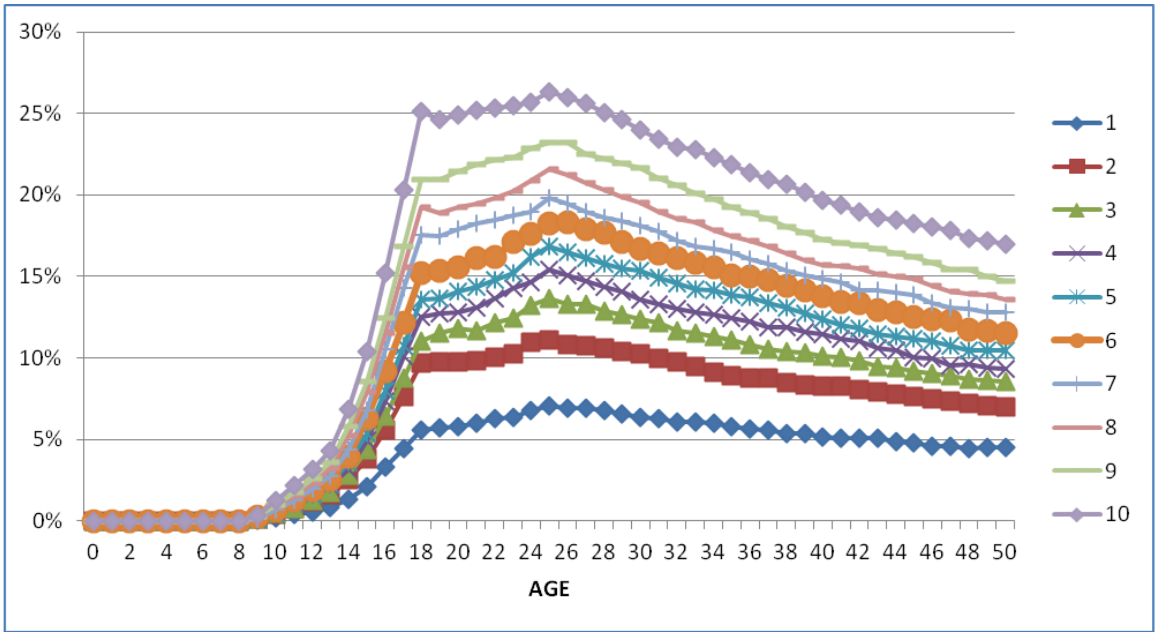
3. Adjust initiation rates in the model by applying a single scalar across all initiation rates.
    a. Four calibration runs were required
    b. The final grouping, relative initiation scalars and number of counties are listed in Table 6

Table 6: Baseline smoking tobacco relapse rates

Band	Relative initiation rate	Current median rate	Low	High	Range	N (counties)
1	0.3494	8%	0.8%	10.0%	9.2%	87
2	0.5872	11%	10.0%	13.0%	3.0%	162
3	0.7026	14%	13.0%	15.0%	2.0%	198
4	0.7833	16%	15.0%	18.0%	3.0%	413
5	0.9215	19%	18.0%	20.0%	2.0%	343
6	**1.0000**	**21%**	**20.0%**	**22.0%**	**2.0%**	**345**
7	1.1042	23%	22.0%	24.0%	2.0%	332
8	1.2709	25%	24.0%	27.0%	3.0%	330
9	1.4227	28%	27.0%	30.0%	3.0%	199
10	1.7057	33%	30.0%	47.0%	17.0%	124

Figure 9 illustrates the pattern of smoking prevalence for a series of modeled birth cohorts by initiation band.

Figure 9: Lifetime tobacco prevalence by initiation band



Determination of cigarette consumption and tax revenue

Upon entering the "smoking state" (i.e. upon smoking initiation of relapse), the agent's daily cigarette consumption (CPD) is determined by a random draw from a Poisson distribution conditioned on age, sex, ethnicity, and education-based averaged estimated from the 2012 NHIS survey. Table 7 summarizes the distribution of cigarettes per day (CPD) by key demographics and consumption categories.

Table 7: Cigarettes per day (CPD)

Cigarettes per day (CPD)		1-9	10-19	20-29	30+
	Overall	37.51%	33.02%	23.61%	5.86%
Sex	**Male**	35.50%	30.99%	26.21%	7.30%
	Female	40.06%	35.59%	20.32%	4.03%
Age (M)	**0-18***	80.00%	10.00%	10.00%	0.00%
	18-24	50.83%	31.68%	15.51%	1.98%
	25-44	41.06%	31.11%	23.86%	3.96%
	45-64	27.11%	31.17%	30.52%	11.20%
	65+	30.46%	29.31%	29.31%	10.92%
Age (F)	**0-18***	80.00%	10.00%	10.00%	0.00%
	18-24	52.41%	34.48%	12.07%	1.03%
	25-44	42.48%	36.98%	17.83%	2.71%
	45-64	35.47%	34.83%	23.76%	5.94%
	65+	37.78%	34.38%	23.58%	4.26%
Ethnicity (M)	**1 (White)**	24.53%	32.32%	33.01%	10.14%
	2 (Black)	45.58%	35.77%	15.96%	2.69%

	3 (Hispanic)	64.62%	21.54%	12.75%	1.10%
	4 (Other)	54.68%	26.60%	14.29%	4.43%
Ethnicity (F)	**1 (White)**	31.94%	38.73%	24.39%	4.93%
	2 (Black)	52.05%	35.45%	11.01%	1.49%
	3 (Hispanic)	67.80%	20.34%	9.83%	2.03%
	4 (Other)	62.70%	18.25%	15.08%	3.97%
Education (M)	**1 (No HS)**	34.35%	28.12%	28.55%	8.99%
	2 (HS)	31.07%	33.59%	28.04%	7.30%
	3 (Post-Secondary)	41.84%	29.42%	22.58%	6.16%
Education (F)	**1 (No HS)**	37.29%	33.28%	23.08%	6.35%
	2 (HS)	37.18%	34.29%	23.72%	4.81%
	3 (Post-Secondary)	46.43%	33.67%	17.19%	2.71%

**Not in NHIS data and derived from YRBS*

As shown in Table 6, cigarette consumption among smokers tends to increase with age. To accommodate this trend, the agent's CPD is reset by a new random draw every five years they remain in the smoking state. For those who relapse, CPD is also determined by another random draw upon returning to the smoking state.

The total number of packs consumed during the year, as well as per-pack tax revenue, is determined by assuming there are 20 cigarettes/pack and applying the following calculation:

Total packs in year (TPY)= (CPD * 365.25)/20

For smokers initiating and those quitting during a year, a proportion of their TPY is applied using a random draw from a uniform distribution.

The Health Impact Module

The Health Impact Module determines the health impact of the policy, or policies, under evaluation by determining how the smoking behavior of simulated individuals impacts disease incidence, morbidity and mortality. It does this by comparing the disease outcomes of each agent that occur in the baseline scenario (i.e. a world *without* the intervention) with those that occur in the intervention scenario. Population-wide estimates of an intervention's impact are determined by aggregating individual effects.

A key benefit of tracking health outcomes at the level of individual agent is that both the incidence and timing of disease outcomes can be determined. For instance, consider an individual who, without intervention, begins smoking at age 18 and suddenly dies of a cardio-vascular event at age 55. With the USPSTF recommendation of primary care referral for smoking cessation counseling and/or the use of a tobacco cessation drug, that individual successfully quits at age 30 and has no cardio-vascular event at age 55. However, at age 70, that same individual is diagnosed with colorectal cancer and dies at age 75. The net health impact of the USPSTF recommendation upon that individual is 15 additional disease-free years and 5 additional years with colorectal cancer. From a population perspective, this agent's adherence to the USPSTF recommendation provided additional quality adjusted life years, a decrease in cardio-vascular disease burden, but an increase in cancer incidence and burden. By tracking policy impact at the individual level, we are able to identify which events are avoided and which additional events occur during the extended lifespan resulting from the policy. The Health Impact Module is capable to tracking health impacts in one of two ways. The baseline method is to track outcome across a variety of tobacco-related diseases simultaneously using age, sex, and smoking status based risks derived from the Smoking-Attributable Mortality, Morbidity, and Economic Costs (SAMMEC) tool maintained by the Center for Disease Control (CDC). This approach provides a broad accounting of all smoking attributable risks and diseases. The second approach provides a detailed examination of cardio-vascular events and disease burden by interfacing with the CVD Prevention Policy Model.

SAMMEC-based disease incidence and burden estimation

The Smoking-Attributable Mortality, Morbidity, and Economic Costs (SAMMEC) tool maintained by the Center for Disease Control (CDC) provides age, sex, and smoking status specific incidence and mortality rates for ten cancers, six cardiovascular outcomes, and three respiratory diseases.

The Health Impact Module independently evaluates incidence of each disease. Given incidence of a particular disease, severity, final outcome (death or recovery), and episode duration is determined. Disease specific quality of life (QofL) decrements are imposed during disease episodes to capture morbidity with the maximum decrement across all concurrent episodes of .5 quality adjusted life years (QALYs)

Evidence indicates that recent quitters have similar smoking-attributable health risks to current smokers for approximately 4 years[1] after quitting. Then their disease risk declined, but never returns to that of a person who has never smoked. Table 5 lists the diseases included in the Health Impact Model only with their assumed duration and quality of life decrement.

Table 8: Summary of diseases included in ModelHealth: Tobacco

	Episode Duration		Quality Adjusted Life Year Decrements	
CANCERS	**Terminal**	**Non-Terminal**	**Initial Year of Event**	**Subsequent Years**
Lip, Oral Cavity, Pharynx	4.3	5	0.2	0.2
Esophagus	1.8	5	0.3	0.3
Stomach	3	5	0.2	0.2
Pancreas	1.24	5	0.3	0.3
Larynx	2	5	0.3	0.3

[1] *This 4-year cut-point is pending thorough literature review on smoking-attributable risks.*[1]

Trachea_Lung_Bronchus	2	5	0.3	0.3
Cervix Uteri	4	5	0.2	0.2
Kidney and Renal Pelvis	4.7	5	0.2	0.2
Urinary Bladder	4.7	5	0.2	0.2
Acute Myeloid Leukemia	4.6	5	0.2	0.2
CVD				
Ischemic Heart Disease	0	0.5	0.1500	
Other Heart Disease	5	0.0769	0.0231	0.3
Cerebrovascular Disease Stroke+	1	until death	0.4000	0.4
Atherosclerosis	5	0.0769	0.0231	0.3
Aortic Aneurism	0	0.0769	0.0231	
Other Arterial Disease	5	0.0769	0.0231	0.3
Respiratory Disease				
Pneumonia Influenza	0	0.0384	0.0115	
Bronchitis Emphysema+	5	until death	0.2	0.2

**Durations are rounded up to the nearest cycle. Episodes with 0 duration indicate instant death and no decrement applied.*

***For CVD and Respiratory Diseases, the initial year decrement is scaled to reflect partial year episode*

+Following initial episode, agent remains at risk for death in future cycles.

The duration of terminal cancer episodes ranges from 1 to 5 years with applicable decrements applied during the terminal episode. The duration of a non-terminal cancer episode was assumed to be 5 years across all cancers. Quality of Life decrements were the same for both terminal and non-terminal cancer episodes and ranges from .2 to .3 QALYs. Once a non-terminal cancer episode ended, the individual is at risk of another episode of that cancer with no addition or reduced risk due to relapse.

From the SAMMEC data, cardiovascular and respiratory disease were modeled as both terminal events and chronic episodes with quality of life decrements ranging from .01 (influenza) to .4 (stroke). Events resulting in death had duration of one year. Non-terminal cardiovascular and respiratory events did not end, and the corresponding quality of life decrement was imposed every year following the event. Individuals experiencing a non-terminal cardiovascular and/or respiratory event could experience a repeat event. Their risk for such a repeat event was the same as that of experiencing the initial event. For instance, a non-terminal cerebrovascular disease episode (i.e. stroke) resulted in a quality of life decrement of .4 QALYs every cycle following that event. The individual experiencing that initial stroke was at risk of another stroke in subsequent years. Similarly, a person could experience repeated cases of pneumonia and/or influenza.

Costs and productivity

Determining the costs of disease and medical utilization attributable to tobacco use can be tracked in two different ways in ModelHealth: Tobacco. These two methods correspond to the two different ways disease incidence and burden may be determined. The first is based on direct observation of the total costs of care (TCOC) of current smokers, as well as a functional relationship between current- and never-smoker costs. The second, which corresponds to the more detailed analysis of CVD risk using ModelHealth: CVD, estimates the costs incurred by specific cardiovascular events.

Total cost of care approach: Smoking-attributable medical costs

We estimated the medical costs of smoking from observed associations between smoking status and medical costs in the Medical Expenditure Panel Survey (MEPS), using smoking status from linked National Health Interview Survey (NHIS) responses. We followed the method of Levy et al.[16,17], including controlling for potentially confounding factors in a two-part model using a gamma distribution and a log-link in the second part. However we combined multiple years of data (2001-2010) to create more stable estimates for age, sex and smoking status subgroups; we also estimated separate models by primary insurer to determine smoking costs by the primary insurer type. MEPS and other claims data are complicated by higher utilization of former smokers (whose quits were likely prompted by diagnoses that lead to increased healthcare utilization in the years following their successful quits). For former smokers, we fit an exponential function to the relationship of current and former risk based on time since quit, as reported by the Congressional Budget Office (Figure 3-5 in CBO report). We applied this function to the costs for current smokers (which we estimated from MEPS data) to obtain estimates of what the medical costs of former smokers would be by age, sex and time since quit, assuming they had a proactive quit:

$$y = 0.9927 - 1.086e(-0.1171t)$$

Where y is the portion of a current smokers' smoking-attributable costs that is reduced according to years since quit (=t). Thus each former's smoker cost will be calculated as a portion of current smokers' costs with the same age, sex and insurance status as estimated from MEPS.

Table 9. Smoking-attributable medical costs by age, gender, smoking status ($2012)

Age categories (in years)	Male Current	Female Current	Male Former*	Female Former*
Private Insurance				
0-34	0	0	0	0
35-44	987	1,210	604	740
45-54	1,265	1,499	774	917
55-64	1,597	1,843	977	1,128
65-74	1,994	2,253	1,220	1,379
75-84	2,465	2,743	1,509	1,679
85+	2,734	3,024	1,673	1,851

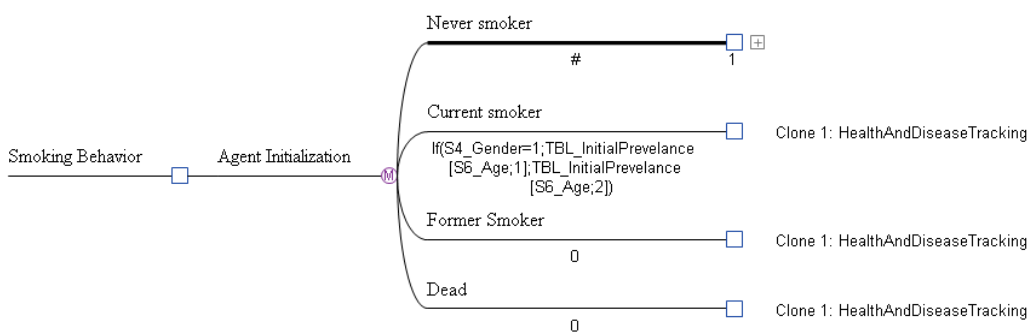
Medicare Insurance				
0-34	0	0	0	0
35-44	1,301	1,531	796	937
45-54	1,639	1,879	1,003	1,150
55-64	2,040	2,296	1,248	1,405
65-74	2,518	2,795	1,541	1,710
75-84	3,089	3,391	1,890	2,075
85+	3,414	3,733	2,090	2,284
Medicaid Insurance				
0-34	0	0	0	0
35-44	1,823	2,117	1,115	1,296
45-54	2,283	2,593	1,397	1,587
55-64	2,830	3,162	1,732	1,935
65-74	3,480	3,842	2,130	2,351
75-84	4,258	4,656	2,606	2,850
85+	4,702	5,123	2,878	3,136
Uninsured				
0-34	0	0	0	0
35-44	374	548	229	335
45-54	517	710	316	435
55-64	695	906	426	554
65-74	914	1,138	559	697
75-84	1,180	1,415	722	866
85+	1,332	1,571	815	962
Other/Multiple Insurance				
0-34	0	0	0	0
35-44	1,536	1,783	940	1,091
45-54	1,922	2,184	1,177	1,337
55-64	2,384	2,664	1,459	1,630
65-74	2,932	3,236	1,795	1,980
75-84	3,587	3,922	2,195	2,400
85+	3,961	4,315	2,424	2,641

*Costs of former smokers are determined by time since quit as described in the text. Former smoker costs here depict 5 years since quit.

ModelHealth: Tobacco overview of model structure

Rather than create a complex decision tree comprised of a diverse set of decision nodes that each represent a different combination of tobacco use and disease state, ModelHealth: Tobacco separates these two processes. The decision tree is used to track tobacco use status and trackers are used to track disease processes. The values of these disease trackers are evaluated each cycle and corresponding actions occur. Figure 10 presents the basic decision tree.

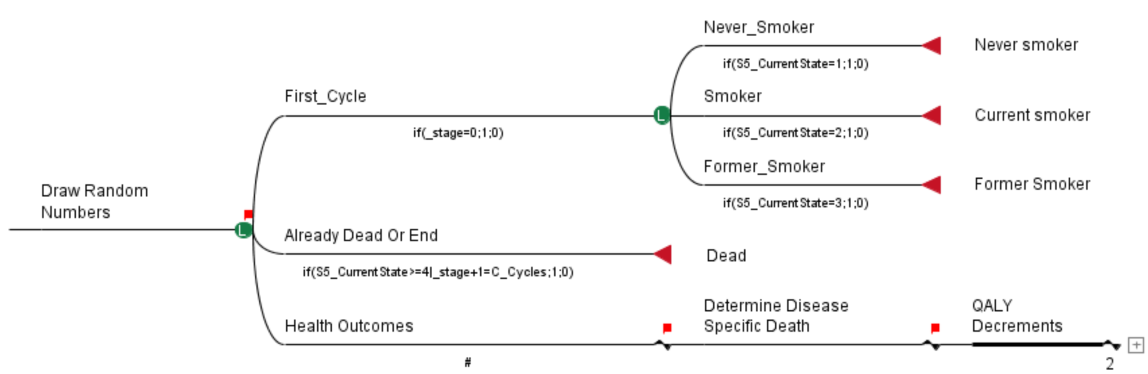
Figure 10. The Basic Decision Tree



The model comprises three tobacco use states (never smoker, current smoker, former smoker) and dead, which is an absorbing state in the model. Agents are initialized at the "Agent Initialization" node and their initial tobacco use determined according to distributions defined by the initial tobacco use estimations presented in Table 9.

Each cycle, each agent's health outcomes and related morbidity and mortality impacts are evaluated. This is done by the Health and Disease Tracking sub-model as presented in Figure 11.

Figure 11: The Health and Disease Tracking Sub-module



The Health and Disease Tracking sub-module begins by drawing a set of random numbers. These will be used to determine tobacco-related disease incidence and outcomes as well as determine the tobacco state transition in the next cycle. Then, it identifies if the agent is currently in the model's absorbing state, death. Although the random number draw may appear redundant for agents currently in the absorbing state, this step is essential in ensuring the virtual self-controlled trial nature of ModelHealth: Tobacco.

If the agent is alive at the start of the cycle, the Health and Disease evaluates disease incidence and burden are to be evaluated.

This evaluation process is the same across all tobacco use states; however, the risk of disease and disease-specific death change according to the agent's current tobacco use state as determined by the SAMMEC tool.

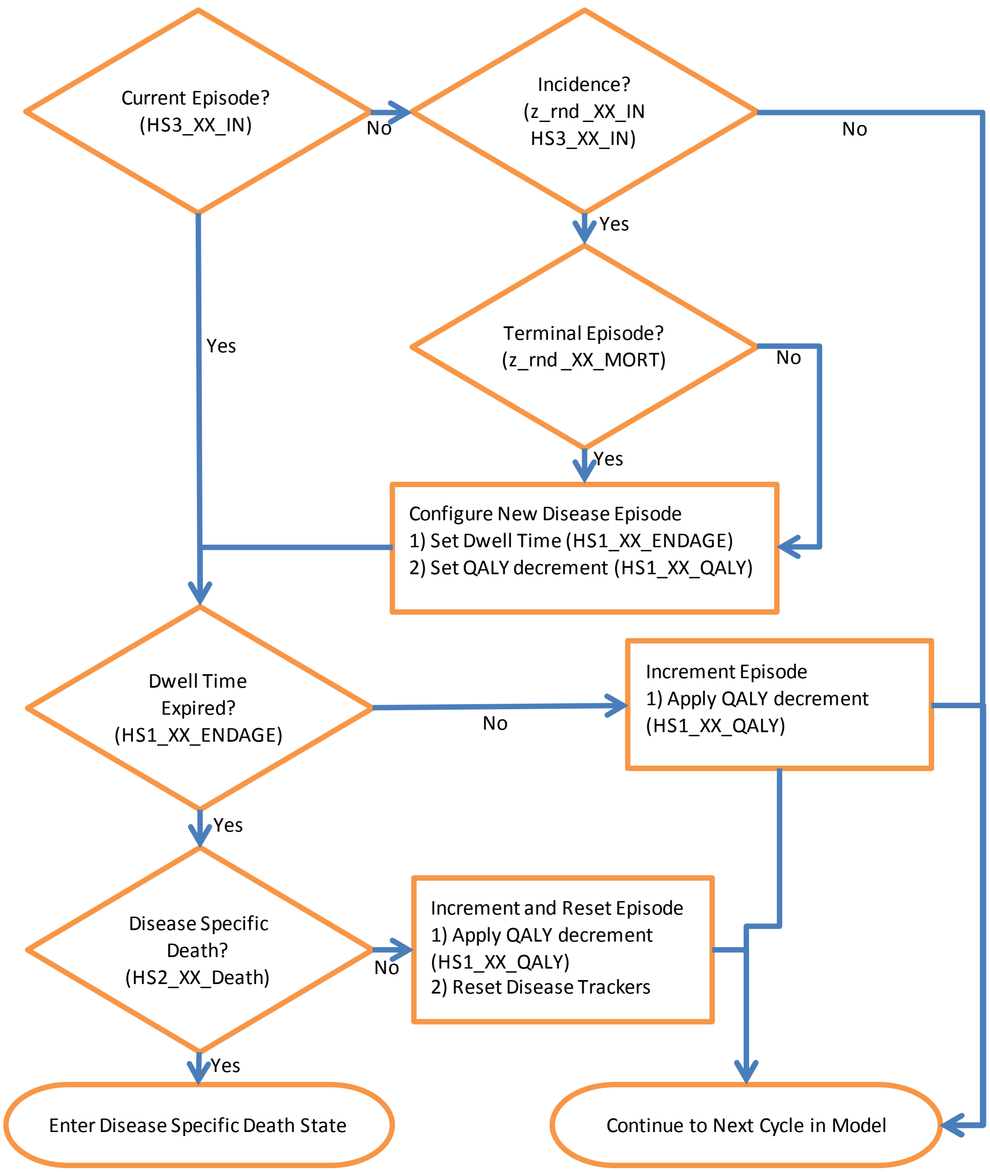
As noted, the SAMMEC-based disease incidence and burden estimation places no precedence on the incidence or mortality burden of any disease. Instead, it simultaneously tracks incidence of all tobacco-related diseases and their corresponding burden simultaneously. Thus, within the model, it is possible for an agent to experience multiple cancers *and* cardiovascular events at the same time. Further, if simultaneously occurring conditions both result in a tobacco-related death within the same cycle, the model attributes the death to both diseases while tracking a single tobacco-related fatality.

Once health status and disease burden have been evaluated, the agent's tobacco use state for the upcoming cycle is determined by the Transition Probabilities sub-model, or clone, as presented in Figure 12.

Determining disease incidence and outcomes (SAMMEC)

When the SAMMEC disease incidence and outcomes sub-model portion of ModelHealth: Tobacco is used to evaluate tobacco-related disease incidence and burden, two groups of trackers are used simultaneously for the tobacco-related diseases. Although each disease has unique risks, durations and quality adjusted life year (QALY) decrements, the basic algorithm employed is the same across all diseases tracked in the model. This algorithm is presented in Figure 12.

Figure 12: Disease tracking algorithm



For each disease, the model first checks to see if the agent is experiencing a current disease episode (i.e. is currently in a disease "state"). If they are, their time in that state (i.e. dwell time) is checked to determine if the episode's dwell time has expired and the episode be terminated during the current cycle. If the dwell time has expired, the episode's terminal condition (death or resolution) is checked and the appropriate actions taken.

If an episode's dwell time has not expired, disease and terminal condition specific quality adjusted life year decrements are applied and the episode continues into the next cycle.

If the agent is not in a disease state for that disease, the model determines if a new incidence of that disease has occurred. If it has, the eventual terminal condition of that state (death or resolution) is determined, and the other trackers for that disease configured.

This process is repeated for each disease and during each cycle the agent is in a non-absorbing state. Each disease is tracked independently. So, the existence of one disease does not affect the likelihood of another's occurrence. In addition, no precedence is given across diseases. Any given agent can be in multiple disease states at the same time, and that agent can experience a disease specific death from multiple causes. Finally, although an agent may experience re-occurrence (multiple episodes) of the same disease, each episode's risks are evaluated independent of one another. No risk adjustments are made for multiple episodes of the same disease.

References

1. Counseling and interventions to prevent tobacco use and tobacco-caused disease in adults and pregnant women: U.S. Preventive Services Task Force reaffirmation recommendation statement. *Ann Intern Med.* Apr 21 2009;150(8):551-555.
2. Moyer VA. Primary care interventions to prevent tobacco use in children and adolescents: U.S. preventive services task force recommendation statement. *Pediatrics.* Sep 2013;132(3):560-565.
3. The Community Preventive Services Task Force. The Community Guide. 2008; www.thecommunityguide.org.
4. The Community Preventive Services Task Force. Interventions to increase the unit price for tobacco products. *Reducing tobacco use and secondhand smoke exposure* http://www.thecommunityguide.org/tobacco/increasingunitprice.html
5. The Community Preventive Services Task Force. Mass-reach health communication interventions. *Reducing tobacco use and secondhand smoke exposure* http://www.thecommunityguide.org/tobacco/massreach.html
6. The Community Preventive Services Task Force. Reducing out-of-pocket costs for evidence–based tobacco cessation treatments. *Reducing tobacco use and secondhand smoke exposure* http://www.thecommunityguide.org/tobacco/RRoutofpocketcosts.html.
7. Community Preventive Services Task Force. Quitline interventions. *Reducing tobacco use and secondhand smoke exposure* http://www.thecommunityguide.org/tobacco/RRquitlines.html.
8. The Community Preventive Services Task Force. Multicomponent Interventions that Include Telephone Support. *Reducing Tobacco Use and Secondhand Smoke Exposure* http://www.thecommunityguide.org/tobacco/multicomponentinterventions_archive.html.
9. Community Preventive Services Task Force. Smoke-free policies. *Reducing tobacco use and secondhand smoke exposure* http://www.thecommunityguide.org/tobacco/smokefreepolicies.html.
10. The Community Preventive Services Task Force. Comprehensive Tobacco Control Programs. *Reducing Tobacco Use and Secondhand Smoke Exposure* http://www.thecommunityguide.org/tobacco/comprehensive.html.
11. Community Preventive Services Task Force. Community mobilization with additional interventions to restrict minors' access to tobacco products. http://www.thecommunityguide.org/tobacco/communityinterventions.html.
12. Community Preventive Services Task Force. Incentives and competitions to increase smoking cessation among workers. *Reducing tobacco use and secondhand smoke exposure* http://www.thecommunityguide.org/tobacco/incentives.html.
13. National Center for Education Statistics: Data Tools in Sciences. 2014. http://nces.ed.gov/datatools.
14. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria2014.
15. *Recursive partitioning and regression trees* [computer program]2014.
16. Levy DT, Friend K. Examining the effects of tobacco treatment policies on smoking rates and smoking related deaths using the SimSmoke computer simulation model. *Tob Control.* Mar 2002;11(1):47-54.

17. Levy DT, Nikolayev L, Mumford E, Compton C. The Healthy People 2010 smoking prevalence and tobacco control objectives: results from the SimSmoke tobacco control policy simulation model (United States). *Cancer Causes Control.* 2005 May 2005;16(4):359-371.