**MGH Institute for Technology Assessment**

**Important note:** This document will be updated periodically. The most current version is available at [http://cisnet.cancer.gov/profiles](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.

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.
  ◦ Population Component
  ◦ Incidence Component
  ◦ Natural History Component
  ◦ Screening Component
  ◦ Treatment Component
  ◦ Survival Mortality Component

Output Overview
Definitions 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.

Further Reading
These topics will provide an intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

Advanced Reading
These topics denote more detailed documentation about specific and important aspects of the model structure.
MODEL PURPOSE

SUMMARY
This document provides a brief overview of how the model will be used.

PURPOSE
The Lung Cancer Policy Model (LCPM) was designed to evaluate the effectiveness, costs, and cost-effectiveness of helical computed tomography (CT) screening for lung cancer in the U.S. The LCPM will inform screening decisions prior to completion of ongoing trials, address limitations of published cost-effectiveness analyses of lung cancer screening and offer an opportunity to evaluate both future screening technologies and advances in treatment effectiveness.

The natural history model at the core of the LCPM was designed to reproduce observed lung cancer incidence and survival rates in a specified cohort, in the absence of screening. A screening component was added to permit validation of the model by reproducing past screening trial results.

Unlike the CISNET-funded models (we are an Affiliate group), the LCPM was originally designed as a single-cohort model rather than a dynamic population model. Most of this Profile document refers to the single-cohort model. An expanded dynamic-cohort (population) version of the LCPM has been developed so that we can evaluate population trends in, for example, incidence.
MODEL OVERVIEW

SUMMARY
This document provides an overview of the lung cancer simulation model developed by researchers at the MGH Institute for Technology Assessment (NCI R01 CA97337, through 2007).

PURPOSE
The LCPM described in this Profile was designed to evaluate screening programs in a specified cohort. Unlike the CISNET-funded models (we are an Affiliate group), the original LCPM does not simulate populations. The population version of the LCPM will be described in a future update of the Model Profile.

Designed to evaluate the effectiveness, costs, and cost-effectiveness of helical computed tomography (CT) screening for lung cancer in the U.S., the LCPM will inform screening decisions prior to completion of ongoing trials, address limitations of published cost-effectiveness analyses of lung cancer screening and offer an opportunity to evaluate both future screening technologies and advances in treatment effectiveness.

BACKGROUND
An effective means of reducing mortality from lung cancer, the leading cause of cancer death in the U.S., is urgently needed. Unfortunately, even a sharp reduction in current smoking rates -- an obvious first step -- would not eliminate lung cancer in the near term: a former smoker’s risk for lung cancer remains elevated for decades after smoking cessation. To date, no screening program has been demonstrated effective at reducing lung cancer mortality.

Ongoing trials of helical CT screening will contribute critical information on effectiveness, but debates over past cancer screening trials (e.g.,\(^1\)) should remind us that publication of completed trial results is unlikely to eliminate uncertainty about the effectiveness of lung cancer screening.

Advances in screening technologies, staging examinations, and therapies are being made simultaneously, yet conducting controlled trials on all of these aspects at once is simply not feasible. The comprehensive modeling approach used in the LCPM, however, permits an evaluation of all three inter-related areas. Specifically, modeling can be used to: 1) estimate effects of several combined screening, workup, and treatment strategies; 2) interpret and reconcile the results of different screening trials; 3) evaluate the potential effects of improved adherence to staging and treatment guidelines; and 4) determine the effect that improvements in staging and treatment might have on screening effectiveness. Finally, by including costs as well as effectiveness outcomes, our model will provide information concerning the relative cost-effectiveness of interventions spanning the spectrum from screening to treatment, and thereby provide information which is useful to physicians, policy makers, legislators and the public.
MODEL DESCRIPTION

The natural history model at the core of the LCPM is a state-transition model, analyzed as Monte Carlo to allow for individual heterogeneity in risk factors (e.g., smoking history) and thus event rates. There are 5 states: general population, follow-up, diagnosis & staging, treatment & survival, and dead. Please see the Component Overview and links provided for further details.

The model employs a lifetime time horizon and a cycle length of one month to capture the short survival times of late-stage lung cancers and to allow for a wide variety of event recurrence frequencies. The model was populated with individuals in an age-, race-, gender-, and calendar year-specific cohort representative of the U.S. in terms of smoking history (Population Component).

Inputs include national survey data for assigning smoking histories, type-specific distributions of doubling times for lung cancers (Natural History Component), rates of thoracic imaging exams performed for reasons unrelated to lung cancer, and response rates of treatments.

Outputs include estimation of incident cancers (Incidence Component), stratified by age, type, and stage, as well as mortality by detected stage and treatment (Survival Mortality Component). Calibration to observed incidence and stage-specific survival curves from the NCI SEER tumor registry allowed estimation of parameters governing unobservable events, such as development of the first cancerous cell and of metastasis. Some endpoints from CT screening trials and other literature sources describing clinical experience were used as secondary calibration targets. Validation of the model was performed by reproducing observed results of a past lung cancer screening trial and cohort studies. See Res Calib Valid for a summary of model calibration and validation.

As with any model, simplifying assumptions were made (Assumption Overview). Increasing complexity of the model must be balanced against the number of parameters that can be estimated using available data; calibration and validation can show that model outputs are consistent with observed data, but do not guarantee that the model accurately represents the underlying biology. The model currently omits radon and second-hand smoke exposure, two known risk factors for lung cancer.

CONTRIBUTORS

Pamela M. Mc Mahon, PhD
Chung Yin Kong, PhD
G. Scott Gazelle, MD, MPH, PhD

REFERENCES:

1 Gøtzsche, P. C., Nielsen, M. “Screening for breast cancer with mammography” in Cochrane Database of Systematic Reviews 2006; 4
ASSUMPTION OVERVIEW

SUMMARY
This document discusses key assumptions underlying the Lung Cancer Policy Model (LCPM) and their possible implications. See the linked Component documents for context and further details.

BACKGROUND
The LCPM is a comprehensive model of lung cancer development, detection, treatment, and survival. Unlike a stage-shift model, the LCPM's underlying natural history model does not require estimates of Screening Biases (eg lead-time) as model inputs. To estimate parameters governing unobservable events (e.g., timing of metastasis), we calibrated to multiple endpoints in observed data (see Calibration Details). Using the calibrated model, we can simulate a screening program and generate estimates of the screening biases as model outputs.

Like the Coronary Heart Disease Policy Model developed by Dr. Milton Weinstein (a Co-Investigator on the LCPM) and colleagues\(^1\),\(^2\), the LCPM was designed to be a model with a long lifespan, anticipating updating and refinements. The LCPM does not rely on data from a single trial to inform the parameter estimates (but rather incorporates trial data as they emerge), so can be used to evaluate screening in populations not included in ongoing trials, and can address the ‘moving target’ problem of improved test sensitivity (e.g., CT resolution), as well as other late-breaking topics, such as treatment interventions.

ASSUMPTION LISTING
Population Component Assumptions

- To allow for undetected lung cancers in the cohort, each individual is first ‘regressed’ to age 20 and assumed to be free of lung cancer. Upon entering the general population state, he can develop one or more lung cancers as he ages and acquires his (known\(^3\),\(^4\)) smoking exposure. Two procedures insure that the cohort still reflects the U.S. population upon reaching the cohort age: 1) individuals face no competing risks of death until reaching the cohort age; and 2) any individual who dies of lung cancer prior to reaching the cohort age is re-started at age 20 with the same smoking history. Aggressive cancers that would have been fatal at ages younger than the cohort age are thus appropriately removed.
- After 1990, all current smokers face a 3% annual chance of quitting, based on estimates of 2.5% to 3.2%.\(^5\),\(^6\)
- Smoking histories (see Population Component) did not incorporate the tendency of beginning smokers to gradually increase the number of cigarettes smoked per day. A closer approximation of such smoking behavior would influence the apportioning of lung cancer risk across the population and yield fewer individuals eligible for screening.

Natural History Component Assumptions
• The risk of developing a lung cancer (≤ 3 per person) is modeled using a tolerance model: increasing age, smoking exposure, and genetic susceptibility contribute to risks of developing one of 5 types of lung cancer.

• The growth rate assigned to each cancer is drawn from a distribution specific to the histologic type, is assumed to decrease with increasing size, and can vary by smoking status.

• Disease progression is modeled through monthly probabilities of involvement of lymph nodes and development of distant metastases. Progression risks are functions of characteristics of existing cancers (location, volume, doubling time, and type), nodal status, and random individual variation.

• Incidence of benign lesions varies with age and geographical region but not with smoking history. Few benign lesions exhibit cancer-like growth.

• The proportion of mixed adenocarcinoma/BAC that is pure BAC (bronchioloalveolar carcinoma) was estimated via calibration, not taken from the literature.

**Incidence Component Assumptions**

• Benign lesions and asymptomatic lung cancers can be detected incidentally during a thoracic imaging exam performed for an unrelated reason (non-screening). Risks of incidental imaging are functions of age, gender, and geography. Sensitivity varies with size and location and was estimated during calibration.

• Symptom detection can occur via symptoms from the largest primary cancer, by distant metastases, or both.

• Incidence rates reported in SEER reflect a negligible rate of lung cancer screening in the population.

**Follow Up Component and Workup And Staging Component Assumptions**

• Lesions suspicious for lung cancer (from symptoms or incidental detection) are biopsied if over a minimum diameter or followed with serial high-resolution CT exams (even in the absence of screening).

• Lesions that exhibit no detectable growth after 2 years of follow-up are assumed to be benign and to require no further surveillance. Reflecting clinical practice, a proportion of benign lesions are diagnosed as benign on the basis of a high-resolution CT (a proxy for modeling calcification patterns).

• Biopsy-confirmed malignancies are clinically staged based on guidelines recommended by the National Comprehensive Cancer Network (NCCN).

**Treatment Component Assumptions**

• Treatments are assigned following NCCN guidelines.

• Effectiveness of systemic treatments are based on probabilities of complete or partial response. See below for relationship of treatment effectiveness to survival.

• Effectiveness of resection depends on the existence of undetected second lung cancers and/or occult metastases.

**Survival Mortality Component Assumptions**
• Survival is a function of both underlying disease state and treatment received (which itself depends on the accuracy of staging). Patients with M1 (stage IV or ES) cancers are assigned exponential survival, based on observed median survival rates. (Observed stage-specific survival rates for patients with M0 cancers are used as calibration targets, not inputs.)

• Once a patient is diagnosed as stage IV, survival is as observed in SEER (by age, decade, race, gender, and cell type).

REFERENCES:


PARAMETER OVERVIEW

SUMMARY
This document provides an overview of the major classes of parameters in the LCPM, and provides links to parameter documents.

BACKGROUND
The NCI’s SEER registry was the primary data source for calibration targets. In the absence of screening, the model should accurately predict observed tumor registry (SEER) incidence by year, age, sex, and race. Characteristics of incident cancers predicted by the model should also correspond to observed distributions of cell types, stages, and sizes, and stage-specific survival rates for M0 cancers should be accurately predicted. Additional calibration targets were defined from the literature (see Calibration Details and Output Overview).

PARAMETER LISTING OVERVIEW

1. Unobservable parameters define unobservable events: development of lung cancer, disease progression, and symptom detection. See Parameters Natural History and Natural History Component for details.

Note that the actual values of many of these parameters are not meaningful outside the context of the LCPM, but their relative magnitudes may reveal insights into biology.

2. Uncertain parameters were those for which literature estimates provided ranges of values. Categories of uncertain parameters included test characteristics, operative mortality rates for interventions, response rates for systemic therapies, and probabilities of clinical events such as wedge biopsy of a growing pulmonary lesion.

See Parameters Test Performance and Parameters Treatment for details.

3. Structural parameters were fixed during calibration, but included for future analyses. These included a parameter to allow simulation of African-American cohorts. Additional structural parameters are described in their relevant model components.

The Assumption Overview describes major assumptions underlying the LCPM.
COMPONENT OVERVIEW

SUMMARY
This document describes typical sequences of component processes for a hypothetical individual simulated by the LCPM.

OVERVIEW
Persons start the model in the general population state. See schematic.

Each month, persons in the general population state face competing risks of death from causes other than lung cancer. While in the general population state, benign pulmonary nodules and lung cancers can develop, and lung cancers can grow, progress to nodal involvement and/or distant metastases, or cause symptoms. Additionally, persons face risks of undergoing non-screening thoracic imaging exams for reasons unrelated to lung cancer (e.g., for trauma). In a screening scenario, persons can undergo screening if they are eligible for the specific program and adherent to the screening protocol.

Persons with small incidentally detected lesions undergo sequential imaging exams in the follow up state.

Larger incidentally detected lesions, lesions exhibiting growth on serial imaging exams, and symptomatic cancers are sent for work up and staging.

Once the diagnosis of lung cancer is made, the cancer is staged, and the person moves to the treatment and survival state.
In the next section, we provide available links to component processes for each of the states indicated above.

**COMPONENT LISTING**

**General population**
The Natural History Component is included, as well as the Screening Component and the Incidental Imaging Component.

**Follow-Up**
In the Follow Up Component, incidentally-detected nodules smaller than the cutoff threshold are managed expectantly with periodic high-resolution CT exams. While a patient is being followed up, he also cycles through the Natural History Component.

**Work-up and Staging**
In a single cycle (one month), workup and staging tests are used to establish both the presence of lung cancer as well as the extent of disease progression. See the Workup And Staging Component. Patients also cycle through the Natural History Component.

**Treatment and Survival**
In addition to the Treatment Component and the Survival Mortality Component, the Natural History Component is also included here. This allows for development of second lung cancers as well as disease progression of existing primary cancers or occult metastases.

See the Assumption Overview for key assumptions and links to parameter documents from the corresponding component documents.

**REFERENCES:**

POPULATION COMPONENT

SUMMARY
This document describes how the simulated population is modeled.

OVERVIEW
The population component defines the initial characteristics of the population entering the LCPM for a given simulation. This is the first component in the modeling process. Once each individual in the population is initiated, the individual moves to the general population state (see Component Overview).

Initial characteristics assigned to each hypothetical person include gender, race, ethnicity, age, and smoking history. Smoking history includes current status, age at smoking initiation (if applicable), age at smoking cessation (if applicable), and cigarettes per day. In the single-cohort LCPM, the cigarettes smoked per day is assumed constant for cycles in which the individual is a current smoker.

Additional characteristics include an indicator for genetic susceptibility to lung cancer (see Natural History Component).

Simulations begin in a specified calendar year, so that:
1) the proportions of ethnicities and the prevalence of smoking is representative of the cohort being simulated, and
2) the model-predicted incidence rates can be compared to the corresponding SEER data (see Incidence Component).

QUANTITATIVE DESCRIPTION
The LCPM is a state transition (Markov) model, analyzed as Monte Carlo (i.e. it simulates life histories of individuals). The possible states are described in the Component Overview. The model uses a lifetime time horizon and a cycle length of one month to capture the short survival times of late-stage lung cancers and to allow for a wide variety of event recurrence frequencies.

POPULATION DYNAMICS
The version of the LCPM described in the Model Profiler is a single-cohort model. Individuals enter the model in specified calendar years, however, and carry appropriate smoking histories (informed by national survey data).

A multi-cohort population LCPM (to be described in future Model Profile documents) uses the Smoking History Generator.

RECURRENCE
The LCPM does model recurrence, via either clinical detection of (previously) occult metastases or by development of a new primary lung cancer. Each individual in the LCPM can develop up to 3 lung cancers (of any of 5 cell types), and up to 3 benign lesions. See the Natural History Component and the Survival Mortality Component.
**DISEASE DISTRIBUTION**

We model the risks of developing each of the 5 cell types of lung cancer (adenocarcinoma/BAC, large cell, small cell, and squamous, as well as other) as independent, conditional on risk factors. (In other words, we do not assign a distribution of lung cancer histologies to the population.)

**DETAIL**

The single-cohort LCPM simulates cohorts of white males and females aged 50, 60 or 70 in 1990.

Approximately 10% of lung cancers occur in life-long non-smokers and SEER data are not stratified by smoking history. Therefore, the LCPM is populated with entire age, race and gender cohorts, representative of the U.S. in terms of smoking history. Cohorts entered the model in calendar year 1990 for calibration to SEER data from 1990 to 2000. Ethnicity (Hispanic/non), region of the country, and smoking history are assigned to each hypothetical individual.

Smoking history includes current status, age at smoking initiation (if applicable), age at smoking cessation (if applicable), and cigarettes per day (currently this is assumed constant for cycles in which the individual is a current smoker). Using the 1990 National Health Interview Survey we fit a multinomial logistic model to estimate the proportion of each smoking status using the predictors age, sex, race, ethnicity, and region. Data from the third National Health and Nutrition Examination Survey fielded in 1988-1994, were used to estimate normal distributions of ages of starting and stopping smoking and the average number of cigarettes smoked per day, conditional on smoking status, age group, and ethnicity.

An indicator for genetic susceptibility to lung cancer (see Natural History Component) is assigned randomly.

(Trial populations can also be simulated - see Screening Component.)

**RELEVANT ASSUMPTIONS**

See the Assumption Overview.

**RELEVANT PARAMETERS**

For the single-cohort LCPM, input parameters control which cohort is simulated (age in 1990, gender). Joint distributions of ethnicity and geographic region of the U.S. were derived from the 1990 Census.

Input distributions of smoking characteristics were derived from the NHIS (multinomial model for proportions of current, former, and never smokers, by age, gender, race, ethnicity, and region of the U.S.) and NHANES (ages of starting and stopping smoking, and cigarettes per day while smoking regularly).
RELEVANT COMPONENTS
The Population Component is necessary to specify the characteristics of the cohort entering the LCPM. Different cohorts will have different lung cancer risks and therefore outcomes.

DEPENDENT OUTPUTS
All outputs will be affected by the characteristics of the population being simulated. Heavier smokers, for instance, will have higher rates of lung cancer death and possibly poorer outcomes from treatment.

RELEVANT RESULTS
The base case outputs of the single-cohort LCPM are described in the Output Overview document.
INCIDENCE COMPONENT

SUMMARY
To be counted as an incident case, a lung cancer must first develop, then be detected (by any of several possible modalities) and finally be diagnosed.

This document describes how the LCPM counts incident lung cancers and provides links to descriptions of components that involve development of cancer, detection, and diagnosis.

OVERVIEW
In the LCPM, we model the development of lung cancers, followed by tumor growth and metastasis (see Natural History Component). An individual with undetected lung cancer remains in the general population state (see Component Overview).

After the last individual in a cohort is simulated, we essentially count up the numbers of cancers in various categories. We count as incident cancers only cancers that were diagnosed during the patient's lifetime. Incident cancers are further categorized by stage, size, type, etc.

Non-screening scenarios:
Lung cancer can be diagnosed symptomatically (either the primary cancer obstructing an airway or from distant metastases) or asymptotically (found incidentally during a thoracic imaging exam performed for unrelated causes - see Relevant Components, below).

Age-specific incidence rates are calculated and then compared to observed data.

Screening scenarios:
Cancers may also be detected by screening (see Screening Component). The model tracks the mode of detection of each cancer.

DISEASE RISK
For each of the 5 lung cancer cell types, we estimate a logistic function to predict monthly risks of developing a cancer. For each cancer type, we estimated independent coefficients for age, age squared, cigarettes per day, years of smoking, and years since quitting. There is also a randomly-assigned indicator for increased genetic risk (equivalent to HR=2). See Natural History Component.

We have not yet modeled races other than white. To account for observed birth cohort trends in lung cancer risks and allow for differences in baseline risk by gender, we added a term that modifies the monthly risk of lung cancer development (all cell types), stratified by gender. See Calibration Details.

IMPACT OF SCREENING
As described above, we do distinguish between incidence in the absence of screening vs. the presence.
We calibrate to SEER (no screening) and validate with screening trial data (with screening).

**DETAIL**

**Development of cancer**
Further details on the way the LCPM simulates development of cancer is provided in the Natural History Component.

**Detection of cancer**
Three modes of detection are possible in the LCPM:

1. Symptoms of previously undiagnosed lung cancers (either the primary cancer or distant metastases) can prompt detection. See the Symptom Detection Component.

2. During each cycle spent in the general population, persons may undergo imaging exams (thoracic CT, or CXR) performed for reasons unrelated to screening for lung cancer. See the Incidental Imaging Component.

3. Screen detection can occur in eligible individuals, in scenarios which include screening. See Screening Component.

**Diagnosis of cancer**
In the LCPM, a diagnosis of lung cancer is required before a person transitions into the Treatment Component. Diagnosis is operationalized by a biopsy that returns a specific diagnosis of lung cancer. Biopsies and staging both occur in the one-month Workup And Staging Component.

**RELEVANT ASSUMPTIONS**
For the single-cohort LCPM, we used national survey data (NHIS, NHANES) to assign smoking histories to the individuals in the cohort (see Population Component), and calibrate to SEER data for incidence.

- If the SEER registries are not representative of the US, calibration to SEER data may yield biased parameter estimates.
- The smoking histories used for the single-cohort LCPM do not reflect the tendency of individuals to increase their smoking intake over time (i.e., light smokers become heavy smokers), which overestimates the pack-years accrued. This could have resulted in biased estimates of the cumulative dose-response relationship between smoking and lung cancer risk.

Also see the Assumption Overview.

**RELEVANT PARAMETERS**
This component relies most directly on the natural history parameters (see Natural History Component and the Symptom Detection Component).
However, parameters in other components can influence the incidence rates, such as patterns of imaging examinations (and their test characteristics) in the general population (see the Incidental Imaging Component) and whether screening is occurring (see the Screening Component).

RELEVANT COMPONENTS
The incidence component operates after the last individual in a simulated cohort 'dies.' It does not contain any other components, per se, but merely functions as a bookkeeping component.

Components that influence the predicted incidence rates include the Natural History Component, as well as the Follow-up, Workup And Staging Component and the Incidental Imaging Component.

DEPENDENT OUTPUTS
Most outputs of interest will depend on the Incidence component, including incidence rates and therefore mortality rates.

RELEVANT RESULTS
See Res Calib Valid for a description of outputs from the LCPM after calibration and validation and links to specific outputs.
NATURAL HISTORY COMPONENT

SUMMARY
This document describes various aspects of the model processes responsible for generating the natural history of lung cancer. Benign pulmonary nodules are described in the Benign Component.

OVERVIEW
The natural history component occurs in every cycle of the model, so that new lung cancers may develop (and existing lung cancers grow and progress) throughout life.

The Population Component initiates the population entering the LCPM and therefore precedes the natural history component. The natural history component has sub-components for lung cancer development, disease (tumor) growth, disease progression, and symptom detection. All of the sub-components are described below and/or in linked documents.

Approximately 6% of patients with lung cancer develop more than one primary tumor, and only half of synchronous multiple primaries are the same type. Therefore, we model up to three cancers per person, of any of the 4 main types of lung cancer, plus a 5th type to represent Carcinoma, Not Otherwise Specified (ICD-O-2 code 80103). We modeled pure bronchioloalveolar carcinoma (BAC) as a subset of adenocarcinoma +/- some BAC, reflecting their differences yet typically mixed histology and misclassification.

DISEASE STAGES
A ‘true’ disease stage is assigned based on the individual’s simulated disease characteristics (tumor size, nodal involvement, distant spread). This true stage is updated every cycle. See Details, below. An observed disease stage is also assigned, based on the individual’s ‘true’ disease characteristics and the results from any diagnostic or staging tests performed. Observed and true stages may not match if a cancer is undiagnosed or mis-staged by a test result.

DISEASE GROWTH
We assume continuous Gompertz tumor growth, assigning a growth parameter for each new cancer that is drawn from distributions specific for the 5 cell types of lung cancer. We also include a term to allow cancers in smokers to exhibit accelerated growth. See Details, below and Parameters Natural History.

STAGE TRANSITION TRENDS
No temporal trends are imposed on stage transitions.

DISEASE EVOLUTION
One birth cohort parameter is changed over calendar time:

To account for observed birth cohort trends in lung cancer risks and allow for differences in baseline risk by gender, we added a term that modifies the monthly risk.
of lung cancer development (all cell types), stratified by decade of birth and gender. See Calibration Details.

Remaining natural history parameters are not changed over (calendar) time. (Smoking histories do change over time, however, so will influence lung cancer trends.)

**REGRESSION**

The model assumes an irreversible (in the absence of resection) progression of lung cancer disease stages. The speed of progression varies greatly, however, so that some cancers would never be detected during life in the absence of screening. The growth of BACs was truncated at a maximum diameter of 1 cm (detectable by X-ray).

**DETAIL**

**Lung Cancer Development**

The LCPM employs a simple ‘tolerance’ model of cancer development (so-called because cancer may only develop after an individual’s tolerance to risk factors has been exceeded).

The monthly probability of developing the first malignant cell of cancer type \( k = 1-5 \) is a logistic function with a type-specific intercept and type-specific coefficients for age, age\(^2\), years of cigarette exposure (smoke-years, SY), average number of cigarettes smoked per day (cigarettes per day, CPD), and the years since quitting (YSQ) smoking, if applicable. We also allow for random individual variation (highrisk, a proxy for genetic susceptibility), constant for all 5 types.

A logistic model produced nearly as high an \( R^2 \) goodness-of-fit statistic as a two-stage model\(^7\) (\( R^2 \) of 0.61 and 0.67, respectively) in a comparison of 5 models for lung cancer’s dose-response to tobacco,\(^8\) and studies of case control data showed good fit using a logistic function to predict lung cancer (all types combined).\(^9\) The MVK 2-stage model\(^7\) models each initiated cell as growing instantaneously into a malignant tumor after a fixed period of time,\(^10\) an assumption that precludes size-dependent sensitivity of imaging exams.

**Lung Cancer Characteristics and Growth**

Indicators are assigned to each new cancer for cell type, size (initial diameter of 0.01 mm), lobe in the lung, and central or peripheral location (varied by type).

In each cycle, the diameter and volume of existing cancers (and any growing benign lesions) are incremented according to a Gompertz function for tumor growth. Consistent with biological mechanisms of tumor growth (e.g., angiogenesis and necrosis of the tumor core), tumor doubling times decrease as the volume asymptotes to its maximum possible.

Mean doubling times for large, small, and squamous cell cancers were estimated from the literature (see Table Growth Parameters) and used to derive distributions of growth rate parameters. Distributions of growth rate parameters for adenocarcinoma/BAC and “other” cell types were estimated via calibration.

A modification term (estimated during calibration) allows slower growth rates in non-
Lung Cancer Progression

Disease progression of an existing lung cancer can occur via nodal involvement and distant metastasis. Risks of disease progression depend on characteristics of any cancers present, and random individual variation that allows for more or less aggressive cancers, given a cancer’s size and growth rate.

For each individual, 8 threshold volumes are drawn randomly from distributions for each nodal stage (N1, N2, N3) and for distant spread (M1), stratified by cell type (NSCLC/SCLC). Threshold volumes are adjusted to allow variation by growth rate. In each cycle, development of metastases and involvement of lymph nodes (N1, 2, 3) occurs if and only if the current volume of the largest cancer is greater than the corresponding adjusted threshold volume.

Symptom Detection

Each month, individuals with distant metastases and/or a primary lung cancer may develop symptoms that result in lung cancer detection. A person with symptom-detected cancer begins the following cycle in the Workup And Staging Component. See Symptom Detection Component for details.

Relevant Assumptions

Lung Cancer Development

The probabilities of developing each cancer type are assumed independent, conditional on the covariates (see Relevant Parameters, below). Each month >age 20, only one cancer can develop. Because the monthly probabilities are on the order of 10E-7, bias resulting from development of more than one cancer type is negligible.

Lung Cancer Growth

After Spratt, a maximum possible tumor size of 277 mm is assumed (this is consistent with the largest reported size of 201-300mm diameter in the SEER*Stat database for 60-64 year old white males, 1990-1994). As a simplification, we assume equal growth in all directions (i.e., spherical), allowing only one diameter to be tracked. The growth of BACs was truncated at a maximum diameter of 1cm.

Lung Cancer Progression

By definition, BACs do not progress. Because AJCC stage T3 cancers (i.e., cancers with extension into adjacent organs) represent only about 5% of NSCLC, we modeled tumor stage as T1 (≤3cm) or T2+ (>3cm). Involvement of lymph nodes (stages N0, N1, N2, and N3) dictates treatment options, so nodal status is modeled explicitly (but not specific nodes within each stage). Once distant spread (M1) has occurred, survival is poor, so explicit modeling of types of metastases was assumed to be unnecessary. As hypothesized for breast cancer, growth rate is related to the probability of metastasis. To reflect observed variations in propensity to metastasize for each histological type, adenocarcinomas are often more indolent, while small cell lung cancers develop metastases earlier. We assume that lymph nodes typically (but not always) become involved before distant spread occurs.

Lung Cancer Symptom Detection

We assume that peripheral cancers must be at least 10mm in diameter to cause
symptoms. Central cancers have a smaller minimum diameter, because they are more likely to obstruct airways. We assume that metastases from SCLC cause symptoms faster than metastases from NSCLC. Benign nodules and lymph node involvement do not cause symptoms that result in lung cancer detection.

**RELEVANT PARAMETERS**
The parameters in the Natural History Component are informed by calibration (see Calibration Details and Parameters Natural History).

**RELEVANT COMPONENTS**
The Natural History Component occurs in every cycle, so can be thought of as a sub-component of the major states in the LCPM (see Component Overview for schematic).

Sub-components in the Natural History Component are described above or in linked documents:
lung cancer development, lung cancer growth, lung cancer progression, Symptom Detection Component, and Benign Component.

**DEPENDENT OUTPUTS**
The natural history component primarily determines the lung cancer incidence rate, as well as the type and stage distributions of incident cancers. The natural history component also primarily determines the survival rates of incident cancers, in conjunction with the Treatment Component.

The particular staging, work-up, and follow-up algorithms used in a scenario will also influence the stage distribution and rate of incident cancers, as well as the stage-specific survival rates (see the Workup And Staging Component). And the rates of thoracic imaging exams performed for reasons unrelated to screening (Incidental Imaging Component) will also influence incidence rates, although to a lesser extent than the natural history components.

**RELEVANT RESULTS**
See the Results Overview for a description of the outputs from the base case LCPM.

**REFERENCES:**


SCREENING COMPONENT

SUMMARY
This document describes the processes in the model responsible for screen-detection of asymptomatic lung cancers.

OVERVIEW
Note that during model calibration (to SEER registry data), the screening component was turned off.

To define a screening program, we can specify eligibility in terms of age and pack-year histories, as well as screen frequencies and probabilities of adherence to recommended screenings.

To reproduce results from a particular screening study (e.g., for validation), the model is populated with simulated trial participants and the screening component is turned on.

An individual with a positive screening exam proceeds in the next cycle to either the Follow Up Component or the Workup And Staging Component. The particular scenario being modeled determines which of these components a person will transition into.

DISEASE DETECTION MECHANISM
Lung cancer detection can occur in one of 4 ways: 1) by symptom detection of distant metastases, as a function of the time since metastases developed, varied by NSC/SCLC; 2) symptom detection of the primary cancer as a function of size and location; 3) by incidental detection of an asymptomatic lung cancer on a chest imaging exam performed for unrelated reasons; and 4) by a screening exam, limited to imaging exams at the moment, in a scenario with screening operating.

For any imaging exam, the probability of detection of asymptomatic cancers is a function of size, location (peripheral/central) and test characteristics. The test characteristics do not change over time, although it would be easy enough to incorporate such a trend.

SCREENING DISSEMINATION
For screening exams, individuals are screened if they are both 1) eligible for the screening program being modeled and 2) compliant, which is assigned randomly according to the population-wide probability of compliance.

The probability of incidental (non-screening) imaging exams is a function of age, region of the U.S., and race. See Incidental Imaging Component. Temporal trends in these background rates have not yet been incorporated.
TYPE / DETECTION INTERACTION

The probability of detection on an imaging exam is a function of nodule size (and
dependently growth rate, indirectly) and location: central lesions are less likely to be
detected. Both growth rate and the proportion central vs. peripheral vary by cell type.
See the Natural History Component.

STAGE / DETECTION INTERACTION

As stated above, the probability of detection of a pulmonary lesion on an imaging
exam is a function of nodule size and location (central/peripheral) and the test
characteristics of the imaging exam. Nodal involvement and distant metastases are not
detected on a screening imaging exam so do not influence screen detection (but do of
course influence symptom-detection - see the Natural History Component).

LENGTH BIAS

Slower-growing lesions persist in the asymptomatic state and are therefore more often
'available' to be screen-detected, on average, than faster-growing lesions. The
probability of detection on an imaging exam is a function of lesion size (and therefore
growth rate, indirectly). The growth rate varies by cell type. See the Natural History
Component.

On average, lung cancers detected on annual screening exams would be expected to
have longer doubling times (i.e., slower growth rates) than interval-detected lung
cancers. Note that a small, slow-growing lung cancer may be referred for follow-up
serial CT exams; if no growth is detectable over a two-year period, the cancer would be
incorrectly diagnosed as benign.

See Screening Biases for background information on lead-time, length-time and
overdiagnosis biases.

DETAIL

Parallel random number generation allows simulation of the same individuals in
screening vs. non-screening scenarios. This allows us to compare the outcomes of
individuals in the two scenarios, as well as the mean life expectancy across a cohort, for
a better understanding of the range of individual outcomes attributable to screening.

When simulating a specific screening study, individual-level data from the study (if
available) is used to populate the LCPM with a cohort similar to the study participants.
See Protected Health Information.

RELEVANT ASSUMPTIONS

We assume that nodal involvement and distant metastases are not detected on a
screening imaging exam. See also the Assumption Overview.

RELEVANT PARAMETERS

To enable simulation of screening, one parameter is set in the input file (intervention =
1 for screening, vs. 0 for no screening), and additional parameters define eligibility
(based on pack-years of smoking exposure and age), compliance rates, and screening
frequency (modality, frequency, maximum number of screens, and follow-up
algorithm).

Sensitivity and specificity of the screening exam also affect the efficacy of the screening program. See Parameters Test Performance.

Indicators record screen results and cancers detected.

RELEVANT COMPONENTS
Under a screening scenario, individuals in the general population state are screened if they are 1) eligible, and 2) adherent. Persons in the Follow Up Component, the Workup And Staging Component, and the Treatment Component are not screened.

Screening will increase the rate of detection of lung cancer in a population and therefore impact the Incidence Component. Similarly, by detecting a lung cancer earlier, screening can alter the treatment a patient would receive in the Treatment Component.

DEPENDENT OUTPUTS
To date, we have used outputs from simulations of two single-arm screening studies to calibrate certain endpoints, to validate the LCPM, and to predict outcomes from hypothetical control arms.

The Mayo Clinic and LSS studies are discussed separately below.

RELEVANT RESULTS
The Mayo Clinic conducted a single-arm study of helical CT screening for lung cancer in current and former smokers. Using data provided by the Mayo Clinic (see Protected Health Information), we replicated the trial population by bootstrapping demographics and smoking histories from individual records.

One endpoint (baseline prevalence) was used to calibrate the proportion of adenocarcinoma that was BAC (see Natural History Component). Remaining endpoints were reserved for use as validation endpoints.

Results from an analysis of the Mayo CT study are forthcoming in Radiology.

See the validation section of Res Calib Valid for a link to a description of validation of the LCPM using the LSS study endpoints.
TREATMENT COMPONENT

SUMMARY
This document describes how treatment after diagnosis is modeled.

OVERVIEW
To enter the treatment component, patients must have been diagnosed with lung cancer in the Workup And Staging Component. Treatment is modeled as occurring in the month(s) after reaching the 'treatment and survival' state. Patients remain in this state until death (from any cause). See also the Survival Mortality Component.

Treatment consists of either removal of the primary lung cancer (i.e., resection) or systematic therapy. Tumors which respond to systemic therapy are reduced in size (diameter), following conventional guidelines for solid tumors.¹

TREATMENT DISSEMINATION
Treatment is assigned based on the diagnosed stage and type (NSCLC/SCLC). We currently assume all patients receive care according to consensus guidelines (e.g., National Comprehensive Cancer Network, NCCN).

We are in the process of adding a 'usual care' option that more closely approximates observed practice patterns. The usual care option will allow us to explicitly vary treatments with calendar year, which will be particularly important in the Population LCPM.

TREATMENT EFFICACY
Treatment effectiveness is incorporated as follows: a person with no occult metastases whose primary cancer is resected is assigned competing risks consistent with a person of the same smoking history – not stage I survival from SEER. On the other hand, if occult metastases are present in a person who undergoes resection for an apparent stage I cancer, the metastases continue to develop as before. (The presence of undetected micro metastases is likely the cause of the poor observed survival after "curative" resection in many patients.)

If a second, undetected primary tumor remains (in a non-resected lobe), metastasis can occur. Note that removal or sampling nodes at resection can result in re-assigning stage at diagnosis, but provides no survival benefit.

For systemic therapies, we use probabilities of partial and complete responses as published in the literature (see Parameters Treatment). A response results in a reduction of the size(s) of existing lung cancer(s), and thereby may delay disease progression.

We do model adverse effects of screening and treatment. Operative mortality can occur during resection, mediastinoscopy, or VATS, in diseased or non-diseased persons (iatrogenic deaths are tracked). We have not yet incorporated complications (e.g., pneumothorax) or quality of life.
**DETAIL**

The default Treatment Component simulates all patients as receiving guideline care.

Only individuals assigned the status of operative candidate were eligible for surgical resection, regardless of stage. To account for patients who were not operative candidates, we estimated proportions of SEER cases who were NSCLC stage I and II and either refused surgery or had contraindications.

As observed in clinical practice, a small proportion of operative candidates with stage LS (limited stage) SCLC underwent resection, with the remaining patients receiving chemoradiation.

Operative candidates with NSCLC stages I through IIIa were assigned resection, with the remainder and all stage IIIb cases receiving chemoradiation.

Stages IV (NSCLC) and ES (SCLC) were assigned chemotherapy.

See Parameters Treatment.

**RELEVANT ASSUMPTIONS**

Death from lung cancer is unlikely to occur without detection of metastases (due to symptoms or otherwise), so we assume that once metastases are detected (by symptoms or any modality), survival is as observed in SEER for stage IV-detected cancer. We estimated cause-specific (net) survival for cases diagnosed as stage IV in the years 1990 – 2000, stratified by 10-year age group, race, and gender. Net survival (i.e., in absence of other causes) was used because persons face competing risks elsewhere in the model.  

Median survival of stage IV lung cancer is uniformly less than one year, so survival for M1 (stages IV or ES) is modeled as exponential.

See also the Assumption Overview.

**RELEVANT PARAMETERS**

Parameter values that define treatment efficacy are probabilities of complete or partial response, using the definition of complete as no visible cancer at 4 week follow-up and partial as ≥30% decrease in diameter. Probabilities of complete and partial responses vary by type, with estimates taken from the literature. A cancer that partially responds to therapy is decreased in diameter by 30%, and a cancer that completely responds to therapy is reduced to 1.5mm diameter, or below the 2mm detection threshold assumed for helical CT.

Based on the new diameter, an adjusted ‘time since cancer developed’ is calculated, retaining the original growth parameter, alpha. The new ‘time since cancer developed’ is used to increment growth in all future cycles.

To account for observed differences in growth rates of cancers pre- and post-therapy, we include a parameter that allows faster-growing cancers (cancers with a parameters
over a specified cutoff) to be more or less likely to respond to therapy. These parameters were initially set to values that conferred no effect (probabilities of response vary by cell type and treatment) but were varied during calibration.

See Parameters Treatment.

RELEVANT COMPONENTS
Treatment assignment occurs as the final step in the Workup And Staging Component (i.e., after the stage at diagnosis is assigned). The treatment itself is the first step in the Treatment and Survival state.

The sensitivity and specificity of the staging algorithm influences the treatment assignment. The specific treatment assigned and the treatment’s effectiveness both influence the survival rate.

DEPENDENT OUTPUTS
Survival by stage is dependent on the treatment assigned and the treatment’s effectiveness. For example, assigning systemic therapy instead of resection to a patient with resectable disease will result in a shorter survival time.

Further, the sensitivity and specificity of the staging algorithm (Parameters Test Performance) influence the treatment assignment. For example, if the staging tests performed on a hypothetical patient miss the involvement in a contralateral node, the individual will be understaged and receive an inappropriate treatment.

Mortality rates are calculated as a secondary output, based on age at death among lung cancer patients.

Incidence rates, on the other hand, depend on neither the treatment assigned nor the treatment effectiveness.

RELEVANT RESULTS
See Output Overview and Results Overview.

REFERENCES:
SURVIVAL MORTALITY COMPONENT

SUMMARY
This document describes how lung cancer survival is modeled and describes estimation of both lung cancer-specific and other-cause mortality.

OVERVIEW
Once a hypothetical patient is diagnosed with lung cancer, he or she moves into the 'treatment and survival state', and remains there until death. The Treatment Component precedes the Survival Mortality Component.

The patient can die from any cause while in this state. Relative risks of each cause of death are a function of underlying disease state, treatment received, any surveillance performed, age, gender, race, and smoking history.

SURVIVAL ESTIMATION COVARIATES
Lung cancer-specific survival for patients diagnosed at stages I-III (i.e., M0) is based on the true, underlying disease stage and the treatment assigned. Stage-specific survival for patients with M0 cancers is a calibration target for the LCPM, not an input. See Calibration Details and below.

Once a person is diagnosed (by symptoms or any modality) as stage IV (i.e., M1), lung cancer specific survival is assumed exponential. Cell-type (N/SCLC) specific median survival by age, gender, and race group was estimated from SEER for appropriate calendar years as inputs.

SURVIVAL AFTER CLINICAL DETECTION
Same as above.

SCREEN DETECTION BENEFIT
Screening may detect cancers that have not yet metastasized.

MORTALITY REDUCTION
Resection of an early-stage cancer is curative if 1) no occult metastases remain and 2) no additional lung cancers arise. The person is subject to competing mortality risks appropriate to his/her age and smoking history (see below).

Resection of an early-stage cancer in a patient with occult metastases does not confer a survival benefit.

Mortality rates are calculated after the last individual in a simulation run 'dies', and are merely counts of lung cancer deaths by age at death, divided by the population at risk. Mortality rates are therefore completely dependent on the incidence and survival rates, and are merely an additional way to present the same information.
OTHER CAUSE MORTALITY

The default single-cohort LCPM assigns other-cause mortality risks using results from an independently-conducted analysis. The Population LCPM will be described in future Model Profiler documents and will use the Smoking History Generator developed by members of CISNET.

We developed a Bayesian evidence synthesis model to estimate cause-specific mortality rates stratified by age, sex, race, and smoking status. We combined three data sources: 1) individual survey data on smoking status, demographics, and date and cause of death; 2) population data on numbers of deaths by cause; and 3) cohort study estimates of smoking-related mortality risks, correcting for known inconsistencies between two of the data sources. Risks of heart disease and other causes (i.e., non-lung cancer, non-heart disease) are used as inputs for the LCPM.

LEADTIME

Stage-specific survival is not an input, but rather a calibration target (see Calibration Details and Output Overview).

Estimates of lead, length, and overdiagnosis biases (see Screening Biases) are outputs of the model, obtained by simulating the same cohort with and without screening.

DETAIL

If a patient undergoes curative resection of an early stage lung cancer and harbors no occult metastases, the patient is assigned competing mortality risks appropriate for his or her age, gender, race, and smoking status.

If a patient undergoes resection of an early stage lung cancer but does harbor occult metastases, the metastases continue to develop as before, and can cause symptoms. Once symptomatic, the person is assigned a stage IV survival as above.

Patients who undergo systemic therapies may respond to the therapy, which results in a reduction in the size of the primary lung cancer(s), and therefore a reduction in the monthly probability of disease progression (see Natural History Component).

RELEVANT ASSUMPTIONS

See the Assumption Overview, the Treatment Component and the Natural History Component.

RELEVANT PARAMETERS

See Parameters Treatment for values of parameters that govern effectiveness of treatment and influence survival rates.

RELEVANT COMPONENTS

See the Treatment Component and the Natural History Component.
DEPENDENT OUTPUTS
Stage-specific survival is dependent on this component (Out Calib3). Incidence rates, however, are not dependent on this component.

RELEVANT RESULTS
See Results Overview and Output Overview for more information on calibration and validation outputs.

REFERENCES:
OUTPUT OVERVIEW

SUMMARY
This document describes the types of outputs generated by the LCPM.

To date, outputs are from the original single-cohort version LCPM. Future model profile documents will describe the Population LCPM, which simulates multiple cohorts over calendar time.

OVERVIEW
See Results Overview for a summary of how the various outputs are used for calibration/validation and analyses.

Some general categories of outputs include:

incidence rates;

characteristics of incident cancers;

survival and mortality rates;

screening test results;

estimation of screening biases.

OUTPUT LISTING
Within each general category, the outputs include:

incidence rates;

- Age-specific incidence rates, by gender, race, and calendar year (Out Calib1)
- Age-adjusted incidence rates are outputs from the Population LCPM, described in future model profile documents.

characteristics of incident cancers;

- Size, type and stage distributions of incident cancers (Out Calib2)

survival and mortality rates;

- Survival curves by type (SCLC vs. NSCLC) and stage at diagnosis (Out Calib3)
- Comparison to published cohort studies (Out Valid1)

screening test results;

- Reproduction of observed endpoints in the LSS screening trial (Out Valid2)
- Screening trial endpoints:
  - stage shift
  - number of surgeries (appropriate and inappropriate)
  - number of invasive work-up procedures (appropriate and inappropriate)

estimation of screening biases
By simulating both screened and unscreened scenarios, the model estimates lead-time, length-time, and overdiagnosis (see Screening Biases).
RESULTS OVERVIEW

SUMMARY
This document will discuss results from the LCPM.

OVERVIEW
We see simulation as a tool for evidence synthesis that will allow us to pose a wide range of interesting questions regarding cancer control interventions in a wide variety of populations.

The LCPM has integrated data from national sources such as SEER with individual-level data from screening trials. The LCPM generates results consistent with multiple sources of data simultaneously (i.e., we do not vary natural history parameters to fit each new data source). This is conceptually similar to the ‘borrowing strength’ idea from Bayesian statistics.

Simulation has advantages over meta-analyses as an approach to evidence synthesis. Classical meta-analyses in medicine typically combine published endpoints (hazard or risk ratios) from studies with different designs or discard outliers. Unfortunately, discarded outliers may be from rare studies in a particular subpopulation. More sophisticated statistical approaches fit models with random effects and/or random intercepts, so that trials with different designs can have varying but related effects. However, the LCPM allows us to extract the maximum information from valuable trials by explicitly incorporating trial-level study designs and follow-up protocols.

RESULTS LIST

Calibration and Validation
See Res Calib Valid for results of calibration and validation. A manuscript on Calibration Methods Research is under review (June 2008).

Screening Evaluations

See also Out Valid2 for results from simulating the CT-screened arm of the LSS study.

An evaluation of the U.S. adoption of helical CT screening is in press in Cancer (July 2008). A document called Index Supplement Cancer contains information and links to information that might be helpful for readers of this analysis.

Treatment Evaluations
Evaluations of specific treatments (ablation, other) will be described in future results documents.

Population Trends
Trends in lung cancer incidence and mortality under various scenarios (e.g., the Smoking Base Case) will be described in future results documents.
RES CALIB VALID

SUMMARY
This document summarizes the calibration of the LCPM and validation to additional endpoints.

RESULT TYPE
Other

OVERVIEW
Calibration and validation results indicate no immediately obvious departure from observed data, lending credence to simulations of hypothetical scenarios (i.e., those that extrapolate past observed data).

Many limitations of the LCPM are common to all studies employing simulation models. Tradeoffs must be made between increasing complexity and practical limits on the number of unknown parameters that can be identified using available data. A ‘deep’ model like the LCPM has more complexity (which allows us to evaluate different workup algorithms) than a ‘shallow’ statistical model that estimates transition probabilities (e.g., stage I to stage II), but at the cost of greater risk of identifiability problems. To reduce the risk of identifiability issues biasing results, we continue to select additional calibration targets, refine calibration approaches, remove parameters where possible, and identify additional sources of data for inputs.

METHODS
For calibration and validation, the LCPM was populated as described in the Population Component.

We chose a large cohort (white males aged 60 in 1990) as the initial calibration cohort, setting the birth cohort term to 1.0 (reference group). Once calibration to this large cohort was completed, the same parameter set was used to generate incidence by age predictions for cohorts of 50- and 70-year old white males and 50, 60, and 70-year old white females. To account for observed birth cohort trends in lung cancer risks and allow for differences in baseline risk by gender (85-87) we added a term that modifies the monthly risk of lung cancer development (all cell types), stratified by gender. The birth cohort term was adjusted in these cohorts such that the LCPM generated observed age-specific incidence rates.

See also Calibration Details.

RESULT
Fit to Calibration Targets
Primary Targets, Derived from SEER

The best fitting parameter set produces a good fit to incidence by age for cohorts of 50-, 60-, and 70 year-old whites (Out Calib1). We also achieved good fits to size, type, and stage distributions (Out Calib2). The best-fitting set slightly overestimated ≥3-year
survival for NSCLC stages I and II (Out Calib3). This overestimation is due to our assumption that all patients underwent guideline staging and treatment; many patients represented in SEER did not receive guideline treatments. Survival of patients with limited stage small-cell lung cancer was accurately predicted.

Secondary Targets, Derived from Cohort Studies and Literature
The LCPM predicted annual mortality rates per 100,000 non-smokers and lung cancer-specific mortality ratios for current (vs. never) smokers (by 5-year age group) that agree with observed data (Out Calib4).

The LCPM predicted lung cancer outcomes in non-smokers and in autopsy studies that agreed with published findings (Out Calib5).

Validation
Validation is documented here (Out Valid1 and Out Valid2).

DISCUSSION
After calibration and validation of the LCPM, the model could be used to evaluate screening programs.

Because screening is not part of usual clinical practice, most lung cancers in the SEER registry were diagnosed on the basis of symptoms. The SEER calibration targets used to inform estimates of incidence and survival were therefore supplemented with screening trial data to refine estimates regarding noninvasive BACs, which appear with greater frequency in screening studies.

CONCLUSION
The LCPM generates outputs consistent with multiple data sources. Predictions from the model regarding the effectiveness of screening or other interventions are however extrapolations beyond available data, and are subject to all assumptions built in to the model.

RELEVANT ASSUMPTIONS
The base case assumption that all individuals receive guideline care is necessary, given the lack of data on staging practices in the US. However, this assumption likely yields fewer understaged patients and therefore higher survival for early-stage cancers than in SEER data used for calibration.

See also Assumption Overview.

RELEVANT PARAMETERS
Parameters Natural History

RELEVANT OUTPUTS
See Output Overview.
SCREENING BIASES

Screening trial results are affected by several well-known biases that make interpretation of results challenging. A test that detects earlier-stage disease, will, by definition, prolong observed survival times (lead-time bias). Assuming some individual heterogeneity in disease progression rates, periodic screening will preferentially detect slowly progressing cases, simply because such cases persist longer in the asymptomatic state (length-time bias). The extreme of length-time bias, overdiagnosis refers to both screen detection of cases that would not have caused symptom detection or death (i.e., without screening, the person would die of competing causes, unaware of the presence of the disease) and to detection of pseudo-disease (e.g., cases with a self-resolving clinical course).

REFERENCES:

1 Reich, J. M. “Improved survival and higher mortality: the conundrum of lung cancer screening.[see comment].” in Chest 2002; 122: 1: 329-37
Calibration Details

Overview

Calibration was used to estimate values of unobservable natural history parameters and uncertain parameters (those for which literature estimates provided ranges of values). Each unique combination of model inputs (tables, distributions, and scalar values) is referred to as a parameter set. A combination of grid search and simulated annealing was used to identify the parameter set that minimized the total sum of squared errors between model output and 8 primary calibration targets. Of the 25 parameter sets with the best fit to the primary calibration targets, we chose the set with the best fit to 5 secondary calibration targets. Extensive debugging was performed throughout model development and prior to final calibration.

Please see Calibration Methods Research for information on a comparison of calibration approaches.

Cohorts

We chose a large cohort (white males aged 60 in 1990) as the initial calibration cohort, setting the birth cohort term to 1.0 (reference group). Once calibration to this large cohort was completed, the same parameter set was used to generate incidence by age predictions for cohorts of 50 and 70-year old white males and 50, 60, and 70-year old white females. To account for observed birth cohort trends in lung cancer risks and allow for differences in baseline risk by gender, we added a term that modifies the monthly risk of lung cancer development (all cell types), stratified by gender. The birth cohort term was adjusted in these cohorts such that the LCPM generated observed age-specific incidence rates.

Targets

See Output Overview for links to comparisons of targets and outputs from the calibrated LCPM.

Defining Ranges for Unobservable Natural History Parameters (see Natural History Component and Parameters Natural History)

During calibration, some parameter values could be ruled out as implausible, after consultation with clinical experts and past research. For example, the intercept terms were ordered to reflect observed risks of each cell type among non-smokers. Lung cancer risks increase with age and SY and decrease with YSQ. SY has the strongest effect on development of small cell cancers, and the effect of YSQ was weakest for adenocarcinoma. The amount of BAC as a proportion of adenocarcinoma was varied from 0 to 0.4 and estimated to be 0.2.

Initial values of symptom detection parameters were selected so that the cumulative probability of symptom detection from (true) distant metastases was nearly 1.0 by 3 years, i.e., very few patients had asymptomatic/undetected metastases at 3 years after diagnosis, but it was not impossible. By comparison, the estimated growth duration of
metastases was 3.8 years (faster growth than the primary tumor) in a published breast cancer model. Initial values of the intercept term and coefficient on tumor volume for symptom detection of primary cancers were chosen so that once a cancer passed the threshold size, the probability of symptom detection increased slowly to yield lung cancers of similar sizes as those observed in SEER. Adenocarcinomas were assumed to be less aggressive than SCLCs.

Defining Ranges for Uncertain Parameters

We classified parameters as uncertain if literature estimates provided ranges of values. During calibration, test characteristics were allowed to vary because verification bias likely affects many published sensitivity and specificity values. See Parameters Test Performance and Parameters Treatment for details.

See Res Calib Valid for model outputs after calibration and validation.

REFERENCES:

FOLLOW UP COMPONENT

As described in the Incidental Imaging Component, patients with incidentally detected nodules suspicious for lung cancer are triaged according to the size of the nodule. Patients with nodules over the threshold diameter are sent to the Workup And Staging Component.

In sensitivity analyses of a manuscript in press, we examined scenarios where nodules smaller than a cutoff size (e.g., 4mm diameter) are ignored, with patients returning to the general population.

Smaller nodules are followed-up with serial high-resolution CT exams (even in the absence of screening), with a specified periodicity (see below) over 24 months. Detection of new small nodules re-starts the 24-month follow-up sequence. Nodules that exhibit no detectable growth (see below) after 2 years of follow-up are diagnosed as benign; detectable growth on any subsequent imaging exam is considered sufficient to cause suspicion for lung cancer.

Structural Parameters in the Follow Up Component
(See Parameter Overview for definition):
1) A threshold (or ‘cutoff’) diameter of 8mm was used and is generally in agreement with a low (5%) biopsy rate for 4-9mm nodules in a recent trial.
2) Depending on the scenario, follow-up could occur with a fixed periodicity of 1, 3, 6, 12, and 24 months or be managed according to the size of the largest nodule found (similar to published algorithms from CT screening trials).
3) For the base case, the minimum detectable growth on sequential exams was 2mm.
4) An estimated 50% of growing nodules are excisionally biopsied using VATS (video-assisted thoracic surgery).

REFERENCES:
1 Benjamin, MS, Drucker, EA, McLoud, TC, Shepard, JO “Small pulmonary nodules: detection at chest CT and outcome” in Radiology 2002; 226: : 489-493
WORKUP AND STAGING COMPONENT

SUMMARY
This document describes the one-month cycle during which a patient with a pulmonary nodule suspicious for lung cancer undergoes clinical workup to establish a diagnosis of cancer or benign histology. Patients with cancer then undergo staging tests to establish the extent of disease progression.

OVERVIEW
Clinical algorithms for workup and staging are modeled explicitly, so that differences in patient management strategies can be compared.

Patients enter this component if they have a pulmonary nodule suspicious for lung cancer, detected by any modality and large enough to be biopsied. Patients with incidentally detected nodules smaller than the threshold go to the Follow Up Component.

Biopsy-confirmed malignancies are clinically staged (in the same cycle) based on practice guidelines from the National Comprehensive Cancer Network (NCCN, version 2000, for calibration to 1990-2000) and assigned both TNM and AJCC stages.

Patients whose pulmonary nodules are definitively diagnosed as benign start the next cycle in the general population state. Patients with a diagnosis of lung cancer begin the next cycle in the Treatment Component.

DETAIL
Patients presenting with symptom-detected cancers undergo biopsy to establish the histological type and a high-resolution CT examination to stage lymph nodes and determine tumor size, if not already known.

All patients undergo one high-resolution CT examination to determine calcification pattern and/or stage lymph nodes, per NCCN guidelines.

An estimated 50% of patients with N0/1 and evidence of primary tumor growth on CT are sent for excisional biopsies using VATS (video-assisted thoracic surgery). Non-operative candidates and remaining N0/1 patients undergo biopsy of the primary tumor (bronchoscopy for central nodules and TTNA for peripheral nodules).

Patients with clinically evident enlarged mediastinal lymph nodes (N2/3 on CT) undergo mediastinoscopy, which can establish a diagnosis of lung cancer and provide staging information. Patients with negative mediastinoscopy results are treated as N0/1.
RELEVANT ASSUMPTIONS

The assumption that all patients undergo staging tests in accordance with consensus practice guidelines is a limitation of the current LCPM. Many publications have demonstrated that a large proportion of lung cancer patients do not receive guideline treatments, but information on staging practices in the U.S. is limited.

Our assumption of guideline staging, which is often more invasive/aggressive than usual care, will result in higher survival rates for earlier staged cancers, because patients with more advanced disease (and lower survival) will be correctly identified and categorized into later stages.

RELEVANT PARAMETERS

Sensitivity and specificity of the diagnostic and staging tests will influence patient trajectories through this state (Parameters Test Performance).

Other input parameters define the scenario being simulated. For example, we will have the flexibility to simulate either guideline or usual care staging patterns.

RELEVANT COMPONENTS

The Workup and Staging Component includes sub-components specific for NSCLC and SCLC, which have different characteristics and therefore different staging systems.

A “Usual Care” version of the staging component is in development and will be available for comparison to guideline staging.

DEPENDENT OUTPUTS

Distributions of stage at diagnosis and survival outputs rely on this component. Natural History and Incidence do not depend on this component.

RELEVANT RESULTS
PARAMETERS NATURAL HISTORY

See the Natural History Component and Calibration Details for explanations of the way the LCPM models lung cancer natural history and how the unobservable parameters were estimated. This document contains additional detail not already provided.

Lung Cancer Development

The logistic model for development of a lung cancer was described in the Natural History Component. There are type-specific parameters for the effects of age and smoking history. We assume that beta_highrisk = ln(2), equivalent to a HR of 2 for individuals positive for the susceptible genotype, independent of age, gender, smoking status and pack-years. As a candidate ‘susceptibility’ genotype, we model the combined genotype of GSTP1 (GG) and p53 (Arg/Pro or Pro/Pro) to occur with an estimated population frequency of 4.7% (no linkage).\(^1\)

The amount of BAC as a proportion of adenocarcinoma was varied from 0 to 0.4 during calibration and was estimated to equal 0.2.

Lung Cancer Growth

See Table Growth Parameters. Lung cancers was assumed to grow 2-fold faster in smokers, although the difference may be due at least in part to ‘type mix’, i.e., non-smokers are more likely to develop slow-growing adenocarcinomas.\(^2,3,4\) The growth of BACs was truncated at 1.0 cm diameter (detectable by CXR). For non-BACs, we assumed a maximum possible tumor size of 27.7 cm,\(^5\) consistent with the largest reported size of 20.1-30.0 cm diameter in the SEER\(^\ast\)Stat database.

Lung Cancer Progression

Final mean (SD) unadjusted threshold volumes for SCLC ranged from 0.61 (0.65) ml for N1 involvement to 4.07 (4.13) for N3 involvement and 4.71 (4.14) for distant metastases. Corresponding unadjusted threshold volumes for NSCLC ranged from 3.34 (4.09) ml for N1 involvement to 3.8 (4.64) for N3 involvement and 2.62 (3.18) for distant metastases. Adjustments were then estimated via calibration to allow the propensity to progress to vary by cell type and be correlated with the growth parameter assigned to the person’s cancer.

Symptom Detection

See the Symptom Detection Component. The cumulative probability of symptom detection from (true) distant metastases was over 95% by 3 years. By comparison, the estimated growth duration of metastases was 3.8 years in a breast cancer model.\(^6\)

REFERENCES:


3 Song, P, Sekhon, HA, Jia, Y, Keller, JA, Blusztajn, JK, Mark, GP, Spindel, ER “Acetylcholine is synthesized by and acts as an autocrine growth factor for small cell lung carcinoma” in Cancer Research 2003; 63: 214-221


PARAMETERS TEST PERFORMANCE

Diagnostic test characteristics (sensitivity and specificity) determine the probabilities of detection and/or diagnosis of lung cancers or benign lesions.

Pulmonary Nodules
Sensitivities of imaging examinations for peripheral pulmonary nodules are input from tables indexed by diameter of the lesion. Sensitivities for a central lesion of the same diameter were assumed to be 25% lower than those for peripheral lesions. We derived sensitivities from the literature to test during calibration.

Computed Tomography (CT)
Helical CT was estimated to have a sensitivity of 0.63 for 1-4mm peripheral nodules, 0.77 for 4-8mm peripheral nodules, and 1.0 for peripheral nodules >8mm. High-resolution CT was assumed to have equivalent sensitivity for detecting presence of a nodule as helical CT (by size), but to have greater resolution for calcification patterns. As in clinical practice, an estimated 11% of benign nodules are diagnosed by high-resolution CT as benign on the basis of calcification pattern (not explicitly modeled). CT could occur in several components: Incidental Imaging Component, Workup And Staging Component, Screening Component, and during surveillance for recurrent disease in the Survival Mortality Component.

Chest X-ray (CXR)
The sensitivity of CXR was estimated to be approximately 25% to 50% of that of helical CT, and to be less than 1.0 at 16mm. The minimum detectable size was assumed to be 7.5mm. CXR could occur in the Incidental Imaging Component.

Specificity for both helical CT and CXR is assumed to be 0.98 (per person, or 0.997 per nodule). Specificity of high-resolution CT for pulmonary nodules was assumed to be 1.0.

Staging Examinations
Computed Tomography
High-resolution CT was estimated to have a sensitivity for nodal involvement of 0.63 (average of N1, N2/3) and a specificity of 0.6. \(^1,^2\) Bronchoscopy
Sensitivity of bronchoscopy is defined as the probability of establishing a definitive diagnosis on the basis of cells recovered from the nodule. The sensitivity increases with increasing size of the nodule. For cancer, the sensitivity is 5% for nodules less than 20mm diameter, 20% for nodules 20-29mm diameter, and 48% for nodules 30-40mm diameter. Establishing a specific diagnosis (of the many possible) for a benign nodule is more difficult; the sensitivities are lower for benign nodules of the same size. \(^3\) Bronchoscopy was assumed to have a sensitivity of 0.5 for malignant nodal involvement. \(^4\) Mediastinoscopy
Sensitivity of mediastinoscopy for cancer in patients with enlarged lymph nodes is estimated at 0.92 (range, 0.88, 0.94), \(^2,^5\) and operative mortality is estimated at 0.3%. \(^5\) Reflecting common practice of not initiating therapy without pathological proof of lung cancer, we assume perfect specificity for mediastinoscopy. \(^5\) TTNA
The sensitivities of trans-thoracic needle aspiration (TTNA) for malignancy and benign diagnoses were indexed by the size of the nodule, informed by literature estimates. \(^6,^7,^8\)
VATS

VATS is assumed to have perfect accuracy at identifying malignant vs. benign disease and to include sampling or removal of nodes for confirmation of stage (perfect sensitivity and specificity). VATS had an operative mortality of 0.5%.

Metastases

A generic test for staging metastases (i.e., those not diagnosed on the basis of symptomatic presentation) is modeled with perfect specificity. Sensitivity for metastases was estimated during calibration; we tested functions of time since metastases developed and constant values of 0.4 to 0.5 derived from published sensitivities for bone scintigraphy and brain/bone MRI.

Omitted Tests

We omitted sputum cytology due to its low sensitivity relative to bronchoscopy and positron emission tomography (PET) staging because it was uncommon during our calibration period (1990 to 2000).

REFERENCES:

1 Beadsmoore, C. J., Screaton, N. J. “Classification, staging and prognosis of lung cancer” in European Journal of Radiology 2003; 45: 1: 8-17
3 Weinberger, SE “Differential diagnosis and evaluation of the solitary pulmonary nodule” in UpToDate 2004; version 12:
4 Mandel, J, Weinberger, SE “Overview of non-small cell lung cancer staging” in UpToDate 2005; version 13:
8 Wallace, MJ, Krishnamurthy, S, Broemeling, LD, Gupta, S, Ahrar, K, Morello, FA, Hicks, ME “CT-guided percutaneous fine-needle aspiration biopsy of small (less than or equal to 1-cm) pulmonary lesions” in Radiology 2002; 225: : 823-828
9 Deterbeck, FC, Rivera, MP “Clinical presentation and diagnosis” in Diagnosis and Treatment of Lung Cancer 2001;
PARAMETERS TREATMENT

Eligibility for Surgery
Individuals were randomly assigned as ineligible for surgical resection based on proportions of NSCLC stage I and II (all ages) cases that did not undergo surgery (where the reason was documented). Surgery was explicitly contraindicated for 5.6%, and offered but refused in 2.1% of cases. (Estimated from public release files using SEER*Stat 4.2.3 software.) We allow a small proportion (base case 13%, SEER-Medicare) of LS to be resected, reflecting the minority of cases which present with localized SCLC.

Resection
Effectiveness of resection is incorporated as follows: a person with no occult metastases whose single primary cancer is resected is assigned competing risks consistent with a person of the same smoking history – not stage I survival from SEER. However, if a second, undetected primary tumor remains (in a non-resected lobe), lung cancer can recur (see Natural History Component). The presence of undetected micro metastases is likely the cause of the poor observed survival after “curative” resection in many patients. \(^1\) Removal of (or sampling from) nodes at resection can result in re-assigning stage at diagnosis, but provides no survival benefit. \(^2\) The base case operative mortality rate for lobectomy is estimated at 4% \(^3\) (value in sensitivity analysis, 3%). No increase in mortality due to late (post-30 day) effects of surgery (e.g., infection) was modeled.

Systemic Therapies
Parameter values that define efficacies of chemotherapy and radiotherapy are the probabilities of complete or partial response, using the definition of complete as no visible cancer at 4 week follow-up and partial as greater than or equal to a 30% decrease in diameter. \(^4\) Probabilities of complete and partial responses vary by histologic type, with estimates taken from the literature. A cancer that partially responds to therapy is decreased in diameter by 30%, and a cancer that completely responds to therapy is reduced to 1.5mm diameter, \(^4\) or below the 2mm detection threshold assumed for helical CT. Based on the new diameter, an adjusted ‘time since cancer developed’ is calculated and used to increment growth in all future cycles, retaining the original growth parameter alpha (Table Growth Parameters). To account for differences in growth rates of cancers pre- and post-therapy, we included parameters (estimated during calibration) that allowed faster-growing cancers to be more or less likely to respond to therapy.

Surveillance
Surveillance for recurrent lung cancer is modeled as helical CT at 6, 12, 24, 36, 48, and 60 months. \(^5\)

REFERENCES:


INCIDENTAL IMAGING COMPONENT

During each cycle spent in the general population, persons may undergo imaging exams (thoracic CT, or CXR) performed for reasons unrelated to screening for lung cancer.

We fit generalized linear models to insurance claims data from 1999. Increasing age predicted higher likelihood of both CT and CXR (p

Imaging results are compared to results of prior imaging exams, if available. Persons with no detected nodules or exclusively stable nodules return to the general population. Persons with newly detected nodules undergo follow-up and are managed according to the size of the largest nodule found. Persons with nodules large enough for biopsy start the next cycle in the Workup And Staging Component. If the new nodule is smaller than the threshold diameter, the person begins the following cycle in the Follow Up Component. A threshold diameter of 8mm was used as a proxy for clinical practice and is generally in agreement with a low (5%) biopsy rate for 4-9mm nodules in a recent trial. ¹

See Parameters Test Performance for details on test characteristics of imaging examinations.

REFERENCES:

SYMPTOM DETECTION COMPONENT

Symptom detection can occur via symptoms from the primary cancer or from distant metastases.

Each month, individuals with distant metastases and/or a primary lung cancer (but not those with exclusively benign nodules) may develop symptoms that result in lung cancer detection and begin the following cycle in the Workup And Staging Component. The probability of symptom detection from primary cancers varies by location (central cancers have a greater propensity to cause symptoms, given size) and cell type (NSCLC vs. SCLC) and is a logistic function of the size of the largest cancer. We assume the minimum diameter for peripheral cancers to cause symptoms is 10mm, approximately the size at which airways are obstructed. The probability of symptom detection from metastases is a logistic function of the months since metastases developed (varied by NSCLC vs. SCLC).

Symptom detection parameters were estimated during calibration (see Calibration Details). The most relevant calibration targets included the proportion of lung cancers detected via symptomatic presentation, the stage and type distributions of incident cancers, and survival by type and stage at diagnosis. The background rate of chest imaging in the population (Incidental Imaging Component) will also influence the symptom detection rates.

See also the Assumption Overview.
**Benign Component**

The benign component occurs in every cycle of the model, so that benign lesions may develop throughout life. Benign lesions may also spontaneously resorb (e.g., due to clearing an infection).

To incorporate the costs and risks of follow-up procedures for benign lesions, the natural history model allows up to 3 benign lesions (with no regard for histological type) per person.

**Overview**

Using a polynomial fit to raw Mayo CT screening trial prevalence data, we estimated the average number of benign lesions per person, by age. Cumulative incidences of one or more benign lesions were converted to age-specific annual probabilities of developing new lesions (ages 35 to 68) and of existing lesions regressing (after age 68). For simulations of Mayo CT participants, therefore, prevalence of benign nodules was as observed in the study.

We estimate that 3% of benign nodules exhibit growth, and assign these lesions doubling times consistent with adenocarcinomas. Non-growing benign lesions are modeled as appearing fully formed in the previous month, consistent with a range of biological mechanisms (e.g., pleural effusion, edema, and infection).

Each benign lesion’s location (i.e., specific lobe) was assigned based on a study of (n=185) nodules from the Mayo Clinic, which stratified by right/left. An indicator for central/peripheral is randomly assigned. Size (diameter) was derived from the Mayo Clinic data, expressed as a lognormal distribution (mean = 0.9, variance 0.36).

**Assumptions**

Solidly calcified lesions are not considered suspicious for lung cancer, and are ignored.

The probability of developing benign lesions is assumed independent of smoking history; no significant correlations were observed between any of the smoking factors and numbers of lesions in the Mayo CT data; and to our knowledge, no literature sources refute this observation.

**Extrapolating from Mayo CT study data on prevalence of benign nodules**

The base case LCPM incorporates no regional variation in the prevalence of benign nodules. Infection with histoplasmosis is a common cause of small (less than 3mm diameter) benign nodules. Histoplasmosis rates vary geographically, with nearly 100% prevalence in persons residing in the major river valleys of the central U.S. The Mayo Clinic (Rochester, MN) is not in an area of the highest histoplasmosis rates.

**REFERENCES:**

## Table Growth Parameters

### Natural History Parameters

Distribution of alpha parameters (rate of decay of growth rate) used in Gompertz, and doubling times (in days) for lung cancers at various sizes (cm diameter), by cell type.

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Distribution of alpha parameter</th>
<th>Mean (SD) DT at 0.5cm</th>
<th>Mean (SD) DT at 1.0cm</th>
<th>Mean (SD) DT at 1.5cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma/ BAC</td>
<td>logN(-7.765, 0.5504)</td>
<td>187(160)</td>
<td>227(194)</td>
<td>260(222)</td>
</tr>
<tr>
<td>Large cell</td>
<td>logN(-6.59942, 0.68862)</td>
<td>61(61)</td>
<td>74(74)</td>
<td>85(85)</td>
</tr>
<tr>
<td>Small cell</td>
<td>logN(-5.44357, 0.611485)</td>
<td>19(16)</td>
<td>23(20)</td>
<td>26(23)</td>
</tr>
<tr>
<td>Squamous cell</td>
<td>logN(-6.6111, 0.7935)</td>
<td>65(72)</td>
<td>79(87)</td>
<td>90(100)</td>
</tr>
<tr>
<td>Other</td>
<td>logN(-6.714, 0.6634)</td>
<td>67(66)</td>
<td>81(80)</td>
<td>93(92)</td>
</tr>
</tbody>
</table>

Notes: Other refers to NSCLC not otherwise specified; DTs shown are for all cancers at specified size (both diagnosed and undiagnosed) and do not reflect adjustment for smoking status.

See Parameters Natural History and Natural History Component for details.
The LCPM study’s protocols for use of human subject data underwent expedited review (secondary use of medical records) and was approved by the human subjects institutional review board as compliant with HIPAA guidelines.

De-identified records (including demographics, smoking histories, and screen results) from two studies were provided to our institution for model calibration and validation.

- 1,520 current and former smokers aged 50-85 years participating in a Mayo Clinic study of annual CT screening for early detection of lung cancer\(^1\). Participants signed informed consent waivers approved by the Mayo Clinic institutional review board before enrollment in the screening study. Transfer of the de-identified data was approved by both institutions’ human subjects review boards and was exempt from further informed consent requirements.

- 3,318 current and former smokers aged 55-74 years participating in the Lung Screening Study (LSS, a pilot study for the National Lung Screening Trial) of annual CT or CXR screening for early detection of lung cancer\(^2\). Transfer of the de-identified data was exempt from further informed consent requirements.

REFERENCES:


CALIB3

Calibration Endpoints - Survival curves for NSCLC (left) and SCLC (right).

Notes: Our use of published estimates for response rates from systemic therapies and our assumption that all patients underwent guideline staging and treatment may be irreconcilable with observed survival used for calibration targets, because many patients represented in SEER did not receive guideline treatments. Defining survival calibration targets that vary by treatment as well as stage, or by finer gradations of stage (i.e., T1N0M0 vs. T1N1M0) would address this limitation (ongoing work).
Figure 1. Age- (in single years), gender-, race-, and calendar year-specific lung cancer incidence rates, derived from SEER*Stat case listing files and counts of populations at risk from the NCI. Shown are incidence rates of all lung cancer types combined, for cohorts of whites. Shaded regions are acceptance windows (95% CIs) around SEER calibration targets; crosses indicate LCPM output.
OUT CALIB2

Size, Type, and Stage Distributions of Incident Cancers; LCPM vs. SEER

The LCPM predicted a mean size of incident cancers of 29mm, vs. 38mm in SEER.

Figure 3 (Left). Calibration to type distribution of incident lung cancers, white males 60-70 years; 1990-2000. Derived from SEER*Stat case listing files, stratified by gender, race, calendar decade, and 10-year age group.

Figure 4 (Right). Calibration to stage distribution of incident lung cancers, white males 60-70 years; 1990-2000. NB stratified by NSCLC and SCLC. Derived from a weighted average of 3 studies (all genders, all races) and SEER data (stratified by gender, race, decade, and age group).
Mortality for individuals with similar smoking histories

Lung cancer incidence rates in male smokers predicted by the LCPM were within the 95% confidence bands observed in the Health Professionals Follow-up Study of male physicians\(^1\) with similar smoking histories. Similarly, lung cancer mortality rates in male smokers predicted by the LCPM were within calculated binomial 95% confidence bands in CPS II participants\(^2\) with similar smoking histories.

REFERENCES:


Reproduction of observed endpoints in the LSS screening trial
LCPM Populated with Lung Screening Study Population in Presence of Screening

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Study result²</th>
<th>LCPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>participants with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>positive baseline CT screen</td>
<td>20.5%</td>
<td>20.9%</td>
</tr>
<tr>
<td>lung cancer at baseline CT screen</td>
<td>1.9% (95% CI, 1.2%, 2.6)</td>
<td></td>
</tr>
<tr>
<td>prevalent lung cancers that were:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenocarcinoma</td>
<td>63% (n=16/30)</td>
<td>73.7%</td>
</tr>
<tr>
<td>small cell</td>
<td>3%</td>
<td>3.1%</td>
</tr>
<tr>
<td>NSCLC, NOS</td>
<td>10%</td>
<td>6.0%</td>
</tr>
<tr>
<td>prevalent lung cancers that were:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stage I</td>
<td>53% (n=16/30)</td>
<td>67.1%</td>
</tr>
<tr>
<td>stage II</td>
<td>10%</td>
<td>8.0%</td>
</tr>
<tr>
<td>stage III</td>
<td>20%</td>
<td>18.6%</td>
</tr>
<tr>
<td>stage IV</td>
<td>10%</td>
<td>6.3%</td>
</tr>
<tr>
<td>unstaged</td>
<td>7%</td>
<td>n/a</td>
</tr>
<tr>
<td>diameter of prevalent lung cancers:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>median (mean, SD)</td>
<td>18mm (27, 23)</td>
<td>10mm (14.6, 8.4)</td>
</tr>
<tr>
<td>patients with prevalent lung cancers:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean cigarettes/day (SD)</td>
<td>27.7 (9.3)</td>
<td>32.4</td>
</tr>
<tr>
<td>mean years of smoking (SD)</td>
<td>47.3 (4.8)</td>
<td>50.3</td>
</tr>
<tr>
<td>proportion male</td>
<td>0.57</td>
<td>0.58</td>
</tr>
<tr>
<td>participants with lung cancer detected at screen #2</td>
<td>0.57% (n=8/1398)</td>
<td>0.29%</td>
</tr>
</tbody>
</table>

Notes: LCPM-predicted endpoints calculated from 250,000 simulated participants. Endpoints not provided in references were calculated directly from study data (see Protected Health Information). CI = confidence interval; SD = standard deviation; positive CT screen defined as detection of at least one non-solidly calcified pulmonary nodule at least 4mm in diameter. Retrospectively identified nodules not included in prevalence estimate. Adenocarcinoma includes bronchioloalveolar carcinoma (BAC) and mixed adenocarcinoma/BAC.

REFERENCES:

In a manuscript in preparation, Kong, et al. compare approaches for calibration of the LCPM, using an earlier model version that included only 4 lung cancer cell types.

Since completion of the work described in this manuscript, we have added a 5th cell type (‘Other’, represented by ICD-O-2 code 80103) to both the Natural History Component and the calibration targets for age-specific incidence lung cancer incidence.
INDEX SUPPLEMENT CANCER

This document is intended as a guide to this Model Profiler for readers of "Adopting helical CT screening for lung cancer: Potential health consequences over a fifteen-year period."

Overview
A description of the model structure, major components, and purpose is available in the Model Overview

Input parameters and sources
Test characteristics and mortality risks are described in Parameters Test Performance
Treatment effects and mortality risks are described in Parameters Treatment
Natural history parameters are described in Parameters Natural History
Smoking histories were derived from survey data as described in the Population Component

Assumptions
The Assumption Overview describes the major assumptions underlying the LCPM and their possible implications.

Omissions from current LCPM that may influence the estimate of screening effectiveness
Harms and benefits from incidental detection of other diseases (e.g., other cancers) found at screening
Tumor seeding of surgical or biopsy site, which is mainly described in case studies (e.g., Raja and Bessman, JCO 2003) and is thought to be a rare event
In these studies, a small minority of ‘late’ (>30 day) post-resection deaths were due to late surgical mortality (1 infection/16 late deaths in Handy, et al., and 5 late surgical mortality death/51 total late deaths in Toker, et al.). A higher proportion of late mortality was due to cancer progression (7/16 in Handy and 16/51 in Toker), which the LCPM models explicitly as a cause of post-operative death. Other-cause deaths due to respiratory failure and heart disease are to a large extent captured by the increased competing mortality risks faced by smokers in the LCPM.

Economic consequences (costs) and influence of screening on quality of life (QALYs) were not considered in this analysis.
Secondary Calibration targets

Two targets were derived from an earlier (1980s) cohort study with a highly selected population (i.e., volunteers with the American Cancer Society or their friends) that experienced only 70% of the all-cause mortality in the general U.S. population.

Figures 7 and 8 show that the LCPM predicted annual mortality rates per 100,000 non-smokers and lung cancer-specific mortality ratios for current (vs. never) smokers (by 5-year age group) that agree with observed data.

REFERENCES:


Lung Cancer in Non-Smokers

The LCPM predicted a percentage of non-smokers among lung cancer cases of 5.4%, in the range of reported values of 2% to 15%\(^1\),\(^2\). As expected, the model predicted a lower proportion of SCLC cases among non-smokers (4.3%) than among all lung cancer cases (18%)\(^3\).

Lung cancer detected at autopsy

Estimates of rates of undetected (“surprise”) lung cancers at autopsy range from 0.34% to 55%\(^4\),\(^5\),\(^6\),\(^7\),\(^8\),\(^9\),\(^10\),\(^11\),\(^12\). Furthermore, autopsy techniques varied in unknown ways and there was no way to correct for potentially large biases due to unreported variability in age ranges and case mix (especially smoking prevalence and the selection bias inherent in autopsy series)\(^13\). Assuming that all lung cancers >15mm diameter would be diagnosed on autopsy, the LCPM predicts a 3.6% autopsy surprise rate, in the reported range.

REFERENCES:

5 Saracci, R. “Problems with the use of autopsy results as a yardstick in medical audit and epidemiology” in Quality Assurance in Health Care 1993; 5: 4: 339-44
10 Sternby, NH “The role of autopsy in cancer registration in Sweden, with particular reference to findings in Malmö” in Autopsy in Epidemiology and Medical Research, International Agency for Research on Cancer 1991;
KEY REFERENCES


Detterbeck, FC, Rivera, MP (2001) Clinical presentation and diagnosis in Diagnosis and Treatment of Lung Cancer


Health and Human Services (1990) The Health Benefits of Smoking Cessation


Mandel, J, Weinberger, SE (2005) Overview of non-small cell lung cancer staging in *UpToDate* version 13:


Saracci, R. (1993) Problems with the use of autopsy results as a yardstick in medical audit and epidemiology in *Quality Assurance in Health Care* 5:, p 339-44


Sternby, NH (1991) The role of autopsy in cancer registration in Sweden, with particular reference to findings in Malmö in *Autopsy in Epidemiology and Medical Research, International Agency for Research on Cancer*


Thun, M. J., Myers, D. G., et al. (1997) Chapter 5. Age and the exposure-response relationships between cigarette smoking and premature death in Cancer Prevention Study II. National Cancer Institute, Smoking and Tobacco Control, Monograph 8:


Usuda, K., Saito, Y., et al. (1994) Tumor doubling time and prognostic assessment of patients with primary lung cancer in *Cancer* 74 : p 2239-2244


Wallace, MJ, Krishnamurthy, S, Broemeling, LD, Gupta, S, Ahrar, K, Morello, FA, Hicks, ME (2002) CT-guided percutaneous fine-needle aspiration biopsy of small (less than or equal to 1-cm) pulmonary lesions in *Radiology* 225 : p 823-828

Weinberger, SE (2004) Differential diagnosis and evaluation of the solitary pulmonary nodule in *UpToDate* version 12:


