



FLEXKB DOCUMENT
Version: HI.001.01132012.83611
Document generated: 01/13/2012

FRED HUTCHINSON CANCER RESEARCH CENTER (FHLUNG)

[Readers Guide](#)
[Model Overview](#)
[Assumption Overview](#)
[Parameter Overview](#)
[Component Overview](#)
[Output Overview](#)
[Results Overview](#)
[Key References](#)

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at <http://cisnet.cancer.gov/profiles>. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: [Reader's Guide](#).



READERS GUIDE

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5-10 minutes. Each contains links to more detailed information if required.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

- [Smoking History Generator Component](#)
- [Population Component](#)
- [Natural History Component](#)
- [Survival Mortality Component](#)

Output Overview

Definitons and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.



- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

MODEL PURPOSE

SUMMARY

This document provides a brief overview of the Fred Hutchinson Cancer Research Center (FHCRC) lung cancer model. This model uses an underlying biologically based multistage model (with additional age, period and cohort effects) to represent the effects of smoking on the natural history of lung cancer. The FHCRC lung cancer model may be utilized to address questions about the impact of public health information on US lung cancer trends, and to predict the hypothetical impact of alternative tobacco control policies.

PURPOSE

The purpose of the FHCRC lung cancer model is to serve as an effective tool for evaluating lung cancer trends in the US population, and the effects of possible interventions. The FHCRC lung cancer model combines an underlying biologically based natural history model that is calibrated to individual smoking histories in substantial US lung cancer mortality cohorts with additional age, period, and birth cohort effects to improve the calibration to US lung cancer mortality data. Limitations include that it is not calibrated to US lung cancer incidence or CT screening cohort data.

Two distinct modeling projects contributed to the development of the FHCRC lung cancer model. The first project consisted of calibrating a biologically based natural history model to individual smoking histories in several substantial lung cancer cohorts. The calibrated model parameters (including background and smoking dose response parameters) were shared with other Cancer Intervention and Surveillance Modeling Network (CISNET) lung cancer modeling groups. The second project (called the Lung Smoking Base Case) consists of combining the biologically based natural history model with additional age, period, and birth cohort effects. This model is calibrated to lung cancer deaths in the US population by single years of age and calendar year. It should be of use in evaluating the effects of alternative tobacco control policies, and in making projections of future US lung cancer mortality.

Development of the FHCRC lung cancer model took place in two distinct projects:

[Project One](#) : Calibrating a natural history model to smoking cohort data

[Project Two](#) : Lung Smoking Base Case - Modeling US lung cancer mortality trends

CATEGORIES

[Core Docs](#)



MODEL OVERVIEW

SUMMARY

This document describes previous work leading to this model and model itself in general terms.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

PURPOSE

We wish to understand the effects of smoking and other factors on US lung cancer mortality. We modeled the effects of smoking using the biologically based two-stage clonal expansion (TSCE) model (Moolgavkar et al. 1979, 1981, 1990; Heidenreich et al., 1997). The TSCE model relates smoking to biological rates for cell initiation, promotion, and malignant conversion processes. The effects of additional unknown factors that may have influenced US lung cancer mortality rates were modeled using period and cohort effects.

BACKGROUND

Lung cancer is the leading cause of cancer death in the US, and smoking is the most important risk factor for developing lung cancer. Thus in modeling lung cancer in the US, we felt it was important to use the best available methods to relate smoking to lung cancer risk. The biologically based TSCE model seemed best for this purpose.

The FHCRC lung cancer project began with calibration of the TSCE model to several large smoking cohorts, modeling individual smoking histories in relation to lung cancer incidence and mortality [[Project One](#)]. This was followed by the Lung Smoking Base Case [[Project Two](#)] in which we used the calibrated TSCE model to represent effects of smoking on lung cancer mortality in the US population. We also introduced additional corrections as a function of period and birth cohort to improve the fit to US lung cancer mortality.

The TSCE model was initially developed by Moolgavkar, Venzon, and Knudson. This model has been applied to analyze many types of cancer, including the effects of smoking and other exposures. Calibrating the TSCE model to cohort data consists of estimating dose-response relationships for these exposures as they affect cell initiation, promotion, and malignant conversion rates. Maximum likelihood methods allow optimization of the model to represent temporal patterns of risk associated with different exposure histories of individuals in the cohort.

The TSCE natural history model represents basic cellular processes, including cell division, apoptosis, and mutation, that contribute to three distinct phases in the carcinogenic process: initiation, promotion (birth minus death of initiated cells) and malignant conversion ([TSCEModel Details](#)). The TSCE model represents a significant simplification of the biological processes associate with lung cancer. The model ignores the possibility of multiple cancer pathways and disease subtypes. However, it does provide a rigorous mathematical representation of processes that are considered as the



rate-limiting events in carcinogenesis, and has provided excellent fits to individual and population data for many cancer types (Moolgavkar and Luebeck, 2003).

The FHCRC lung cancer model builds on previous analyses using the TSCE model to relate lung cancer risk to individual exposure patterns for smoking, radon, arsenic, fibers, and radiation (Castren et al., 1999; Hazelton et al., 2001, 2005, 2006; Haylock et al., 2004; Heidenreich et al, 2002; Kai et al, 1997; Little et al., 2002; Luebeck et al., 1999, 2000; Meza et al., 2008; Moolgavkar et al., 1989, 1993, 1998, 1999, 2000, 2001a, 2001b, 2001c; Stevens et al., 1979, 1984). These analyses consistently show that the most important lung cancer risk factor is tobacco smoke, with the risk increasing non-linearly with smoking duration (Hazelton et al., 2005; Meza et al., 2008).

The FHCRC lung cancer model was applied to the Smoking Base Case [[Project Two](#)], using the TSCE model to represent effects of smoking, and period and cohort effects to represent other unknown factors. Inputs included US population data and lung cancer deaths for males and females binned by single year of age and calendar year, and a smoking history generator developed by NCI to simulate smoking histories and other cause mortality for individuals in the US. Outputs are estimates of lung cancer deaths by gender, age, and calendar year given historical smoking patterns, and also counterfactual estimates for lung cancer deaths given alternative US smoking patterns.

MODEL DESCRIPTION

The FHCRC lung cancer model consists of a biologically based TSCE natural history model of the effects of smoking on lung cancer mortality, along with period and birth cohort effects to represent lung cancer mortality in the US population (called the TSCE-PC model, representing the age effects given by the TSCE natural history model, along with period and birth cohort effects). A second model (called the TSCE-APC model) includes additional age effects to capture possible discrepancies between age effects in the TSCE calibration to US lung cancer mortality, and compensate for possible limitations of the TSCE model in representing the effects of tobacco smoke on lung cancer mortality.

For more details see: [TSCModel Details](#)

CONTRIBUTORS

William D Hazelton
Jihyoun Jeon
Rafael Meza
Suresh Moolgavkar

CATEGORIES

[Core Docs](#)



ASSUMPTION OVERVIEW

SUMMARY

This document discusses assumptions underlying the model as well as some of their implications.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

BACKGROUND

The TSCE natural history model relates individual non-smoking and smoking histories to cellular processes that contribute to the development of lung cancer. The TSCE model allows calculation of the time-dependent probability for lung cancer mortality at each age. The TSCE model includes an initial rate for mutation or epi-genetic change leading to initiated cells, a birth rate and death rate for the initiated cells, and a rate for second mutation or epi-genetic change that occurs during the cell division process of an initiated cell to generate a malignant cell as well as another initiated cell. A lag time or lag time distribution is used to represent the time from the first malignant cell to cancer incidence or mortality from cancer. Smoking is assumed to affect any or all of the rates through flexible dose-response functions.

To assess model reliability ([Project One](#)), we looked at consistency of model parameter estimates between the different lung cancer mortality cohorts. We also analyzed separately lung cancer incidence by subtype in the NHS and HPFS cohorts, where that data is available.

In working on the Lung Smoking Base Case - [Project Two](#), the FHCRC group found that the TSCE natural history model, when calibrated to cohort data for lung cancer in relation to smoking, does not fully account for lung cancer mortality in the US population. Thus the FHCRC group found it necessary to include additional age, period, and birth cohort adjustments that may correct for limitations of the TSCE natural history model, for changing cigarette composition, and other exposures and environmental factors that contribute to lung cancer in the US population. The FHCRC lung cancer model combines the TSCE natural history model for smoking with these additional age, period, and birth cohort adjustments.

ASSUMPTION LISTING

1. The TSCE natural history model (used in the [Project One](#) and [Project Two](#) by the FHCRC group) assumes two stochastic rate-limiting mutation events, clonal expansion of initiated cells, and a lag time from the first occurrence of a malignant cell to the time of lung cancer death.
2. Calibration of the TSCE model to the HPFS and NHS lung cancer mortality data (see [Project One](#) and Calibration and Validation sections) and subsequent estimation of lung cancer deaths for the Lung Smoking Base Case included a model assumption that there are a fixed number (10^7) normal stem cells in lung. Clearly the number must increase during embryogenesis, and any trend throughout life has not been ascertained, nor has the total number, as lung stem cells are difficult to identify. However, the likelihood based modeling approach will estimate an initiation rate that will compensate any error on the assumed number of stem cells at risk for initiation.



3. The background rates for initiation and malignant conversion rates set equal to each other, $\nu = \mu$. This assures identifiability of the TSCE model parameters.
4. Normal stem cells may undergo faulty division to create an initiated cell.
5. Growth of the population of initiated cells (promotion) is modeled stochastically through cell birth and death process. This process can not be observed directly, but is consistently estimated between different cohorts as the most important mechanism whereby cigarettes influence the risk of lung cancer.
6. Malignant conversion in the TSCE model is assumed equivalent to first occurrence of a malignant cell arising through faulty division of an initiated cell.
7. A lag time or lag time distribution is assumed to represent the time between the occurrence of the first malignant cell and cancer death.
8. The effects of cigarette smoke is modeled as a constant dose rate during periods of smoking, and the smoking dose has separate non-linear (power-law) influences on the initiation, promotion, and malignant conversion rates in the TSCE model.
9. The smoking dose-response from the HPFS/NHS calibration is applicable to the simulated US population data, given subsequent adjustments for additional age, period, and birth cohort.
10. The SHG provided by NCI provides individual smoking histories consistent with the historical patterns of smoking in the US.
11. The other cause mortality input provided by the SHG reflects historical trends in the US population.
12. An additional age effect may correct for deficiencies of the TSCE model, and differences between age effects in the calibration cohort and the US population.
13. A period effect is adequate to account for historical changes in cigarette composition, demographic changes, and changes in health care that may influence lung cancer mortality.
14. A birth cohort effect is able to capture effects of environmental, nutritional, and other factors that influences the lifetime risk for lung cancer.

CATEGORIES

[Core Docs](#)



PARAMETER OVERVIEW

SUMMARY

This document provides an overview of the major parameters in the model, their sources, and general implications they have on model outputs.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

BACKGROUND

Parameters related to natural history for the Lung Smoking Base Case ([Project Two](#)) were estimated from the TSCE model calibration to lung cancer mortality in smoking cohorts ([Project One](#)), focusing on the HPFS and NHS cohorts. Additional age, period and birth cohort parameters were estimated using US population and lung cancer mortality data. Additional demographic parameters are embedded in the SHG to reflect historical smoking trends and rates for other cause mortality.

PARAMETER LISTING OVERVIEW

The FHCRC lung cancer model parameters are categorized into:

1. Background parameters of the TSCE model ([Natural History Component](#)).
2. Dose-response parameters that relate current cigarette smoke exposure to the rates for initiation, promotion (birth minus death of initiated cells) and malignant conversion ([Natural History Component](#)).
3. Lag time parameters that describe the time lag or gamma lag time distribution from the first malignant cell to cancer death (See [Natural History Component](#) and [Survival Mortality Component](#)).
4. Additional demographic parameters used in the Lung Smoking Base Case ([Project Two](#)): additional age, period, and birth cohort parameters applied to the simulated US population ([Population Component](#)) to adjust the lung cancer mortality calculations to represent US lung cancer deaths ([Survival Mortality Component](#)).

The additional age, period, and birth cohort parameters are based on:

1. Smoothing parameters used to generate single year US population data from census data
2. Smoothing parameters used to generate annual US lung cancer mortality data

BACKGROUND PARAMETERS:

1. Background initiation, malignant conversion rate
2. Background initiated cell division rate
3. Background net initiated cell promotion rate

DOSE-RESPONSE PARAMETERS FOR FULL MODEL:

1. Coefficient multiplying dose response for initiation
2. Power of dose for initiation



3. Coefficient multiplying dose response for promotion
4. Power of dose for promotion
5. Coefficient multiplying dose response for malignant conversion
6. Power of dose for malignant conversion

NOTE 1. Typically only three or four dose-response parameters are required to model lung cancer incidence or mortality due to cigarettes - two for the dominant effect of promotion, and one or two describing the much smaller effect on initiation or malignant conversion.

LAG TIME PARAMETERS:

Mean and standard deviation for gamma distribution, or fixed lag time

CATEGORIES

[Core Docs](#)



COMPONENT OVERVIEW

SUMMARY

This is a description of the basic computational building blocks/components of the model.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

OVERVIEW

Several components are involved to construct the FHCRC lung cancer model for the Lung Smoking Base Case. A [Population Component](#) uses individual simulated smoking and other cause mortality histories generated by the Smoking History Generator to generate a simulated US population. A [Natural History Component](#) utilizes the TSCE model, previously calibrated to smoking cohort data in [Project One](#), to estimate lung cancer deaths in the simulated US population based on the TSCE model. A [Survival Mortality Component](#) includes effects of the lag time from first malignant cell to lung cancer death in the TSCE model, and adjustments for additional age, period, and birth cohort to improve the fit to US lung cancer mortality.

COMPONENT LISTING

The components used to construct the FHCRC lung cancer model include:

- [Population Component](#)
- [Natural History Component](#)
- [Survival Mortality Component](#)

CATEGORIES

[Core Docs](#)



SMOKING HISTORY GENERATOR COMPONENT

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

SUMMARY

The smoking history generator (SHG) is a shared precursor micro-simulation model that produces cohort-specific smoking histories and deaths due to causes other than lung cancer as inputs for the dose-response models used by members of the CISNET lung cancer consortium.

OVERVIEW

The core SHG software was parameterized using three tobacco control scenarios to produce the requisite input data for the models. The first, called the actual tobacco control (ATC) scenario, is a quantitative description of actual smoking behaviors of males and females born in the United States between 1890 and 1984. The second, called no tobacco control (NTC), is a quantitative description of predicted smoking behaviors of males and females in the United States under the assumption that tobacco control efforts starting mid-century had never been implemented. The third, called complete tobacco control (CTC), is a quantitative description of predicted smoking behaviors of males and females in the United States under the assumption that tobacco control activities yielded perfect compliance, with all cigarette smoking coming to an end in the mid-sixties. The ATC scenario used inputs derived directly from observed data in the National Health Interview Surveys (NHIS) and the Substance Abuse and Mental Health Services Administration (SAMHSA) National Survey on Drug Use and Health. The NTC scenario used inputs derived by extrapolating from trends in the observed histories before 1954, i.e., before any tobacco control in the decade leading up to the publication of the Surgeon General's Report in 1964. The CTC scenario was simulated by setting cessation rates to one (i.e., transferring all current smokers to former smokers) and allowing no further initiation starting in 1965 while using the observed values in earlier years.

DETAIL

The SHG accepts parameters supportive of the three tobacco control scenarios described above (see Table SGH-I below). The ATC scenario uses initiation, cessation and smoking intensity (CPD) rates directly derived from the NHIS and SAMHSA datasets. The NTC scenario uses initiation and cessation rates derived by fitting an age-



period-cohort model to the ATC rates upto 1954, i.e., before the appearance of any tobacco control measures, and by projecting those into the future maintaining them consistent with the patterns observed in 1954. The CTC scenario uses initiation and cessation rates identical to those of the ATC scenario upto 1965, and then sets the cessation rates equal to one and the initiation rates equal to zero, i.e., all smokers are forced to quit in 1965, and no new smokers are allowed to appear thereafter. All scenarios use smoking dependent other cause mortality (OCD) rates derived from several sources as mentioned above.

Computational process in the usage of the SHG

The CISNET SHG is implemented in C++ and consists of a single simulation class, that receives file system paths to five parameter files, four integer pseudorandom number generator (PRNG) seeds, and an optional immediate smoking cessation year parameter. The SHG simulation class employs four independent random selection processes that are implemented via a class-based wrapper of the Mersenne Twister PRNG.¹

Here we briefly describe the outline for computational process in the usage of the SHG:

1. Initialization

- a. Load input data
- b. Initialize random number streams

3. Start Simulation

- a. Validate inputs
- b. Determine Initiation Age (if any)
- c. Determine Cessation Age (if any)
- d. Compute cigarettes smoked per day (CPD) vector for those who initiate
 1. Determine smoking intensity group (based on initiation age)
 2. Determine CPD based on smoking intensity and age at initiation
 3. Determine uptake period and attenuate CPD during uptake period
 4. Generate CPD vector from initiation to cessation or simulation cutoff
- e. Compute other cause of death (OCD) age

5. Write individual outputs

6. Loop simulation if repeats are specified



RELEVANT PARAMETERS

The SHG utilizes input data from several sources: the NHIS data from 1965 to 2001, the SAMHSA data, the Berkeley mortality database cohort life-tables, the National Center for Health Statistics (NCHS), the Cancer Prevention Study I and II (CPS-I and CPS-II), and the Nutrition follow-up studies sponsored by the American Cancer Society. The NHIS and the SAMHSA datasets provide estimates for prevalence of never, former (by years quit) and current smokers by age and year, and data on smoking intensity (in terms of the average number of cigarettes smoked per day (CPD)). These data were used to create implicit initiation and cessation rates. Using the average initiation rate, the SHG is able to determine the likelihood that a never smoker becomes a smoker. For those individuals that are smokers, the cessation rates are used to determine the likelihood that a smoker becomes an ex-smoker. The Berkeley life-tables, combined with smoking prevalence estimates from NHIS and the relative risks of death for smokers and former smokers in comparison to never smokers from CPS-I and CPS-II, are used to produce the probability of death from causes other than lung cancer based on age, sex, birth cohort, and smoking status. Table SHG-I summarizes the input source for the SHG for the three CISNET tobacco control scenarios.

Table SHG-I

Input	ATC	NTC	CTC
Initiation rates	NHIS	Derived	Derived (no new smokers after 1965)
Cessation rates	NHIS	Derived	Derived (all smokers quit in 1965)
CPD ¹	NHIS, SMAHSA		
OCD ²	Berkeley life-tables, NCHS, NHIS, CPS-I, CPS-III, Nutrition Follow-up studies		
Birth year (1890-1984)	User Defined		
Gender (Male/Female)	User Defined		
Race (All race)	User Defined		

¹ Cigarettes smoked per day, ² Other Cause of Death

ATC: actual tobacco control, NTC: no tobacco control, CTC: complete tobacco control. To simulate life histories for individuals using the SHG, for any given run, the following parameters must be provided:



Table SHG-II

Parameter	Valid Values
Seed value for PRNG used for Initiation, Cessation, OCD ¹ , Smoking intensity quintile	Integer from -1 to 2147483647 (A value of -1 uses the clock time as the seed)
Race	0 = All Races
Sex	0=Male, 1=Female
Year of Birth	Integer from 1890 to 1984
Immediate Cessation year ²	0 or Integer from 1910 to 2000
Repeat ³	Integer >1 (number of times to repeat simulation)
File paths to Initiation,Cessation, OCD, Smoking intensity quintile and CPD ⁴ data files	As derived from NHIS depending on the scenario

¹Other cause of death, ² This variable is set to 0 except for CTC scenario. To apply immediate smoking cessation for CTC scenario, the year for immediate cessation must be supplied to the simulator. If the year value supplied is 0, immediate cessation will not be used in the run. If a year value is supplied, immediate cessation will occur on January 1st of year provided. ³Key is optional and can be excluded. If the Repeat value is included and is not a vector value, each set of parameters will be repeated by the amount specified. If the Repeat value is included and is a vector value, the repeat value will pertain to the value set that it corresponds to. ⁴Cigarettes smoked per day.

DEPENDENT OUTPUTS

The inputs of the SHG are used to simulate life histories (up to age 84) for individuals born in the United States between 1890 and 1984. These life histories include a birth year, and age at death from causes other than lung cancer, conditioned on smoking histories. For each simulated individual, the generated life histories include whether the individual was a smoker or not and, if a smoker, the age at smoking initiation, the smoking intensity in cigarettes per day (CPD) by age, and the age of smoking cessation. Smoking relapse, the probability that a former smoker starts smoking again, is not modeled. Table SHG-III summarizes the output of the SHG. Fig. SHG-1 shows two examples of smoking histories simulated by the SHG; a) an individual born in 1910 who begins smoking at age 17, quits at age 56 and dies at age 67 due to causes other than lung cancer, and b) an individual born in 1920 who begins smoking at age 22 and dies at age 53 due to causes other than lung cancer.

Table SHG-III

Table SHG-III

Initiation Age	Age at smoking initiation
Cessation Age	Age at smoking cessation
OCD ¹ Age	Age at death from cause other than lung cancer
Smoking History	Smoking intensity quintile (5 quintiles ranging from light to heavy smoking), Yearly smoking dose (CPD ²)

¹Other cause of death, ²Cigarettes smoked per day.

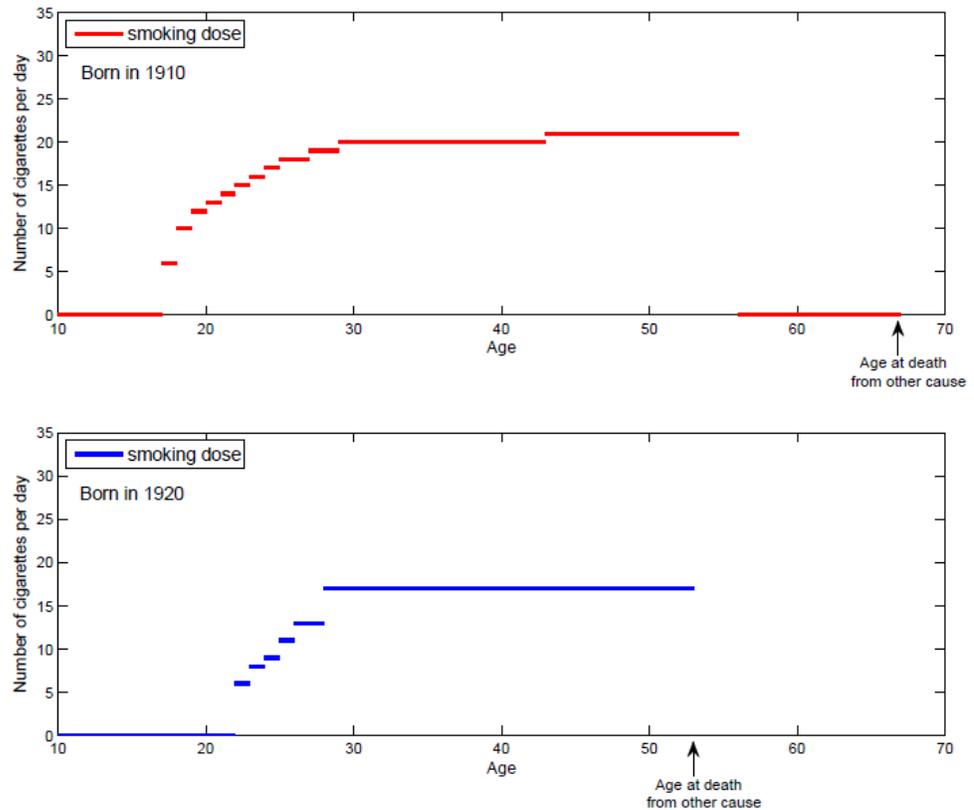


Figure SHG-1: Examples of the SHG-Generated Events

Simulation results by the SHG can be formatted in four different ways:

1. Text (formatted, human readable text depicting smoking history);
2. Tab Delimited Data (plain text, suitable for post-processing);
3. Annotated text-based timeline (visual representation in text);
4. XML (plain text, suitable for parsing). The outputs from the SHG are made up of individual life histories, each of which includes the following variables: birth year, age of smoking initiation, the corresponding smoking intensity (CPD) by age, age of smoking cessation, and age at death from causes other than lung cancer, conditioned on smoking histories.

REFERENCES:

-
- ¹ Matsumoto M., Nishimura T. "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator." in ACM Transactions on Modeling and Computer Simulation 1998; 8: 1: 3-30
-



POPULATION COMPONENT

SUMMARY

We use individual level simulation of the full US population stratified by single year age and birth year, and gender.

OVERVIEW

The purpose of the population component was to model smoking and other cause mortality in the US. The Smoking History Generator provided simulated individual histories, including birth year, age and smoking intensity at start of smoking, and ages and intensities at subsequent ages when smoking habits changed or when smoking stopped, and projected age of other cause death.

QUANTITATIVE DESCRIPTION

We use individual-based microsimulation of the full US population stratified by single year age and birth year, and gender. These smoking histories were sampled and used to build up a synthetic population that matched the full US population binned by single year ages 30-84 and calendar years 1975-2000. The TSCE model (a Markov transition model at the cellular level) was applied to each of the time-dependent smoking histories to estimate lung cancer deaths. This was combined with a statistical model representing additional age, period, and birth cohort effects, based on the stratification of the US population by age and calendar year.

POPULATION DYNAMICS

The TSCE model relates smoking to cell dynamics from birth to death for each individual in the population. However, changes in lung cancer over time in the US population were not fully captured by the TSCE model. Thus we made a second calibration by estimating additional period and birth cohort effects. These factors may adjust for model misspecification, effects of changing cigarette composition, changing demographics, and other exposures and environmental factors.

RECURRENCE

The FHCRC model represents lung cancer mortality, not detection or recurrence.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)



NATURAL HISTORY COMPONENT

SUMMARY

The FHCRC lung model represents lung cancer development as a stochastic (two-stage) cellular process that is influenced by an individual's smoking history.

OVERVIEW

The natural history component relates the probability of lung cancer mortality to each individual's smoking history using the biologically-motivated TSCE model. In the TSCE model, smoking influences the carcinogenic process throughout life, with the model representing survival until occurrence of a first malignant cell. A lag time is used to represent time from first malignant cell to lung cancer death ([Survival Mortality Component](#)). The natural history model builds on the ([Population Component](#)) that simulates the full US population based on simulated individual smoking histories and date of other cause mortality. The natural history provides estimates of US lung cancer mortality by age, gender and calendar year. These estimates from the natural history component are further adjusted by calendar year and birth cohort ([Survival Mortality Component](#)).

DISEASE STAGES

The TSCE model represents carcinogenesis as a lifelong process consisting of initiating mutations, clonal expansion of initiated cells, and malignant conversion. In the TSCE model, malignant conversion is defined as occurrence of the first malignant cell. The probability of lung cancer is related to the evolving joint probability distribution of these different cells throughout life. There is no explicit definition of disease stages in this model.

DISEASE GROWTH

We model the probability distribution for discrete cells as described above, allowing almost continuous growth of the intermediate lesions that consist of initiated cells.

STAGE TRANSITION TRENDS

The cell birth, death, and mutation in the carcinogenic process naturally give rise to changing transition rates because any cell in the increasing mass of intermediate cells can mutate to give rise to cancer. Thus in general, the effective rates increase with age, even if the rates for a single cell are held constant.

REGRESSION

The model allows for disease regression as pre-malignant clones may become extinct through random process of cell death.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)



SURVIVAL MORTALITY COMPONENT

SUMMARY

This document describes how survival and mortality are modeled.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

OVERVIEW

The TSCE model (See [Natural History Component](#)) represents cellular dynamics during carcinogenesis until occurrence of the first malignant cell. Thus it is necessary to include a lag time, or distribution of lag times, to represent the time from first malignant cell to lung cancer death. Further adjustments to lung cancer deaths in the US population are modeled as calendar year and birth cohort effects.

SURVIVAL ESTIMATION COVARIATES

Survival and mortality depend on disease progression according to the TSCE natural history model, combined with a statistical model representing additional age, period, and birth cohort effects.

OTHER CAUSE MORTALITY

The Smoking History Generator was developed to reflect historical smoking trends and rates for other cause mortality in the US population. The FHCRC lung cancer model used the simulated individual smoking histories including the age of other cause death provided by the SHG.



OUTPUT OVERVIEW

SUMMARY

This document describes the types of outputs generated by the FHCRC lung cancer model.

- Readers Guide
- Model Overview
- Assumption Overview
- Parameter Overview
- Component Overview
- Output Overview
- Results Overview
- Key References

OVERVIEW

The FHCRC lung cancer model provides estimates of lung cancer deaths in the US population by gender, age, and calendar year, given historical smoking patterns or alternative (counter-factual) smoking scenarios.

OUTPUT LISTING

1. **Project One** : Biological parameters and dose-response parameters related to smoking exposure in the TSCE model (See [Table3](#))

- back ground initiation rate
- back ground division rate of an initiated cell
- back ground net cell proliferation rate of an initiated cell
- back ground malignant conversion rate
- two parameters regarding dose-response effect for the division rate and the net cell proliferation rate
- two parameters regarding dose-response effect for the malignant conversion rate
- lag time parameter for the progression (assumed 5 years)

These estimates for pre-malignant cellular kinetics relate directly to the commonly reported lead time from smoking exposure to lung cancer, and to the growth rate of pre-malignant lesions. Validation of these rates included comparison with calibration to other cohorts. These outputs were the basis for estimating the relation between smoking and lung cancer in the simulation of the US population ([Project Two](#)).

2. **Project Two** : Parameters for secular time trends in both TSCE-PC model and TSCE-APC model

- calendar year effects (calendar years from 1975 to 2000) ([Figure2](#) and [Figure3](#))
- birth cohort effects (birth cohorts 1890-1894,1895-1899,.....,1960-1964,1965- for AC) ([Figure4](#) and [Figure5](#))
- additional age effects in the TSCE-APC model ([Figure6](#))

3. **Lung Smoking Base Case**: Model outputs in both TSCE-PC model and TSCE-APC model



- Age-standardized lung cancer mortality rates (using the 2000 US standard population, Census P25-1130) by calendar year for US males and US females for three different smoking scenarios (ATC, NTC, CTC) (For ATC case, see [Figure7](#) and [Figure8](#). Figures for NTC and CTC cases will be available in the forthcoming paper)
- Avoided lung cancer deaths resulted from the actual tobacco control (# lung cancer deaths in NTC – # lung cancer deaths in ATC) (Figures will appear in the forthcoming paper)
- Avoidable lung cancer deaths by assuming all smoking stops in 1965 (# lung cancer deaths in ATC – # lung cancer deaths in CTC) (Figures will appear in the forthcoming paper).

These estimates for lung cancer deaths, stratified by gender, age and calendar year, may be converted to stratum specific lung cancer rates by dividing by population counts within each stratum. Estimates of avoided and avoidable deaths were compared against historical outcomes for lung cancer mortality in the US to provide some indication of the benefit of the US Surgeon General's warnings about the dangers of smoking, and also an upper limit to the potential benefits of additional efforts at reducing tobacco consumption.

CATEGORIES

[Core Docs](#)



RESULTS OVERVIEW

SUMMARY

This document summarizes results from calibration of the FHCRC lung cancer model to cohort data ([Project One](#)) and application to understand the effects of smoking on lung cancer mortality in the US population ([Project Two](#)).

OVERVIEW

The FHCRC lung cancer model used the TSCE model incorporating the mechanisms of initiation, promotion and malignant conversion in carcinogenesis to analyze lung cancer mortality in five large cohorts; the British Doctors's, CPS I and II, HPFS and NHS cohorts ([Project One](#)). The parameter estimates from calibration to lung cancer mortality using these cohorts are closely tied to the model purpose of understanding the underlying biological mechanisms that relate tobacco smoke to lung cancer in the US population ([Project Two](#)).

By fitting the model to these cohorts data, we estimated the biological parameters related to age-specific cancer rates and dose-response parameters related to smoking exposure. The key biological parameters are the rate of initiation, the rates of cell division and apoptosis/differentiation of initiated cells, and the rate of malignant conversion of initiated cells. Progression from the first appearance of malignancy to death from lung cancer is modeled as a constant lag time. And the effect of smoking habits on age-specific lung cancer mortality is modeled via a dose-response function on each of these parameters. Based on the likelihood approach, we chose the most parsimonious model which is consistent with these cohorts data, and summarized the list of parameters in the final model here.

We began the model calibration to each cohort assuming that smoking could influence promotion, initiation, or malignant conversion. This is called the full model. After optimization of the full model, we progressively eliminated parameters that did not significantly contribute to the likelihood. This is called the reduced model (refer to [Table1](#)).

We found that the contemporaneous CPS-I and British doctors cohorts could be fit with all but one parameter in common. The CPS-II cohort was followed about 20 years later, and represented individuals smoking newer cigarettes. The CPS-II cohort dose-



response differs from the earlier cohort by having no significant effect of cigarette smoking dose-response on initiation, but a larger dose-response on promotion than found in the CPS-I cohort (for details, refer to the paper by Hazelton et al. 2005, See [Table2](#)). The HPFS and NHS cohorts are in the similar period as the CPS-II cohort, and we found no significant effect of smoking dose-response on initiation, which is consistent with the CPS-II cohort. However, we found significant effect of smoking dose-response on promotion as well as malignant conversion in the HPFS and NHS lung cancer mortality cohorts (refer to [Table3](#)).

The FHCRC lung cancer model combines the TSCE model, which was calibrated to the lung cancer mortality in the HPFS cohort for males and the NHS cohort for females, with US population-based adjustments for secular time trends (period and birth cohort in the TSCE-PC model, and additional age, period and birth cohort in the TSCE-APC model). The NCI's SHG can simulate individual-level smoking histories and generate cohorts of individuals for different smoking scenarios. We use the FHCRC lung cancer model to predict lung cancer mortality for various smoking scenarios using the cohorts generated by the SHG.

Three smoking scenarios are explored in the Lung Smoking Base Case study.

1. Actual tobacco control (ATC): Due to the increasing information about the harmful effects of smoking on the public health, including the US Surgeon General's report about the risk of smoking around 1964, smoking habits have been changed in the US over last several decades starting in the early 1950s. The SHG generated individuals with smoking histories mimicking actual smoking trends in the US population.
2. No tobacco control (NTC): To investigate the effects of tobacco control on lung cancer mortality, individuals with comparative smoking histories by assuming that no tobacco control occurred were generated by the SHG.
3. Complete tobacco control (CTC): To explore the maximum potential benefits from tobacco control on lung cancer mortality, the SHG generated individual histories that all smokers quit smoking in 1965 and no individuals began smoking after that.

There are two separate cohorts generated by the SHG and we present the results for both cohorts.

1. Empirical Cohorts (EC): Starting with birth cohort 1900, based on empirical data
2. All Cohorts (AC): Starting with birth cohort 1890, involves extrapolated smoking history data

RESULTS LIST

PROJECT 1 -TSCE SMOKING NATURAL HISTORY MODEL OUTPUT

During the calibration of the TSCE natural history model to different lung cancer mortality cohorts [See [Project One](#)], we found that smoking exposure tends to increase



rates of mutation, and more importantly, smoking significantly increases the growth (promotion) of initiated cells, leading to increased net cell proliferation rates of pre-malignant cells, and a subsequent rise in risk for lung cancer.

The FHCRC group chose to use the lung cancer mortality calibration of the TSCE model to the HPFS cohort for males and the NHS cohort for females that were performed as the first project [[Project One](#)]. This TSCE model calibration was used as the primary component in constructing the FHCRC lung cancer model. The HPFS (1986-2002) and NHS (1976-2000) cohorts calibration was chosen on the basis that these cohorts may take account for the risks from current cigarette compositions, and may well represent the contemporary US lung cancer trend and demographics for the Lung Smoking Base Case, which are ranged over the period 1975-2000. Furthermore these cohorts have information for former smokers as well as never and current smokers, with extensive cross tabulation of lung cancer deaths and population at risk by gender, race, age, duration of smoking, and smoking rate. Because of the explicit information for former smokers, these cohorts may be adequate to reflect the impact of quitting smoking on lung cancer mortality. To compensate for possible limitations of the cohort-calibrated TSCE model, the FHCRC group introduced additional age, period, and birth cohort effects when calibrating to lung cancer mortality in the US population.

PROJECT 2 - OUTPUTS OF ADDITIONAL AGE, PERIOD, AND BIRTH COHORT ADJUSTMENTS IN LUNG SMOKING BASE CASE

For the Lung Smoking Base Case (see [Project Two](#)) the expected number of lung cancer deaths were calculated from the cohort-calibrated TSCE model, using the individual histories generated by the NCI's SHG that were selected to fill up the full US population table by single year of age from 30 to 84, and calendar years from 1975 to 2000. The outputs from the additional age, period, and birth cohort adjustment consist of a set of parameters (see [Parameter Overview](#)) to optimize the FHCRC lung cancer model projections to observed US population counts and numbers of lung cancer deaths by gender and single year of age from 30 to 84, and calendar years from 1975 to 2000. The Lung Smoking Base Case [[Project Two](#)] study utilizes calibration to a lung cancer mortality cohort [[Project One](#)] that includes background, dose-response, and lag time parameters relating to the effects of smoking to lung cancer mortality. These parameters are used in calibrating to US population and lung cancer mortality data. Outputs from this process include projections of lung cancer mortality by single years for ages 30-84, and calendar years 1975-2000. In the Lung Smoking Base Case, we predict the number of avoided deaths from lung cancer under the actual tobacco control scenario compared to no tobacco control scenario, and also the potential avoidable deaths from lung cancer if all smokers quit smoking in the year 1965, and no one started smoking after that. The explicit results from the Lung Smoking Base Case will appear in the forthcoming paper.

CATEGORIES

[Core Docs](#)



PROJECT ONE

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

The first project consisted of calibrating a natural history model to substantial US smoking cohorts to model the probability of lung cancer death at different ages in terms of an individual's smoking behavior up to that age. The two stage clonal expansion (TSCE) model was chosen to represent the natural history of lung cancer (Moolgavkar and Venzon, 1979; Moolgavkar and Knudson, 1981). The TSCE model is a stochastic model that represents the processes of initiation, promotion, and malignant conversion during carcinogenesis. The TSCE model parameters were calibrated to lung cancer mortality among never and current smokers in the American Cancer Society CPS-I and CPS-II cohorts and also the British Doctors cohort from the UK (Hazelton et al., 2005). We also calibrated the model to lung cancer incidence (Meza et al., 2008) and mortality in the prospective Health Professionals Follow-up Study (HPFS) and the Nurses' Health Study (NHS) cohorts. These calibrated natural history model parameters were shared with other CISNET lung cancer modeling groups for use in modeling US lung cancer mortality. The TSCE model provided excellent fits to the data representing smoking histories for individuals in the different cohorts. The FHCRC lung group's use of these natural history parameters is discussed in more detail below.

The natural history model calibrations to different cohorts allowed us to compare estimates for the effects of tobacco on lung cancer through analysis of incidence (HPFS and NHS) or mortality (CPS-I, British Doctors, CPS-II, HPFS and NHS). Follow-up for the CPS-I and British Doctors cohorts occurred about 20 years prior to CPS-II, HPFS and NHS, allowing us compare the effects of earlier versus later cigarette compositions. In general, all models indicated that the most important dose-response effect of tobacco smoke is on promoting the growth (increasing the clonal expansion rate) of pre-malignant cells. This promotion effect is slightly stronger in the CPS-II, HPFS, and NHS analysis with the more recent (lower tar and nicotine) cigarette compositions. Lung cancer risk is also slightly increased by an influence of smoking on initiation, but this effect is insignificant in the newer cigarettes.

Other factors outside the natural history model's domain influence the lung cancer mortality rates in the US population. First, individuals in the US smoking cohorts are not fully representative of the US population. Second, the composition of cigarettes has changed substantially over time. Third, other exposures and environmental factors may contribute to lung cancer risk. In addition, the TSCE model consists of only two stages, whereas the lung cancer process is biologically complex, and progresses along many pathways. Thus, the TSCE natural history model may not be capable of fully capturing the effects of smoking on the subsequent risk of lung cancer. Therefore, in addition to the biologically based smoking model, we introduce additional age, period, and birth cohort effects to adjust for these other factors when calibrating the FHCRC lung cancer model to US lung cancer mortality data.

See Also: [Project One](#), [Project Two](#)



PROJECT TWO

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

Each CISNET lung cancer modeling group has developed age-specific models of US lung cancer mortality trends for ages 30-84 and calendar years 1975-2000. This comparative modeling effort of different groups is called the Lung Smoking Base Case. The FHCRC lung cancer modeling group chose to use the TSCE natural history model, calibrated to relate smoking histories to lung cancer mortality, as the central element in addressing the Lung Smoking Base Case. However, the FHCRC group found that additional age, period, and birth cohort effects were needed to account for changes in cigarette composition over time, other factors that contribute to lung cancer, and limitations of the TSCE natural history model. This combination of biologically based modeling, combined with statistical adjustments for additional age, period, and birth cohort, constitutes the FHCRC lung cancer model. The FHCRC lung cancer model was subsequently used to evaluate the impact of public health messages on lung cancer trends and life years lost (or gained) under alternative tobacco consumption scenarios.

Smoking is the most significant risk factor in modeling US lung cancer trends. The Smoking History Generator (SHG) provided by the National Cancer Institute (NCI) was used to generate individuals with simulated smoking histories. We developed software to sample and combine these simulated individuals to imitate the full US population table for each calendar year (ranging from 1975 to 2000) and age bin (ranging from 30 to 84) separately for males and females. Then the cohort-calibrated TSCE natural history model was applied to each of these simulated individuals with smoking histories to calculate the expected lung cancer deaths in each cell of the US population table for males and females. However, comparison of the number of expected lung cancer deaths against the number of observed lung cancer deaths in each cell revealed discrepancies and apparent age, period, and birth cohort trends are not fully accounted for by the model.

Factors other than smoking influence US lung cancer mortality trends. Although the TSCE natural history model accounted for most lung cancer deaths observed in the US, additional age, period, and birth cohort factors were required to accurately represent the detailed lung cancer mortality outcomes in the US population. Thus, we did further calibration by applying additional age, period, and birth cohort effects to the modeled US lung cancer mortality rates by the TSCE natural history model to match the observed US lung cancer mortality rates.

IMPACT OF PUBLIC HEALTH MESSAGES ON US LUNG CANCER MORTALITY

It is generally thought that increasing US lung cancer trends slowed, and even decreased, in response to increasing public awareness during the 1960's and later about the dangers of smoking (Irvine et al., 2006; Musk et al., 2003). This public knowledge came from many sources, including statements of the US Surgeon General about the risk of smoking (Parascandola et al., 2001, 2006).

See Also: [Project One](#), [Project Two](#)

TSCEMODEL DETAILS

The TSCE model is a mathematical model that represents the carcinogenic process by tracking the probability distribution for the number of cells on the pathway to cancer. The model assumes that any of the normal stem cells in lung may undergo, at random, a first mutation step (called initiation) at rate $\nu(t)$ during the course of cell division to create an initiated cell. Each initiated cell may undergo cell division at rate $\alpha(t)$ or cell death at rate $\beta(t)$. A random second mutation event may occur at rate $\mu(t)$ as any of the initiated cells undergo cell division, producing a malignant cell. After occurrence of the first malignant cell, a lag time is used to represent the time between the appearance of the first malignant cell and lung cancer mortality (See [Figure1](#)).

Let $d(t)$ be the exposure dose to smoking at age t . Then we assume that initiation, promotion, and malignant conversion rates may be altered during periods of exposures through flexible dose-response relationships:

$$\theta(d(t)) = \theta_0(1 + \theta_c d(t)^{\theta_e}),$$

where θ represents a biological parameter in the model, θ_0 is the background parameter, and θ_c and θ_e are the dose-response coefficients corresponding to smoking. Closed form expressions for the hazard and survival function of the TSCE model are known in the case of piecewise constant parameters (Heidenreich et al., 1997).

$$S_2(t) = \exp \left\{ \sum_{j=1}^n \frac{\nu_j X}{\alpha_j} \ln \left(\frac{q_j - p_j}{f_j(t_{j-1}, t_n)} \right) \right\},$$

$$h_2(t) = \sum_{j=1}^n \frac{\nu_j X}{\alpha_j} \frac{\partial}{\partial t_n} \ln (f_j(t_{j-1}, t_n)) = \sum_{j=1}^n \frac{\nu_j X}{\alpha_j} \frac{1}{f_j(t_{j-1}, t_n)} \frac{\partial}{\partial t_n} f_j(t_{j-1}, t_n),$$

where X is the number of normal stem cells, n is the number of age-periods with different parameter values before age $t_n \equiv t; [t_{j-1}, t_j], j = 1, \dots, n$ denote the end-points of the j -th age-period, d_j is the smoking-dose during the j -th age-period, $t_0 = 0$, and $\nu_j, \alpha_j, g_j, \mu_j$ denote the parameter values during the j -th age-period, and

$$g_j = g(1 + g_c d_j^{g_e}), \quad \alpha_j = \alpha(1 + g_c d_j^{g_e}),$$

$$\nu_j = \nu(1 + \nu_c d_j^{\nu_e}), \quad \mu_j = \mu(1 + \mu_c d_j^{\mu_e}),$$

$$p_j, q_j = \frac{1}{2} \left(-g_j \mp \sqrt{g_j^2 + 4\alpha_j \mu_j} \right).$$

$$\tilde{y}_n = 0, \quad \tilde{y}_{j-1} = \frac{\alpha_{j-1} (\tilde{y}_j - p_j) q_j e^{q_j(t_{j-1}-t_j)} + (q_j - \tilde{y}_j) p_j e^{p_j(t_{j-1}-t_j)}}{\alpha_j f_j(t_{j-1}, t_n)},$$

$$f_j(t_{j-1}, t_n) = (\tilde{y}_j - p_j) \exp \{q_j(t_{j-1} - t_j)\} + (q_j - \tilde{y}_j) \exp \{p_j(t_{j-1} - t_j)\},$$

$$\frac{\partial}{\partial t_n} f_n(t_{n-1}, t_n) = [\exp \{q_n(t_{n-1} - t_n)\} - \exp \{p_n(t_{n-1} - t_n)\}] p_n q_n,$$

$$\frac{\partial}{\partial t_n} f_j(t_{j-1}, t_n) = [\exp \{q_j(t_{j-1} - t_j)\} - \exp \{p_j(t_{j-1} - t_j)\}] \frac{\partial}{\partial t_n} \tilde{y}_j,$$

$$\frac{\partial}{\partial t_n} \tilde{y}_{n-1} = \frac{\alpha_{n-1}}{\alpha_n} \frac{(q_n - p_n)^2 e^{(p_n + q_n)(t_{n-1} - t_n)}}{(f_n(t_{n-1}, t_n))^2} p_n q_n,$$

$$\frac{\partial}{\partial t_n} \tilde{y}_{j-1} = \frac{\alpha_{j-1}}{\alpha_j} \frac{(q_j - p_j)^2 e^{(p_j + q_j)(t_{j-1} - t_j)}}{(f_j(t_{j-1}, t_n))^2} \frac{\partial}{\partial t_n} \tilde{y}_j.$$

These equations may be used to calculate the TSCE model survival $S_2(u)$ and hazard $h_2(u)$ at any time u .

If we assume a constant or gamma lag time between the appearance of the first malignant cell and lung cancer death, the survival probability at age t of an individual with smoking history d_j , $S(t; \bar{\theta}(d_j))$, is given by

$$S(t; \bar{\theta}(d_j)) = \begin{cases} S_2(t - t_{lag}; \bar{\theta}(d_j)) & \text{if lag time is constant,} \\ 1 - \int_0^t (1 - S_2(u; \bar{\theta}(d_j))) g(t - u) du & \text{if lag time is gamma distributed,} \end{cases}$$

where $\bar{\theta}(d_j)$ denotes the vector of identifiable model parameters given the smoking history d_j and $g(\cdot)$ is the gamma density. We calculate the probability for lung cancer mortality in each single year of age for the individual, given the gender and the full smoking history of the individual. During calibration to smoking cohort data, the study follow-up times and known outcome for each individual (death from lung cancer, or study censoring) were combined with the model probabilities for individual death from lung cancer at each age to form an individual likelihood.

The individual likelihood $\mathcal{L}_j = \mathcal{L}_j(t_j, s_j; \bar{\theta}(d_j))$ depends on time of entry into the study s_j , censoring or failure time t_j , and on detailed smoking exposure histories in conjunction with general dose-response models for the biological parameters in the TSCE model, and on the lag time or lag time distribution. We assume that each individual is lung cancer free at the beginning of the study, s_j . The individual likelihoods for cases and survivors, including left truncation, are given by

$$\mathcal{L}_j(t_j, s_j; \bar{\theta}(d_j)) = \begin{cases} -S'(t_j; \bar{\theta}(d_j))/S(s_j; \bar{\theta}(d_j)) & \text{if death from lung cancer,} \\ S(t_j; \bar{\theta}(d_j))/S(s_j; \bar{\theta}(d_j)) & \text{otherwise,} \end{cases}$$

where the prime denotes derivative with respect to t .

Assuming independence between individuals, the cohort likelihood is the product of individual likelihoods over all subjects j ,

$$\mathcal{L} = \prod \mathcal{L}_j(t_j, s_j; \bar{\theta}(d_j)).$$

Gradient search methods (Bhat FORTRAN software library, Luebeck 2009) were then used to maximize the likelihood, leading to a set of model parameters that relate individual histories to the probability of death from lung cancer at each age.

TSCE SMOKING NATURAL HISTORY MODEL INPUTS

Individual histories include gender, age at start smoking (if a smoker), beginning smoking rate (number of cigarettes per day), age at each change in smoking habit, and smoking rate during each of these periods, the age at quit smoking (if that occurs), age at entry into the study, age at lung cancer death or end of study follow-up. These individual history inputs were input from the cohort records during calibration to smoking cohort data [see [Project One](#)].

The TSCE-PC and TSCE-APC model calibrations utilized histories generated by the NCI provided Smoking History Generator (SHG) to fill up the full US population tables by gender and single year of age from 30 to 84, and calendar years from 1975 to 2000. The TSCE natural history model used the simulated smoking history inputs for individuals that contributed to each cell of the simulated US population table to calculate the expected number of lung cancer deaths by single year of age and calendar year.

INPUTS FOR MODEL CALIBRATION IN LUNG SMOKING BASE CASE

Assume lung cancer mortality data is in tabular form for $i = 1, \dots, I$ age groups and $j = 1, \dots, J$ calendar years. For age group i , the number of lung cancer deaths during calendar year j can be assumed to follow a Poisson distribution with mean $\Lambda_{i,j}$.

We consider two calibration approaches:

1. Age-Period-Cohort model (TSCE-PC model):

$$\Lambda_{i,j} = PY_{i,j} h(a_i) b_{i,j} c_j,$$

where a_i is the mean of the i -th age group, $b_{i,j}$ and c_j are coefficients that adjust for birth cohort and calendar year (period) effects, respectively, $PY_{i,j}$ is the person years at risk, and $h(a_i)$ represents the hazard function of the TSCE natural history model with lag time evaluated at age a_i .

2. Age-Age-Period-Cohort model (TSCE-APC model):

$$\Lambda_{i,j} = PY_{i,j} h(a_i) e_i b_{i,j} c_j,$$

where e_i is the coefficient that adjust for additional age effect for the i -th age group, and other variables are the same as above.

The overall likelihood \mathcal{L} for the observed lung cancer mortality in all age-calendar year groups is given by

$$\mathcal{L} = \prod_{i,j} \frac{\Lambda_{i,j}^{O_{i,j}} e^{-\Lambda_{i,j}}}{O_{i,j}!},$$



where $O_{i,j}$ is the number of lung cancer deaths in the i - th age group during calendar year j .

The parameters in the TSCE natural history model were estimated by calibrating to lung cancer mortality cohorts [refer to the [Project One](#)]. And these estimated parameters are used in the calculation of $h(a_i)$. Thus the above likelihood is used to estimate the secular terms: period and birth cohort effects in the TSCE-PC model, and additional age, period and birth cohort effects in the TSCE-APC model. Note in these two models, 'TSCE' refers the age effect calculated from the TSCE model.

In the Lung Smoking Base Case study, the inputs for model calibration consist of US population counts and numbers of lung cancer deaths by gender and single year of age from 30 to 84, and calendar years from 1975 to 2000.



TABLE3

Table 2: Reduced model estimates for the HPFS and NHS cohorts

Parameter	p_1	p_2	p_3	p_5	p_6	p_{10}	p_{11}
Cohort	Background cell division rate (per cell per year)	Background net cell proliferation rate (per cell per year)	Background initiation and malignant conversion rate (per cell per year)	Tobacco promotion rate coefficient	Tobacco promotion rate power	Tobacco malignant conversion rate coefficient	Tobacco malignant conversion rate power
HPFS	3.000	0.106	7.311e-8	0.134	0.474	0.175	0.552
NHS	3.000	0.083	9.370e-8	0.189	0.491	0.087	0.711

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

FIGURE2

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

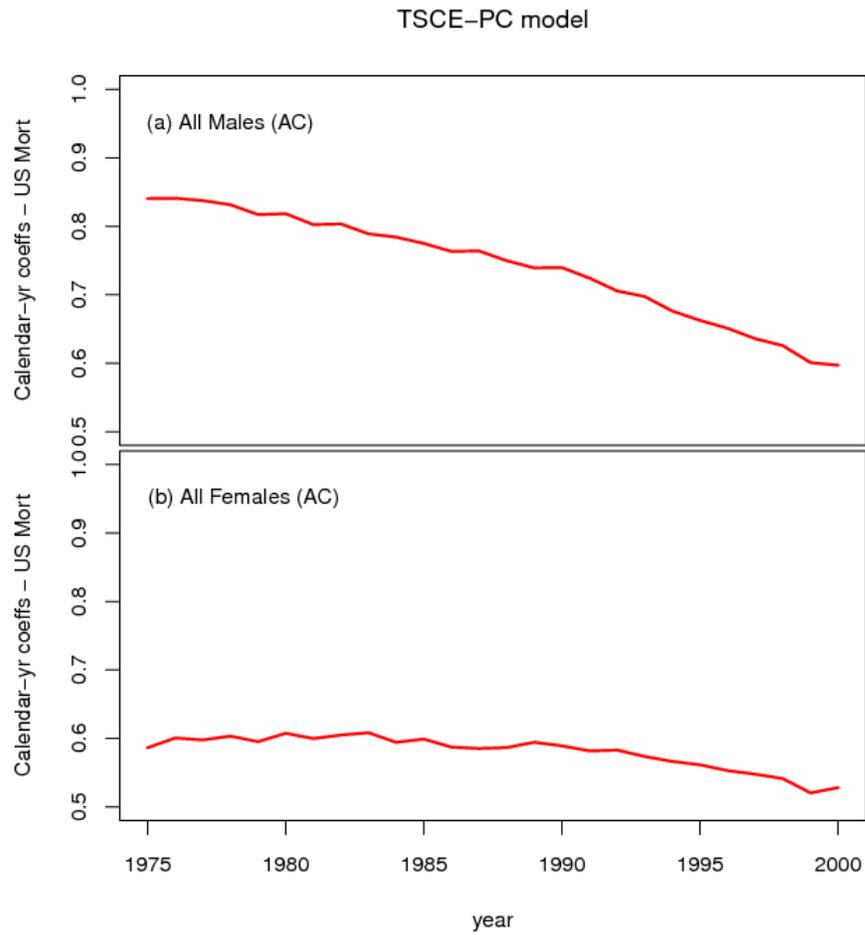


Figure 2: Calendar year effects (1975-2000). AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. Age-Period-Cohort model is used.

FIGURE3

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

TSCE-APC model

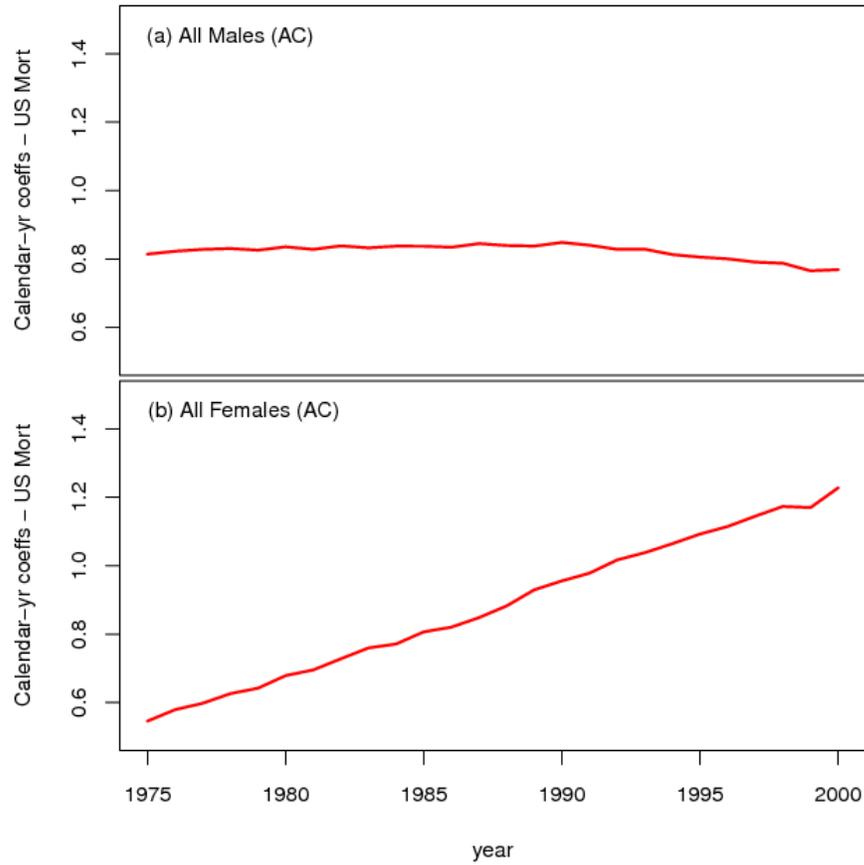


Figure 3: Calendar year effects (1975-2000). AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. Age-Age-Period-Cohort model is used.

FIGURE4

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

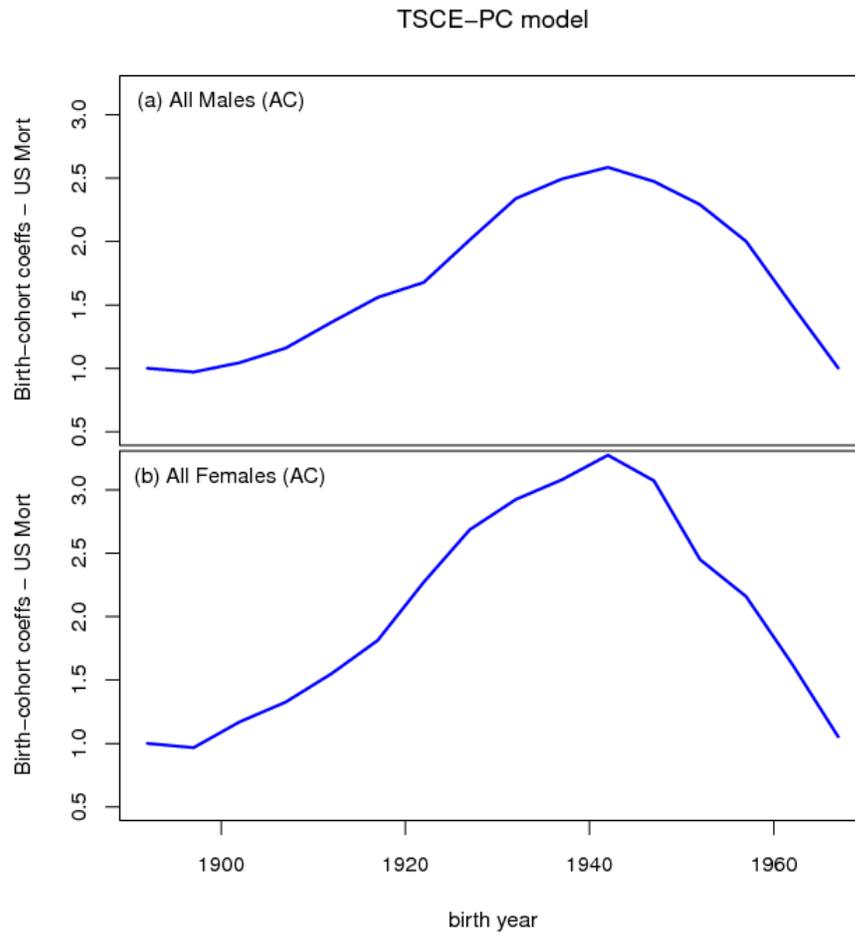


Figure 4: Birth cohort effects (birth cohorts 1890 – 1894, 1895 – 1899, ..., 1960 – 1964, >= 1965). AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. Age-Period-Cohort model is used.

FIGURE 5

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

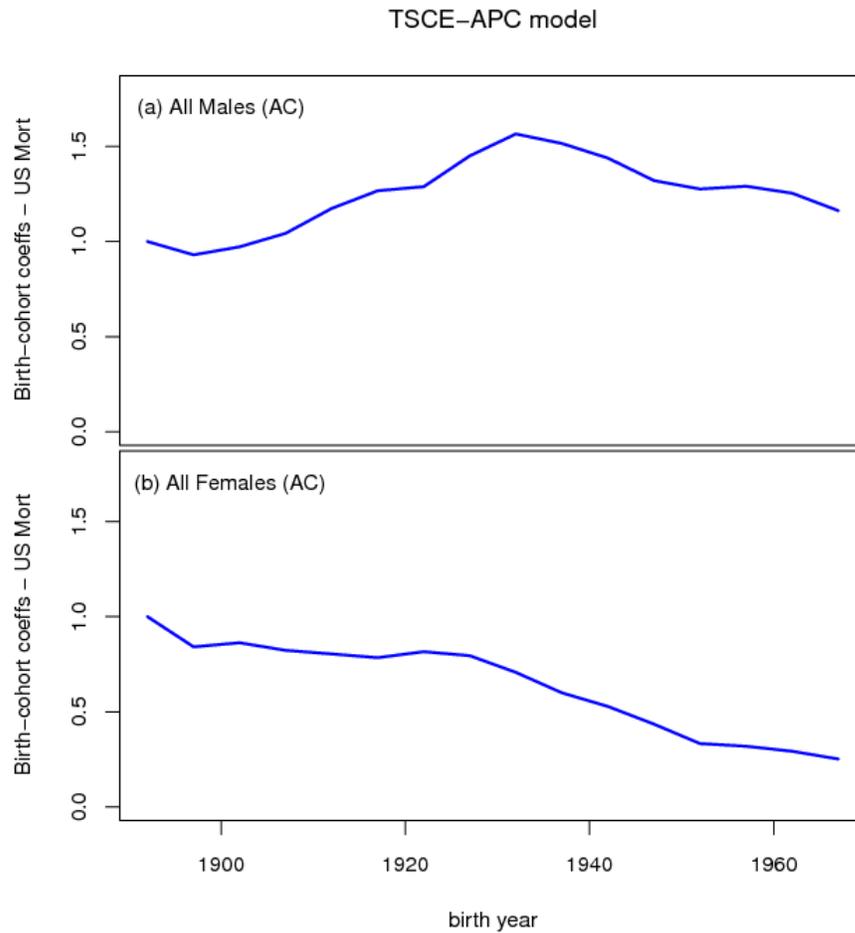


Figure 5: Birth cohort effects (birth cohorts 1890 – 1894, 1895 – 1899, ..., 1960 – 1964, >= 1965). AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. Age-Age-Period-Cohort model is used.

FIGURE6

TSCE-APC model

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

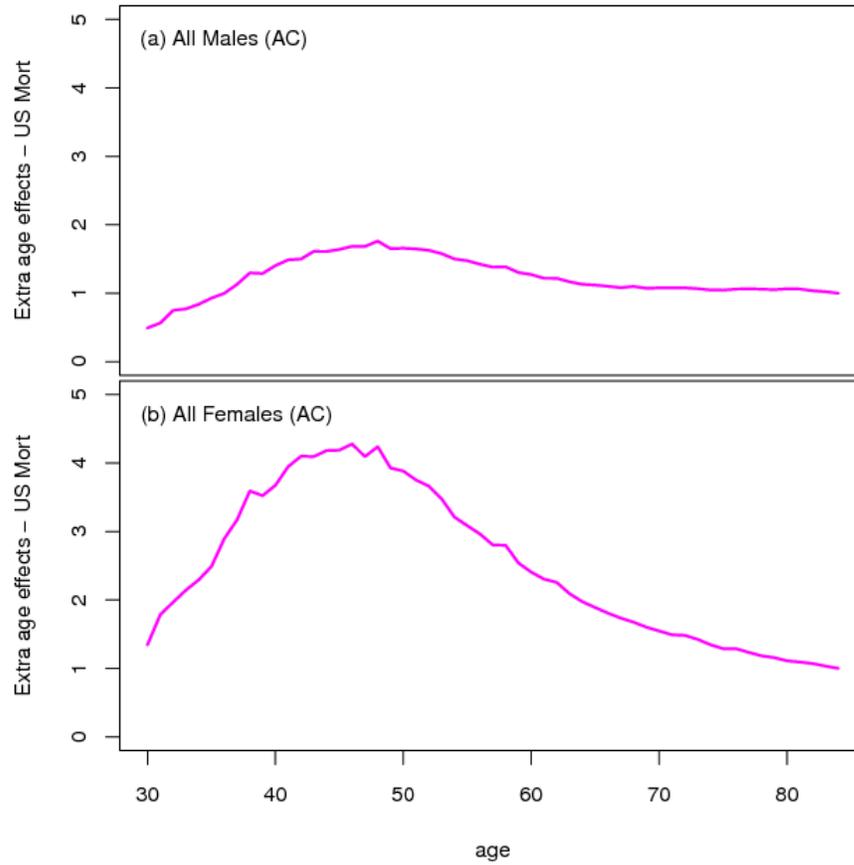


Figure 6: Additional age effects. AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. Age-Age-Period-Cohort model is used.

FIGURE 7

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

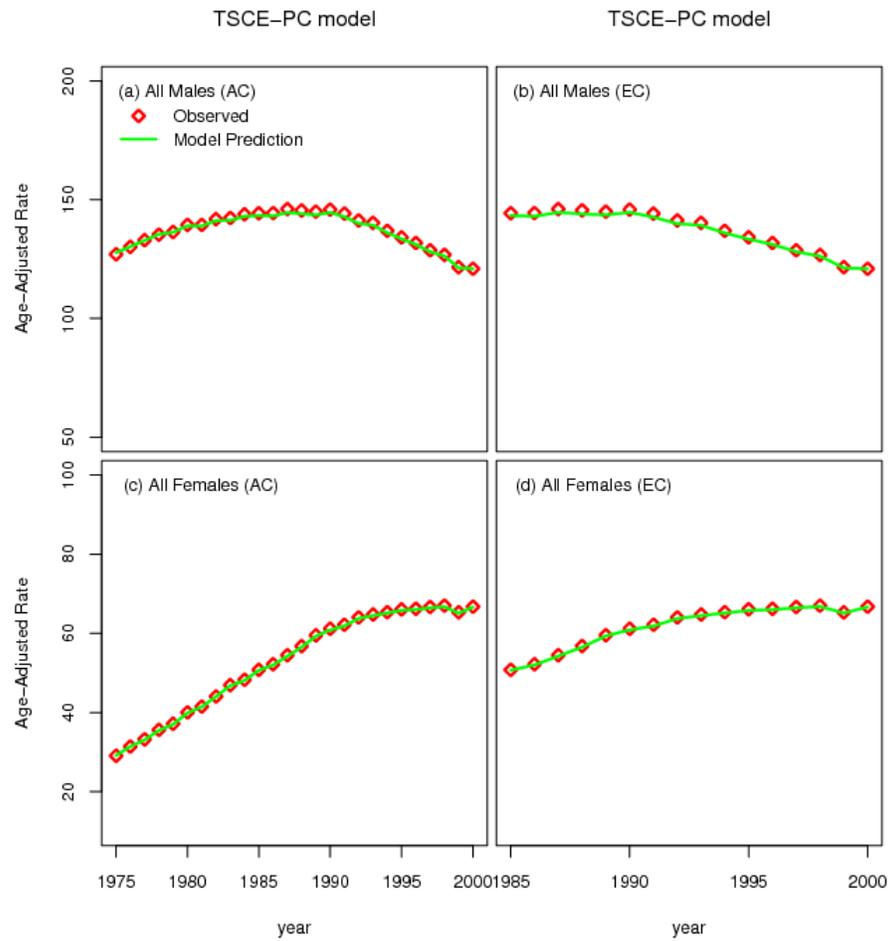


Figure 7: Age-standardized lung cancer mortality rates (using the 2000 US standard population, Census P25-1130) by calendar year. Diamond points represent the age-standardized lung cancer mortality rates in the observed US lung cancer mortality data, and the green line is the model prediction. AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. EC: empirical cohorts starting with birth cohort 1900, based on empirical data. Age-Period-Cohort model is used.

FIGURE 8

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

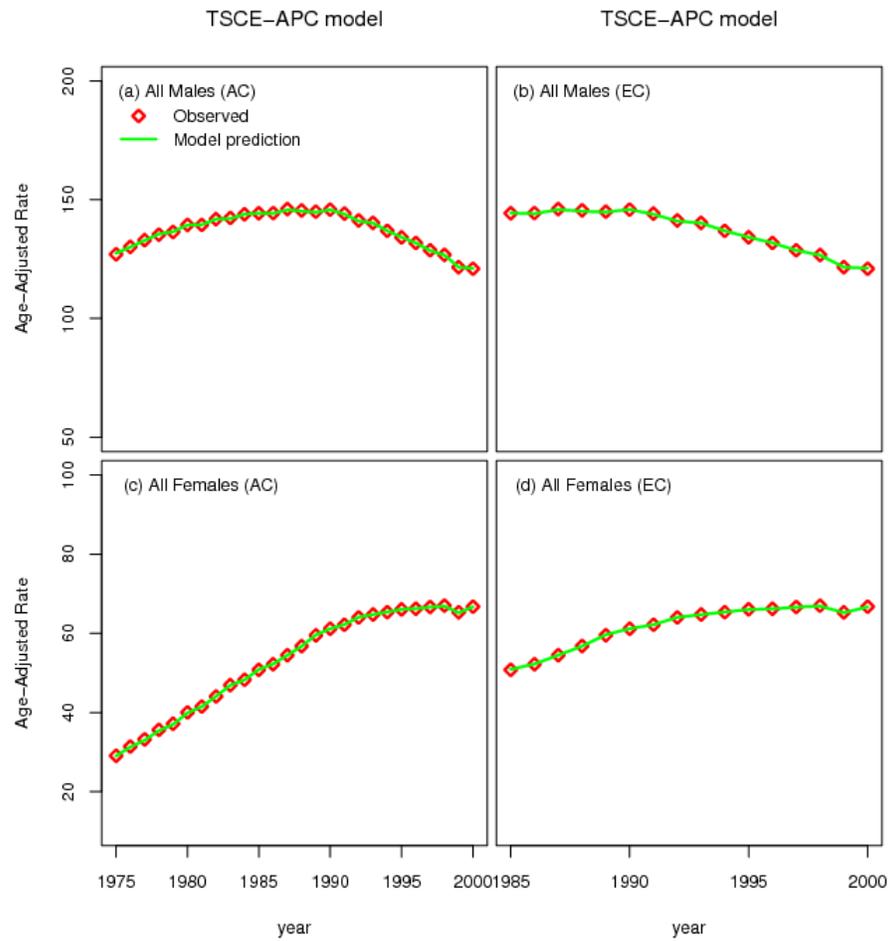


Figure 8: Age-standardized lung cancer mortality rates (using the 2000 US standard population, Census P25-1130) by calendar year. Diamond points represent the age-standardized lung cancer mortality rates in the observed US lung cancer mortality data, and the green line is the model prediction. AC: all cohorts starting with birth cohort 1890, involves extrapolated smoking history data. EC: empirical cohorts starting with birth cohort 1900, based on empirical data. Age-Age-Period-Cohort model is used.

TABLE 1

Table 1: Full model and Reduced model dose-response

- Readers Guide
- Model Overview
- Assumption Overview
- Parameter Overview
- Component Overview
- Output Overview
- Results Overview
- Key References

Full model.		
Initial scale:		
$X = 10^7$		Assume 10^7 normal stem cells in both lungs*
$v_0 = \mu_0$		Equate background initiation, malignant conversion rates*
Background variables for cohort		
Rate	Parameter	
α_0	p_1 (CPS-I)	Background or nonsmoking cell division rate* (per cell per year)
	p_1' (Brit. docs.)	
$g_0 = \alpha_0 - \beta_0 - \mu_0$	p_2	Background or nonsmoking net cell promotion rate (per cell per year)
$v_0 = \mu_0$	p_3	Background initiation, malignant conversion* (per cell per year)
Full model dose-response for cohort		
$v_i = v_0(1 + p_4 \text{dose}_i^{p_6})$		Initiation rate* (per cell per year)
$g_i = g_0(1 + p_5 \text{dose}_i^{p_6})$		Net initiated cell promotion rate (per cell per year)
$\alpha_i = \alpha_0(1 + (p_5 + p_9) \text{dose}_i^{p_6})$		Initiated cell division rate* (per cell per year)
$\mu_i = \mu_0(1 + p_{10} \text{dose}_i^{p_{11}})$		Malignant conversion rate* (per cell per year)
Fixed lag:		
		Gamma distribution with random variable X as lag time:
$t_{lag} = p_7$		$f(x, a, b) = \frac{1}{b\Gamma(a)} x^{a-1} e^{-x/b}$, $\mu_{lag} = p_{12} = a b$, $\sigma_{lag} = p_{13} = \sqrt{a b^2}$
Reduced model.		
X and background rates v_0, α_0, g_0 , and μ_0 are parameterized the same as in the full model above.		
Reduced model dose-response for the British Doctors', CPS I and II cohorts		
$v_i = v_0(1 + p_4)$, $p_4 = 0$ for nonsmokers		Initiation rate* (per cell per year)
$g_i = g_0(1 + p_5 \text{dose}_i^{p_6})$		Net initiated cell promotion rate (per cell per year)
$\alpha_i = \alpha_0(1 + p_5 \text{dose}_i^{p_6})$		Initiated cell division rate* (per cell per year)
$\mu_i = \mu_0$		Malignant conversion rate*, no dose-response (per cell per year)
Reduced model dose-response for the HPFS and NHS cohorts		
$v_i = v_0$		Initiation rate*, no dose-response (per cell per year)
$g_i = g_0(1 + p_5 \text{dose}_i^{p_6})$		Net initiated cell promotion rate (per cell per year)
$\alpha_i = \alpha_0(1 + p_5 \text{dose}_i^{p_6})$		Initiated cell division rate* (per cell per year)
$\mu_i = \mu_0(1 + p_{10} \text{dose}_i^{p_{11}})$		Malignant conversion rate* (per cell per year)
Fixed lag:		
$t_{lag} = 5$ years		

Note: A full model was first tested where initiation, promotion, and malignant conversion are assumed to be potentially modified by smoking through flexible dose-response parameterization of the form: $a(1 + bx^c)$. Likelihood ratio tests were used to eliminate variables that were not significant for any of the cohorts, resulting in the reduced model.

*Only the product vX is identifiable. Thus, without loss of generality, we assume $X = 10^7$ normal stem cells at risk. This number may be changed by reciprocal change in v . A separate identifiability condition relates variables α, μ , and v . Thus, without loss of generality, we set $\mu_0 = v_0$. These variables may be rescaled without changing the model fit.



TABLE2

Table 2: Reduced model estimates for CPS-I, CPS-II and British Doctors' cohorts [MLE (MCMC 95% CI)]

Parameter	p_1	p_2	p_3	p_4	p_5	p_6
Cohort	Background cell division rate (per cell per year)	Background net cell proliferation rate (per cell per year)	Background initiation and malignant conversion rate (per cell per year)	Tobacco initiation rate	Tobacco promotion rate coefficient	Tobacco promotion rate power
CPS-I males	22.65(17.35,31.31)	0.075(0.065,0.085)*	1.40e-7(1.12e-7,1.76e-7)*	1.79(1.11,3.07)*	0.21(0.15,0.28)*	0.47(0.42,0.54)*
British Doctors	5.87(2.66,12.89)	0.075(0.065,0.085)*	1.40e-7(1.12e-7,1.76e-7)*	1.79(1.11,3.07)*	0.21(0.15,0.28)*	0.47(0.42,0.54)*
CPS-II males	7.70(6.45,12.99)	0.090(0.071,0.106)	7.16e-8(4.60e-8,1.21e-7)	~ 0.00(~ 0.00,1.76)	0.60(0.43,0.91)	0.22(0.12,0.30)
CPS-I females	71.56(49.08,100.0)	0.086(0.073,0.101)	8.93e-8(6.50e-8,1.19e-7)	1.23(0.32,2.80)	0.04(0.02,0.07)	0.98(0.80,1.15)
CPS-II females	15.82(13.39,42.12)	0.071(0.055,0.088)	1.07e-7(6.97e-8,1.62e-7)	0.02(~ 0.00,12.50)	0.50(0.27,0.86)	0.32(0.14,0.40)

Note: Estimates for net cell proliferation rates are tightly constrained, whereas cell division rates are not due to compensatory cell death.
* Shared variables in joint fit to White males in CPS-I and British Doctors' cohorts.

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

FIGURE 1

Two-stage clonal expansion model

- [Readers Guide](#)
- [Model Overview](#)
- [Assumption Overview](#)
- [Parameter Overview](#)
- [Component Overview](#)
- [Output Overview](#)
- [Results Overview](#)
- [Key References](#)

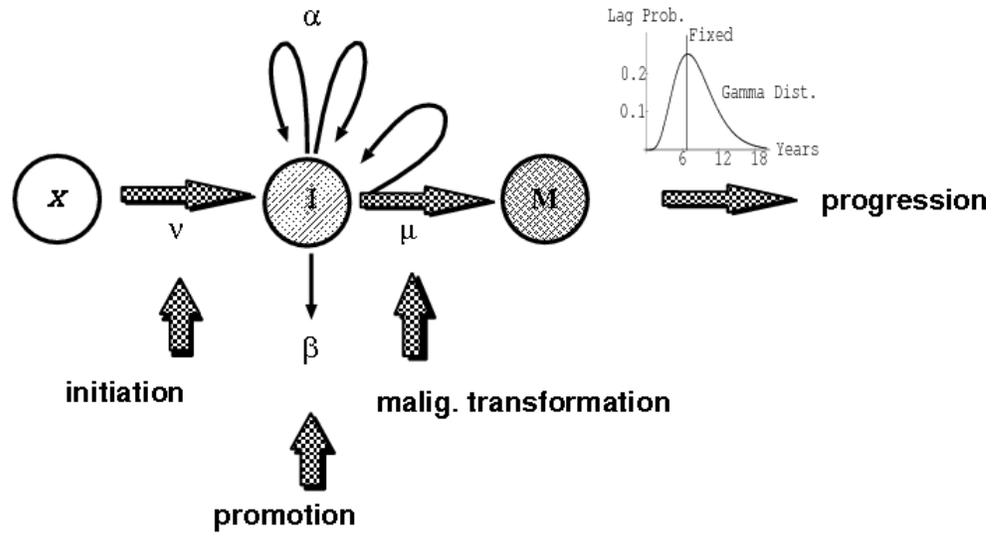


Figure 1: Two-stage clonal expansion (TSCE) model of lung cancer, including initiation, promotion, malignant conversion, and a lag time from first malignant cell to time of death from lung cancer.



KEY REFERENCES

[Readers Guide](#)
[Model Overview](#)
[Assumption Overview](#)
[Parameter Overview](#)
[Component Overview](#)
[Output Overview](#)
[Results Overview](#)
[Key References](#)

- Alberg AJ, Ford JG, Samet JM; American College of Chest Physicians. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* 2007 Sep;132(3 Suppl):29S-55S.
- Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer Statistics, 2008. *CA Cancer J Clin*. 2008 Mar-Apr;58(2):71-96.
- Stewart SL, Cardinez CJ, Richardson LC, Norman L, Kaufmann R, Pechacek TF, Thompson TD, Weir HK, Sabatino SA: Centers for Disease Control and Prevention (CDC). Surveillance for cancers associated with tobacco use: United States, 1999-2004. *MMWR Surveill Summ*. 2008 Sep 5;57(8):1-33.
- Coultas DB, Samet JM. Occupational lung cancer. *Clin Chest Med*. 1992 Jun;13(2):341-54.
- Moolgavkar SH, Knudson AG Jr. Mutation and cancer: a model for human carcinogenesis. *J Natl Cancer Inst*. 1981 Jun;66(6):1037-52.
- Stevens RG, Moolgavkar SH. Estimation of relative risk from vital data: smoking and cancers of the lung and bladder. *J Natl Cancer Inst*. 1979 Dec;63(6):1351-7.
- Moolgavkar SH, Venzon DJ. Two-event models for carcinogenesis: Incidence curves for childhood and adult tumors. *Mathematical Biosciences* 47:55-77, 1979.
- Irvine L, Elliott L, Wallace H, Crombie IK. A review of major influences on current public health policy in developed countries in the second half of the 20th century. *J R Soc Health*. 2006 Mar;126(2):73-8.
- Musk AW, de Klerk NH. History of tobacco and health. *Respirology*. 2003 Sep;8(3):286-90.
- Parascandola M, Weed DL, Dasgupta A. Two Surgeon General's reports on smoking and cancer: a historical investigation of the practice of causal inference. *Emerg Themes Epidemiol*. 2006 Jan 10;3:1.
- Parascandola M. Cigarettes and the US Public Health Service in the 1950s. *Am J Public Health*. 2001 Feb;91(2):196-205.
- Moolgavkar SH, Luebeck EG. Multistage carcinogenesis and the incidence of human cancer. *Genes Chromosomes Cancer*. 2003 Dec;38(4):302-6.
- Moolgavkar SH. Pollution analysis flawed by statistical model. *Nature* 2007 Jan 4;445(7123):21.
- Moolgavkar SH. Fine particles and mortality. Comment on: *Inhal Toxicol*. 2005 Dec 15;17(14):803-16. *Inhal Toxicol*. 2006 Jan;18(1):93-4.
- Thun MJ, Burns DM. Health impact of "reduced yield" cigarettes: a critical assessment of the epidemiological evidence. *Tob Control*. 2001;10 Suppl 1:i4-11. Erratum in: *Tob Control* 2002 Mar;11(1):85.
- Kabat GC. Fifty years' experience of reduced-tar cigarettes: what do we know about their health effects? *Inhal Toxicol*. 2003 Sep 15;15(11):1059-102. Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. *J Toxicol Environ Health*. 1997 Mar;50(4):307-64.
- Wynder EL, Muscat JE. The changing epidemiology of smoking and lung cancer histology. *Environ Health Perspect*. 1995 Nov;103 Suppl 8:143-8. Hoffmann D, Rivenson A, Hecht SS. The biological significance of tobacco-specific N-nitrosamines: smoking and adenocarcinoma of the lung. *Crit Rev Toxicol*. 1996;26(2):199-211.



- Bartsch H, Dally H, Popanda O, Risch A, Schmezer P. Genetic risk profiles for cancer susceptibility and therapy response. *Recent Results Cancer Res.* 2007;174:19-36.
- Clegg LX, Reichman ME, Miller BA, Hankey BF, Singh GK, Lin YD, Goodman MT, Lynch CF, Schwartz SM, Chen VW, Bernstein L, Gomez SL, Graff JJ, Lin CC, Johnson NJ, Edwards BK. *Cancer Causes Control.* 2008 Nov 12.
- Brske-Hohlfeld, I. Environmental and occupational risk factors for lung cancer. *Methods Mol Bio.* 2009;472:3-
- Gonzalez CA. The European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr.* 2006 Feb;9(1A):124-6.
- Benowitz NL, Hall SM, Stewart S, Wilson M, Dempsey D, Jacob P 3rd. Nicotine and carcinogen exposure with smoking of progressively reduced nicotine content cigarette. *Cancer Epidemiology Biomarkers Prev.* 2007 Nov;16(11):2479-85.
- Moolgavkar SH and Luebeck EG. Two event model for carcinogenesis: biological, mathematical, and statistical considerations. *Risk Anal.* 1990;10:323-341.
- Heidenreich WF, Luebeck EG, Moolgavkar SH. Some properties of the hazard function of the two-mutation clonal expansion model. *Risk Anal.* 1997;17:391-399.
- Meza R, Hazelton WD, Colditz GA, Moolgavkar SH. Analysis of lung cancer incidence in the Nurses' Health and the Health Professionals' Follow-Up Studies using a multistage carcinogenesis model. *Cancer Causes Control.* 2008 Apr;19(3):317-28. Epub 2007 Dec 6.
- Hazelton WD, Moolgavkar SH, Curtis SB, Zielinski JM, Ashmore JP, Krewski D. Biologically based analysis of lung cancer incidence in a large Canadian occupational cohort with low-dose ionizing radiation exposure, and comparison with Japanese atomic bomb survivors. *J Toxicol Environ Health A.* 2006 Jun;69(11):1013-
- Hazelton WD, Clements MS, Moolgavkar SH. Multistage carcinogenesis and lung cancer mortality in three cohorts. *Cancer Epidemiol Biomarkers Prev.* 2005 May;14(5):1171-81.
- Curtis SB, Hazelton WD, Luebeck EG, Moolgavkar SH. From mechanisms to risk estimation—bridging the chasm. *Adv Space Res.* 2004;34(6):1404-9.
- Clements MS, Armstrong BK, Moolgavkar SH. Lung cancer rate predictions using generalized additive models. *Biostatistics.* 2005 Oct;6(4):576-89. Epub 2005 Apr 28.
- Knoke JD, Shanks TG, Vaughn JW, Thun MJ, Burns DM. Lung cancer mortality is related to age in addition to duration and intensity of cigarette smoking: an analysis of CPS-I data. *Cancer Epidemiol Biomarkers Prev.* 2004 Jun;13(6):949-57.
- Haylock RG, Muirhead CR. Fitting the two-stage model of carcinogenesis to nested case-control data on the Colorado Plateau uranium miners: dependence on data assumptions. *Radiat Environ Biophys.* 2004 Feb;42(4):257-63. Epub 2003 Nov 15.
- Krewski D, Zielinski JM, Hazelton WD, Garner MJ, Moolgavkar SH. The use of biologically based cancer risk models in radiation epidemiology. *Radiat Prot Dosimetry.* 2003;104(4):367-76. Review.
- Curtis SB, Luebeck EG, Hazelton WD, Moolgavkar SH. A new perspective of carcinogenesis from protracted high-LET radiation arises from the two-stage clonal expansion model. *Adv Space Res.* 2002;30(4):937-44.
- Heidenreich WF, Luebeck EG, Hazelton WD, Paretzke HG, Moolgavkar SH. Multistage models and the incidence of cancer in the cohort of atomic bomb survivors. *Radiat Res.* 2002 Nov;158(5):607-14.



- Curtis SB, Luebeck EG, Hazelton WD, Moolgavkar SH. The role of promotion in carcinogenesis from protracted high-LET exposure. *Phys Med.* 2001;17 Suppl 1:157-60.
- Little MP, Haylock RG, Muirhead CR. Modelling lung tumour risk in radon-exposed uranium miners using generalizations of the two-mutation model of Moolgavkar, Venzon and Knudson. *Int J Radiat Biol.* 2002 Jan;78(1):49-68.
- Moolgavkar SH, Brown RC, Turim J. Biopersistence, fiber length, and cancer risk assessment for inhaled fibers. *Inhal Toxicol.* 2001 Sep;13(9):755-72.
- Hazelton WD, Luebeck EG, Heidenreich WF, Moolgavkar SH. Analysis of a historical cohort of Chinese tin miners with arsenic, radon, cigarette smoke, and pipe smoke exposures using the biologically based two-stage clonal expansion model. *Radiat Res.* 2001 Jul;156(1):78-94.
- Moolgavkar SH, Turim J, Brown RC. The power of the European Union protocol to test for carcinogenicity of inhaled fibers. *Regul Toxicol Pharmacol.* 2001 Jun;33(3):350-5.
- Moolgavkar SH, Turim J, Brown RC, Luebeck EG. Long man-made fibers and lung cancer risk. *Regul Toxicol Pharmacol.* 2001 Apr;33(2):138-46.
- Luebeck EG, Curtis SB, Hazelton WD, Moolgavkar SH. The importance of promotion in lung carcinogenesis for protracted exposures to radon and radon progeny. *Radiat Res.* 2000 Dec;154(6):730;discussion 730-1. No abstract available.
- Moolgavkar SH. Air pollution and daily mortality in three U.S. counties. *Environ Health Perspect.* 2000 Aug;108(8):777-84.
- Moolgavkar SH, Luebeck EG, Turim J, Hanna L. Quantitative assessment of the risk of lung cancer associated with occupational exposure to refractory ceramic fibers. *Risk Anal.* 1999 Aug;19(4):599-611.
- Neus H, Schmann M, Koss G. [Uncertainty analysis of quantitative risk assessment in temporally variable exposures: model observations based on biological and epidemiological risk models] *Schriftenr Ver Wasser Boden Lufthyg.* 1999;103:140-200. German.
- Moolgavkar SH, Luebeck EG, Turim J, Brown RC. Lung cancer risk associated with exposure to man-made fibers. *Drug Chem Toxicol.* 2000 Feb;23(1):223-42.
- Cardis E, Zeise L, Schwarz M, Moolgavkar S. Review of specific examples of QEP. *IARC Sci Publ.* 1999;(131):239-304. Review. No abstract available.
- Luebeck EG, Heidenreich WF, Hazelton WD, Paretzke HG, Moolgavkar SH. Biologically based analysis of the data for the Colorado uranium miners cohort: age, dose and dose-rate effects. *Radiat Res.* 1999 Oct;152(4):339-51.
- Castrn O. Implications of a two-stage clonal expansion model to indoor radon risk assessment. *Health Phys.* 1999 Apr;76(4):393-7.
- Moolgavkar SH, Luebeck EG, Anderson EL. Estimation of unit risk for coke oven emissions. *Risk Anal.* 1998 Dec;18(6):813-25.
- Kai M, Luebeck EG, Moolgavkar SH. Analysis of the incidence of solid cancer among atomic bomb survivors using a two-stage model of carcinogenesis. *Radiat Res.* 1997 Oct;148(4):348-58. Erratum in: *Radiat Res* 1998 Mar;149(3):309.
- Moolgavkar SH, Luebeck EG, Anderson EL. Air pollution and hospital admissions for respiratory causes in Minneapolis-St. Paul and Birmingham. *Epidemiology.* 1997 Jul;8(4):364-70.



- Luebeck EG, Curtis SB, Cross FT, Moolgavkar SH. Two-stage model of radon-induced malignant lung tumors in rats: effects of cell killing. *Radiat Res.* 1996 Feb;145(2):163-73.
- Luebeck EG. <http://cran.r-project.org/web/packages/Bhat/Bhat.pdf> (2009).
- Stayner L, Smith R, Bailer AJ, Luebeck EG, Moolgavkar SH. Modeling epidemiologic studies of occupational cohorts for the quantitative assessment of carcinogenic hazards. *Am J Ind Med.* 1995 Feb;27(2):155-
- Review.
- Moolgavkar SH. Biological models of carcinogenesis and quantitative cancer risk assessment. *Risk Anal.* 1994 Dec;14(6):879-82.
- Moolgavkar SH. Air pollution and mortality. *N Engl J Med.* 1994 Apr 28;330(17):1237-8. No abstract available.
- Moolgavkar SH, Luebeck EG, Krewski D, Zielinski JM. Radon, cigarette smoke, and lung cancer: a reanalysis of the Colorado Plateau uranium miners' data. *Epidemiology.* 1993 May;4(3):204-17.
- Moolgavkar SH. Carcinogenesis models: an overview. *Basic Life Sci.* 1991;58:387-96; discussion 396-9. Review.
- Moolgavkar SH, Luebeck G. Two-event model for carcinogenesis: biological, mathematical, and statistical considerations. *Risk Anal.* 1990 Jun;10(2):323-41. Review.
- Moolgavkar SH, Cross FT, Luebeck G, Dagle GE. A two-mutation model for radon-induced lung tumors in rats. *Radiat Res.* 1990 Jan;121(1):28-37.
- Moolgavkar SH, Dewanji A, Luebeck G. Cigarette smoking and lung cancer: reanalysis of the British doctors' data. *J Natl Cancer Inst.* 1989 Mar 15;81(6):415-20.
- Stevens RG, Moolgavkar SH. A cohort analysis of lung cancer and smoking in British males. *Am J Epidemiol.* 1984 Apr;119(4):624-41.
- Moolgavkar SH. Model for human carcinogenesis: action of environmental agents. *Environ Health Perspect.* 1983 Apr;50:285-91.