FLEXKB DOCUMENT
Version: HI.001.11302018.9737
Document generated: 11/30/2018

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

# RAND Corporation

**Important note:** This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. The CISNET model profile topics are not necessarily meant to be read in sequential fashion, so the reader should feel free to skip around as their interests dictate.

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

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

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

Go directly to the: Reader's Guide.

# READERS GUIDE

**Core Profile Documentation**

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

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

### Model Purpose
This document describes the primary purpose of the model.

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

### Assumption Overview
An overview of the basic assumptions inherent in this model.

### Component Overview
Describes the basic computational building blocks (components) of the RAND model.

- Natural History Component
- Adenoma Risk Component
- Transition To Preclinical CRC Component
- Transition To Clinical CRC Component

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

### Output Overview
Describes basic model outputs. Because we output complete information for our simulated population, we are free to choose a wide range of model outputs. Current model outputs are driven by comparisons with other CISNET models.

### Results Overview
A guide to the results obtained from the model. At this time, our focus is on Bayesian calibration of model parameters.

### Key References
A list of references used in the development of the model.

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

# MODEL PURPOSE

## SUMMARY

The Colorectal Cancer Simulated Population model for Incidence and Natural history (**CRC–SPIN**) was developed to explore trends in colorectal cancer (CRC) incidence and mortality, to compare the effectiveness of different screening modalities, and to extend results from clinical trials to mortality endpoints.

## PURPOSE

CRC–SPIN contains three components that are used in combination to predict outcomes: a natural history model, a calibration component, and a screening component.

1) The **natural history model** describes the development of adenomas, preclinical cancers, clinically detected cancers, and survival after detection. *The purpose of the CRC–SPIN natural history model is to parsimoniously describe the natural history of colorectal cancer.* (see Model Overview, the Natural History Component provides a brief description).

2) The **calibration component** is used to combine information from multiple targets to select good natural history model parameters. CRC–SPIN 1.0 was calibrated using an approximate Markov Chain Monte Carlo approach, with calibration based on data likelihoods (Rutter, Miglioretti, Savarino, 2009). CRC–SPIN 2.0 and later model versions were calibrated using an Incremental Mixture Approximate Bayesian Computation approach, which is a likelihood–free approach (Rutter, Ozik, DeYoreo, Collier, under review). The Bayesian methods used for calibration result in a sample from the posterior parameter distribution, which can be used to estimate the uncertainty of model predictions. *The purpose of the CRC–SPIN calibration component is to provide an objective, automated, data–based method for calibrating natural history model parameters. A secondary purpose is to obtain posterior distribution estimates of model parameters that can be used to describe uncertainty in model predictions.*

3) The **screening component** simulates the action of screening tests by simulating the occurrence of and outcomes from screening tests. CRC–SPIN simulates test performance that depends on disease characteristics. For example, the sensitivity of colonoscopy depends on lesion size. *The purpose of the CRC–SPIN screening component is to simulate and then compare model–predicted outcomes (such as incidence and mortality) under a range of screening scenarios.*

# Model Overview

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

## Summary

The CRC–SPIN model simulates colorectal cancer disease trajectories for agents that are part of population, cohort, or sample.

## Purpose

CRC–SPIN is used to examine the effect of screening on colorectal cancer (**CRC**) outcomes, including incidence and mortality. Additional details are provided in Model Purpose.

## Background

CRC is the third leading cause of cancer death in the United States. Randomized controlled trials (RCTs) have shown that screening for CRC – using either fecal occult blood tests (FOBT) or flexible sigmoidoscopy – can reduce CRC–mortality. Evidence of the effectiveness of other tests is inferred from operating characteristics (sensitivity and specificity) or from observational studies (e.g., case–control studies). Although RCTs are a gold standard for evaluating the effectiveness of screening tests, it is impractical or impossible to use RCTs to answer the full range of health policy questions about CRC screening. This includes questions about the relative effectiveness of different CRC screening strategies. Microsimulation models like CRC–SPIN provide a method for addressing a broad range of questions about CRC screening.

## Model Description

CRC–SPIN is a microsimulation model: it simulates individuals (or 'agents') but not interactions between agents. For each agent, CRC–SPIN simulates disease trajectories over a lifetime, including the occurrence and growth of adenomas, transition of adenomas to preclinical CRC, death from CRC, death from other causes, and the actions of screening on these trajectories. CRC–SPIN 1.0, developed in C#, is now retired. CRC–SPIN was updated and recalibrated in January 2018. CRC–SPIN 2.0 and beyond are written in R.

CRC–SPIN has four components:

1. adenoma risk;
2. adenoma growth;
3. transition from adenoma to preclinical cancer; and
4. transition from preclinical to clinical cancer (sojourn time).

Once CRC is clinically detected, CRC–SPIN stochastically assigns stage and size at detection, and survival given stage at detection. CRC–SPIN simulates events in continuous time, and simulates continuous (rather than categorical) adenoma and cancer size.

CRC–SPIN components are described in detail in the Component Overview. Below, we describe key assumptions, model inputs, and model outputs.

**Key Assumptions:** CRC–SPIN 2.x is built on the assumption that all CRC arises through the adenoma–carcinoma process. (A CRC–SPIN version that incorporates the sessile serrated polyp disease pathway is under development.) Another key

assumption is that neither adenomas nor cancers regress, though adenomas can grow very slowly. CRC–SPIN specifies a minimum adenoma size of 1mm and a maximum adenoma size of 50mm. (Cancer size may be larger.)

Agents may develop multiple adenomas. Every adenoma has the potential to develop into preclinical cancer, so that agents may develop multiple colorectal cancers, and thus may have multiple hypothetical cancer death times. In the absence of screening, the first clinically detected cancer determines CRC survival. In the presence of screening, the first screen– or clinically–detected cancer determines CRC survival, though the removal of adenomas may prevent their transition to CRC.

CRC–SPIN incorporates the overall effect of changes in treatment on CRC survival, which is simulated using a model based on analysis of SEER data. CRC survival is a function of age, sex, cancer location (colon or rectum), stage, and year of diagnosis. CRC–SPIN does not simulate the impact of specific treatments on colorectal cancer outcomes, though this is a potential model extension.

CRC–SPIN is a 'parallel universe' model, and simulates outcomes for the same population under different screening scenarios. The screening component can accommodate complex screening scenarios, and is easily extended to incorporate new screening modalities. The screening component includes two general types of screening tests: an agent–level test that provides a single result for each agent, and a structural exam that provides a result for each adenoma (and an overall agent–level false positive rate). Colonoscopy is a special type of structural exam that can remove adenomas and preclinical cancers. (In some simulations lesions may also be removed at flexible sigmoidoscopy.) Agent–level tests include fecal–based tests of all types (gFOBT, FIT, stool DNA) and could include blood–based tests. Structural tests include flexible sigmoidoscopy, colonoscopy, CT colonography, and could include capsule tests.

A more detailed description of CRC–SPIN assumptions is described in Assumption Overview.

**Model Inputs:** CRC–SPIN includes relatively few calibrated parameters (v1.0: 23, v2.x: 22, see Parameter Overview). Model *inputs* refer to information, based on empirical data, that is directly passed to the model. CRC–SPIN model inputs are:

- the distribution of adenomas over the large intestine (based on both autopsy and screening colonoscopy studies),
- the size distribution of clinically detected CRC (based on 1979 SEER data, prior to the diffusion of screening),
- the stage distribution of clinically detected CRC (based on 1979 SEER data), and
- CRC relative survival, a function of stage, location (colon or rectum), sex, and age at diagnosis (based on SEER data, see Rutter, Johnson, Feuer, et al, 2013)

Additional information about model inputs is provided in the Assumption Overview.

**Model Outputs:** CRC–SPIN simulates life events histories for agents both with and without screening. Generated model outputs include: the prevalence and number of adenomas across agents, rates of preclinical cancer, rates of clinical cancer (by location, sex, and age), and mortality rates (by location, sex and age). Additional information is provided in Output Overview.

# CONTRIBUTORS

Carolyn Rutter (PI), RAND Team (in alphabetic order:) Maria DeYoreo, Florentine Eloundou, Chris Maerzluft, Angel Martinez

past contributors: Jessica Hwang, Eric Johnson, Tracey Marsh, Diana Miglioretti, Chester Pabiniak, Jim Savarino

# ASSUMPTION OVERVIEW

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

## SUMMARY

This document describes basic assumptions made by the CRC–SPIN natural history and screening models.

## BACKGROUND

Microsimulation models are complex and require assumptions about the functional form governing simulated process.

## ASSUMPTION LISTING

The most basic model assumption is that all cancers arise from adenomas. (Development of a CRC–SPIN version that incorporates the sessile serrated pathway is underway.) Model assumptions for each CRC–SPIN component are described below.

### Detailed Description of Model Assumptions

*Adenoma initiation assumptions*

CRC–SPIN uses a non–homogenous Poisson Process to simulate adenoma occurrence.

- Adenoma risk systematically varies with age and sex. Calibration of a CRC–SPIN model that allows adenoma risk to vary systematically by race is underway.
- Adenoma risk stochastically varies across agents (some have higher risk than others), and the distribution of agent–level baseline log–risk follows a normal distribution.
- Adenomas are independently located within each agent's large intestine (e.g., we do not model an agent–level tendency to develop adenomas in a specific location.)
- The distribution of adenomas across the large intestine is uncalibrated and is based on findings from 9 autopsy studies and one colonoscopy study not included as calibration data. We assume that P(cecum)=0.08, P(ascending colon)=0.23, P(transverse colon)=0.24, P(descending colon)=0.12, P(sigmoid colon)=0.24, P(rectum) = 0.09.

*Adenoma growth assumptions*

CRC–SPIN simulates adenoma growth using a Richard's growth model (Tjørve, Tjørve, 2010). This model includes the Janoschek model (used in CRC–SPIN 1.0).

- Adenoma growth *parameters* are constant over time. This does not imply constant growth of adenomas over time.
- Adenomas do not regress.
- Adenoma growth parameters are independent within agents (for agents with multiple adenomas).
- The time to 10mm follows a Frèchet (type II extreme value) distribution.
- The minimum adenoma size is 1mm.
- The maximum adenoma size is 50mm.

*Size at transition to cancer assumptions*

CRC–SPIN simulates the size of adenoma transition to cancer using log–normal distribution.

- Cancer first grows within an adenoma, with the lesion size only increasing once the cancer 'overtakes' the adenoma.
- Most adenomas do not transition to cancer, and so most adenomas do not reach their simulated transition size.
- The minimum cancer size (size at transition) is 0.5mm. This is less than the 1mm adenoma size. Therefore, the minimum malignant lesion size is 1mm.
- The probability of transition to cancer is a function of adenoma size, sex, age at initiation, and location (colon v. rectum). (Calibration of a CRC–SPIN model that allows the size at adenoma transition to preclinical CRC to vary systematically by race is underway.)

*Cancer growth assumptions*

- Cancerous lesions grow exponentially. The exponential growth rate is a function of the size at transition (0.5mm), the size at clinical detection, and the time from initiation to clinical detection (sojourn time). Given the exponential model and the cancer growth parameter, cancer size can be calculated at any time during the preclinical detectable phase.

*Sojourn time and stage at detection assumptions*

- Sojourn time depends only on location within the large intestine (colon or rectum), and is independent across cancers within agents.
  - CRC–SPIN V1.0 used a log–normal model for sojourn time.
  - CRC–SPIN V2.x uses a Weibull model for sojourn time, with calibrated shape and location parameters, and a proportional hazards model used to capture differences in sojourn time by location.
    Calibration of a CRC–SPIN model that allows sojourn time to vary systematically by race is underway.
- Stage and size at clinical detection are model inputs, and are based on 1979 SEER data that describe stage and size at clinical detection. The method for incorporating this information into the model depends on the version.
  - CRC–SPIN V1.0 simulated size at clinical detection then stage at clinical detection given size.
  - CRC–SPIN V2.x simulates stage at clinical detection then size at clinical detection given stage at detection.
- Size and stage at clinical detection are model inputs and are based on the SEER distribution of cancer size in 1975–1979, years prior to widespread CRC screening. CRC–SPIN 2.x uses stage at detection and the size at clinical detection stratified by stage. CRC–SPIN 1.0 used the overall size distribution and stage at detection given size at detection.

*Survival assumptions*

- CRC survival depends only on stage at diagnosis, age at diagnosis, location (colon or rectum), sex, and year of diagnosis.
- Survival following detection is a model input, and is based on relative survival conditional on stage at diagnosis, age at diagnosis and sex (as provided in Rutter, Johnson, Feuer et al. 2013).
- Future models that incorporate race will specify survival that also depends on race.

# Component Overview

## Summary

This page describes the components of the CRC–SPIN natural history model.

## Overview

There are four key components of the CRC–SPIN natural history model: 1) Adenoma risk; 2) Adenoma growth; 3) Size at transition to preclinical CRC; and 4) Time transition to clinical CRC and stage at diagnosis. Together, adenoma growth and size at transiting to preclinical CRC determine the time to transition to CRC. After transition to clinical CRC, the model assigns stage at diagnosis and survival given stage at diagnosis.

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

CRC–SPIN is used to simulate events for individuals (or `agents') in a population, sample, or cohort, by combining an age–sex distribution with a target population size. The age–sex distribution of simulated agents is a model input. The CRC–SPIN model can simulate population cohorts that have identical birthdays or have the same birth–year (or were born in a particular period). It is also possible to specify a sex–specific age distribution at a point in time, and then simulate these agents forward. The ability to specify more flexible age distributions is important for model calibration and validation.

## Component Listing

**Adenoma Risk:** CRC–SPIN simulates the occurrence of adenomas within agents using a non–homogenous Poisson process that allows adenoma risk to vary by age and sex (see Adenoma Risk Component for more details). The adenoma risk model is based on a Bayesian meta–analysis of 14 autopsy studies (Rutter, Miglioretti, Yu. 2007), which showed excellent fit to both the autopsy studies used for estimation, and to 4 screening colonoscopy studies used for validation.

Once adenomas are initiated, the CRC–SPIN model assigns their location using a multinomial distribution across 6 possible sites of the large intestine (from proximal to distal): 1) cecum; 2) ascending colon; 3) transverse colon; 4) descending colon; 5) sigmoid colon; 6) rectum. Overall location probabilities are not calibrated.

**Adenoma Growth:** The adenoma growth model is based on simulating the time it takes an adenoma to reach 10mm. This is then used in combination with a growth model to determine adenoma size at any point in time, which is needed to determine the outcomes of simulated tests. Adenoma size is also needed to determine the time at transition to preclinical CRC.

**Size at Transition to Preclinical CRC:** The model for transition to preclinical cancer is based loosely on autopsy studies of adenoma size and the presence of preclinical cancer. CRC–SPIN simulates the size at adenoma transition to preclinical invasive CRC using a lognormal model.

The **time from adenoma initiation to transition to preclinical cancer** is based on the combination of simulated adenoma growth and the simulated size at transition to preclinical cancer (see Transition To Preclinical CRC Component for more details).

**Time to transition From Preclinical to Clinical CRC:** CRC–SPIN V2.x uses a Weibull distribution for sojourn time. (CRC–SPIN 1.0 model used a log–Normal distribution.) The CRC–SPIN model does not include agent–level covariates in the sojourn–time model, though models that incorporate race will include an effect of race in the proportional hazards sojourn time model.

**Cancer stage and survival** are based on models that use Surveillance Epidemiology and End Results (**SEER**) data. In particular, we model the stage at clinical detection and then size conditional on stage. (The CRC–SPIN 1.0 model simulated cancer size, and then stage given size.) The **size during the preclinical detectable phase** is calculated assuming an exponential cancer growth model in combination with the size at transition to invasive cancer (0.5mm), the size at clinical detection, the time from initiation to clinical detection (sojourn time).

**Survival after CRC detection** is modeled as a function of age at diagnosis, sex, location (colon or rectum, stage, and year of diagnosis. Survival curves are based on analysis of SEER data. Models that include race will specify separate survival functions for black and white agents (in addition to effects of age at diagnosis, sex, location, stage, and year of diagnosis).

# Natural History Component

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

## Summary

This document describes the Natural History model and the specific components we use to model agents' progression from a disease free state to diagnosis.

## Overview

The CRC–SPIN natural history model simulates adenoma development, growth and transition to cancer. CRC–SPIN allows disease processes to depend on age, sex, and adenoma location (colon or rectum), but does not incorporate other risk factors.

## Details

There are four key components of the CRC–SPIN natural history model: 1) Adenoma risk; 2) Adenoma growth; 3) Size at transition to preclinical CRC; and 4) Time transition to clinical CRC and stage at diagnosis. Together, adenoma growth and size at transiting to preclinical CRC determine the time to transition to CRC. After transition to clinical CRC, the model assigns stage at diagnosis and survival given stage at diagnosis. These are described more fully in the Model Overview and in the separate model compents (Adenoma Risk Component, Transition To Preclinical CRC Component, Transition To Clinical CRC Component).

## Relevant Assumptions

The most basic model assumption is that all cancers arise from adenomas. In addition, at this time, we assume that risk for adenomas depends only on sex and age, and does not otherwise vary over time. CRC–SPIN includes stochastic variability in risk, but does not link risk across components. For example, at any given size, fast growing adenomas are no more likely to transition to cancer than slow–growing adenomas. For a complete listing of assumptions see Assumption Overview.

## Relevant Parameters

Parameters associated with the CRC–SPIN natural history model are described in the Parameter Overview

## Relevant Components

The separate components of the natural history model are described in the following pages:

- Adenoma Risk Component : A non–homogenous Poisson Process that allows risk to change with age and to depend on sex.
- Transition To Preclinical CRC Component : This model component is composed of two separate models, one describing adenoma growth and another describing the size at adenoma transition to preclinical cancer. CRC–SPIN simulates the time it takes for each adenoma to reach 10mm, and then simulates the size at any point in time using a Richards growth model that limits the maximum adenoma size to range from 1mm to 50mm. The size at transition to preclinical cancer is simulated using a lognormal model. Together, the growth model and the size at transition determine the time at adenoma transition to preclinical cancer.

- Transition To Clinical CRC Component : Sojourn time is modelled using a Weibull distribution that describes sojourn time and the variability of sojourn time across agents. The effect of location (colon or rectum) is incorporated through a proportional hazards model. (CRC–SPIN 1.0 used a log–normal distribution, with a two parameters for each location. CRC–SPIN 2.0 used a Weibull model with shape parameter set to 5 and one parameters for each location.)

Adenoma Stage and Survival, and Survival after CRC detection are uncalibrated model inputs (see Component Overview).

## Dependent Outputs

The CRC–SPIN natural history model is a 'parallel universe' model, that simulates complete life histories for all agents. These life histories include: age at adenoma initiation, transition(s) to preclinical cancer, age(s) at clinical cancer detection, age at colorectal cancer death, and age at non–CRC death. Transition times and CRC death ages are calculated both with and without screening. CRC–SPIN uses a shared uniform random deviate to link survival when a cancer is screen–detected rather than clinically–detected. .

# ADENOMA RISK COMPONENT

## SUMMARY

The occurrence of adenomas is simulated using a non–homogenous Poisson process that allows risk to depend on sex, and to increase with age.

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

## OVERVIEW

The CRC–SPIN adenoma risk model is based on a Bayesian meta–analysis of 14 autopsy studies (Rutter, Miglioretti, Yu, 2007). The meta–analytic model showed excellent fit to both the autopsy studies used for estimation, and to 4 screening colonoscopy studies used for validation.

## DETAIL

Let $\psi_i(t)$ denote the $i$th agent's instantaneous risk of an adenoma at time $t$. The risk of developing adenomas differs for men and women and increases with age. To allow flexibility, CRC–SPIN describes log–risk as a piecewise linear function of age. The risk of an adenoma developing in the $i$th agent at time $t$ is given by

$$\psi_i(t) = exp\left(\alpha_{0i} + \alpha_1 sex_i + \sum_{k=1}^{4} \delta(A_k < age_i(t) \leq A_{k+1}) \left\{ age_i(t)\alpha_{2k} + \sum_{j=2}^{k} A_j(\alpha_{2\ j-1} - \alpha_{2j}) \right\}\right)$$

where $\alpha_{0i}$ describes an agent's baseline risk; $\alpha_1$ describes the difference in risk for women ($sex_i = 0$) relative men ($sex_i = 1$); $\alpha_{2k}$ describes changes in risk with age (in years) in the $k$th interval; and $\delta()$ is an indicator function, with $\delta(x) = 1$ when $x$ is true and $\delta(x) = 0$ otherwise.

Given $\psi_i(t)$, the number of adenomas an agent develops by time $t$, has a Poisson distribution with mean
$\Psi_i(t) = \int_{20}^{age(t)} \psi_i(u)du$, given by

$$\Psi_i(t) = e^{\alpha_{0i} + \alpha_1 sex_i} \sum_{k=1}^{4} \left\{ \delta(age_i(t) > A_k) \left( \frac{e^{\alpha_{2k} min(A_{k+1}, age_i(t))} - e^{\alpha_{2k} A_k}}{\alpha_{2k}} \right) exp\left( \sum_{j=2}^{k} A_j(\alpha_{2\ j-1} - \alpha_{2j}) \right) \right\}$$

The baseline distribution of adenomas across the large intestine is based on combined information from 9 autopsy studies. These data were combined using a Bayesian Multinomial model with a Dirichlet prior for unknown probabilities. These baseline probabilities are: P(cecum)=0.08, P(ascending colon)=0.23, P(transverse colon)=0.24, P(descending colon)=0.12, P(sigmoid colon)=0.24, P(rectum) = 0.09, a distribution that is similar to the observed distribution in a relatively recent study of virtual and optical colonoscopy in a minimally screened population (Pickhardt et al. NEJM, 2003).

## RELEVANT ASSUMPTIONS

The risk of developing adenomas in childhood is extremely low. The CRC–SPIN model does not simulate the development of adenomas until age 20. The CRC–SPIN adenoma model specifies $K = 4$ fixed age–risk intervals: [20,50), [50,60), [60,70), and $\geq 70$, so that $A_1 = 20$, $A_2 = 50$, $A_3 = 60$, $A_4 = 70$ and $A_5 = \infty$ (effectively 120 years old). Risk increases log–linearly within these age intervals.

Agent–level baseline risk ($\alpha_{0i}$) results in clustering of adenomas within agents, so that high–risk agents develop more adenomas than low–risk agents. Agent–level baseline risk, $\alpha_{0i}$, is assumed to be independently and identically distributed Normal($\alpha_0$,$\sigma_\alpha$) across agents.

## Relevant Parameters

The CRC–SPIN adenoma risk model includes 7 parameters:

- $\alpha_0$, Expected baseline log–risk
- $\sigma_\alpha$, Standard deviation of baseline log–risk
- $\alpha_1$, The effect of sex on risk
- $\alpha_{2k}, k = 1, \ldots, 4$, The effect of age on risk.

## Relevant Components

The adenoma risk model starts the process that eventually leads to colorectal cancer. There are no subcomponents of this process. All subsequent adenoma processes (growth, transition to cancer) depend on the adenoma risk model.

## Dependent Outputs

The number of adenomas within each agent over time, when each was initiated, and their locations in the large intestine.

# TRANSITION TO PRECLINICAL CRC COMPONENT

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

## SUMMARY

Transition to preclinical CRC is modeled as a function of adenoma size. Thus, the CRC–SPIN transition model is based on on two sub–models, one for adenoma growth and another for cancer as a function of size.

## OVERVIEW

Adenoma growth is modeled using the Richards growth model, parameterized in terms of the median time to reach 10mm. Transition to clinical CRC is modeled as a function of adenoma size, with transition probabilities based loosely on autopsy studies of size and presence of invasive cancer.

## DETAIL

**Adenoma growth model**

Adenoma growth is simulated using the Richards growth model:

$$d_{ij}(t) = d_\infty \left[1 + \left(\left(\frac{d_0}{d_\infty}\right)^{1/p} - 1\right)\exp(-\lambda_{ij}t)\right]^p$$

where $d_\infty$ is the maximum possible adenoma diameter, $d_0$ is the minimal detectable adenoma diameter, and $\lambda_{ij}$ is the growth rate for the $_j$th adenoma within the $_i$th agent.

The Richards model is a general growth curve model that is primarily used in studies of animal growth. This model offers several advantages over other models of tumor growth. Unlike the Gompertz and logistic models, it allows relatively fast early growth with an asymptote at $d_\infty$. Adenoma size is assumed to range from a minimum of $d_0$ = 1mm to $d_\infty$ = 50mm. CRC–SPIN 1.0 used a Janoschek model, with $p = 1$. CRC–SPIN 2.x treats $p$ as a calibrated parameter.

Clinical information is not available for growth model parameters. Instead, there is expert opinion about the expected time it takes and adenoma to reach 10mm and information about the size of detected adenomas. To better incorporate this information, the CRC–SPIN model specifies adenoma growth in terms of the time, in years, that it takes for an adenoma to reach 10mm,

$$t_{10mm} = -\frac{1}{\lambda} ln\left(\frac{\left((10/d_\infty)^{1/p} - 1\right)}{\left((d_0/d_\infty)^{1/p} - 1\right)}\right)$$

$t_{10mm}$ is simulated using a Frèchet (or Type 2 Extreme Value) distribution with scale parameter $\beta_1$ and scale parameter $\beta_2$. The cumulative distribution function given by

$$F(t) = exp\left(-\left(\frac{t}{\beta_1}\right)^{-\beta_2}\right)$$

for $t \geq 0$. This is equivalent to using a type I extreme value distribution on $ln(t_{10mm})$. The Frèchet distribution has a long right tail but does not heavily weight small values that indicate fast growth.

Calibration of the CRC–SPIN 2.x model incorporated information about adenoma growth from a recent study that examined individuals with two screening colonoscopies that were approximately ten years apart (Ponugoti and Rex, 2017) and found that advanced adenomas were detected in only 3% of individuals at the second screening. Based on this, adenoma growth parameters were bounded so that the probability of an adenoma reaching 10mm within 10 years ranged from 0.001 to 0.25.

CRC–SPIN specifies separate growth distributions for colon and rectal adenomas, with parameters ($\beta_{1c}$, $\beta_{2c}$) and ($\beta_{1r}$, $\beta_{2r}$), respectively.

**Model for Size at Transition to Preclinical Cancer**

Information about adenoma transition to preclinical invasive disease comes from autopsy and colonoscopy studies of adenomas examining the rate of preclinical invasive disease by adenoma size. The CRC–SPIN adenoma transition model is loosely based on on autopsy study results of Nusko and colleagues (1997). This study included information about preclinical cancer rates in the colon and rectum from 11380 adenomas removed endoscopically or by surgical resection between January 1978 and December 1993. Other information comes from a study of follow–up colonoscopy that provides evidence that the probability of transition depends on the age of the individual at the time of adenoma initiation (Yamaji et al., 2006).

Adenomas in the rectum appear to transition to cancer earlier than adenomas located in the colon. This possibility is further supported by clinical cancer rates. Relatively few adenomas occur in the rectum (approximately 9%), yet nearly a third of clinically detected colorectal cancers are located in the rectum (based on 1975–1979 SEER data).

CRC–SPIN uses a log–normal model for the size at adenoma transition as a function of sex, location, and age at adenoma initiation, that is, the log–size at transition preclinical invasive CRC has a normal distribution. CRC–SPIN 1.0 assumed that standard deviation of log–size at transition was 0.5, with mean:

$$\begin{aligned}\mu_\gamma = \quad & \gamma_0 + \gamma_1\delta_f + \gamma_2\delta_r + \gamma_3\delta_f\delta_r + \\ & \Big(\gamma_4 + \gamma_5\delta_f + \gamma_6\delta_r + \gamma_7\delta_f\delta_r\Big)(\text{age at initiation} - 50).\end{aligned}$$

Where $\delta_f = 1$ if the agent is female and is zero if male, and $\delta_r = 1$ if the adenoma is located in the rectum and is zero if in the colon.

CRC–SPIN 2.x calibrates the standard deviation of the log–size at transition and assumes it has mean:

$$\begin{aligned}\mu_\gamma = \quad & \gamma_0 + \gamma_1\delta_f + \gamma_2\delta_r + \gamma_3\delta_f\delta_r + \\ & \gamma_4(\text{age at initiation} - 50) + \gamma_5(\text{age at initiation} - 50)^2\end{aligned}$$

## Relevant Assumptions

Key assumptions made by the CRC–SPIN adenoma transition model are:

- Adenomas do not regress, though they may grow very slowly;
- The minimum adenoma size (size at initiation) is 1mm and the maximum *adenoma* size is 50mm;
- The probability of transition to cancer is a function of adenoma size, adenoma location, and age at adenoma initiation.
- The Richards model adequately describes adenoma growth, the type 2 extreme value distribution adequately describes the variability in time to 10mm across agents, and the Lognormal model adequately describes the probability of transition as a function of size.

## Relevant Parameters

A full description of the parameters included in this component is provided in our Parameter Overview.

The CRC–SPIN adenoma transition model includes 11 parameters, 4 are associated with the adenoma growth and 7 are associated with the transition to invasive CRC.

*Adenoma Growth:*

- 4 parameters are associated with the Type 2 extreme value distribution used to model median time to 10mm: $\beta_{1\sigma}$, $\beta_{2\sigma}$, $\beta_{1r}$ and $\beta_{2r}$.

*Transition to Preclinical (Invasive) CRC:*

- 7 parameters are associated with the location–specific logistic regression models: $\gamma_0, \gamma_1, \ldots, \gamma_7$, and $\sigma_\gamma$, the standard deviation of the underlying standard deviation.

## Relevant Components

The adenoma transition component includes two subcomponents, one describing adenoma growth and the other describing the transition of adenomas to cancer as a function of size.

## Dependent Outputs

The growth model is used to simulate adenoma size at any point in time (size is used to determine the accuracy of some screening tests). The size at adenoma transition to preclinical invasive CRC is used to calculate the time/age at transition to preclinical cancer.

## Relevant Results

The key result from this component is the time from adenoma occurrence to transition to preclinical cancer. As noted above, adenoma size is also important because of its effect on screening accuracy.

# Transition To Clinical CRC Component

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

## Time to Clinical Cancer and Stage at detection

These components are described separately, below.

### Time to Clinical Cancer

The time from preclinical detectable cancer to clinical disease is known as **sojourn time**. For modeling purposes, all preclinical cancer is detectable, and so sojourn time begins at the time of transition to preclinical cancer and ends at transition to clinically detectable cancer.

- CRC–SPIN 1.0 used a log–normal distribution for sojourn time.
- CRC–SPIN 2.x simulates sojourn time using a Weibull proportional hazards model. Both the shape and scale of the Weibull distribution are calibrated. The proportional hazard regression incorporates differences in sojourn time for adenomas in the colon and rectum, and will be used to simulate differences in sojourn time by risk factors (e.g., race).

$$F(t) = [1 - exp(-(t/\mu_1)^{\mu_2})]^{exp(\mu_3 \mathrm{rectum}_i)}$$

Under this model, mean sojourn time is given by $[\mu_1 \Gamma(1 + 1/\mu_2)]^{exp(\mu_3 \mathrm{rectum}_i)}$.

### Size and Stage at Detection

Size at clinical detection is needed to simulate cancer size during the preclinical detectable period. Cancer size affects the sensitivity of screening tests, especially endoscopic screen detection. Size at detection is also related to stage at detection which is used to simulate survival.

- CRC–SPIN 1.0 simulated size at clinical detection and then stage at detection conditional on size.
- CRC–SPIN 2.x simulates stage at clinical detection, and then size at detection conditional on stage. Simulating stage at clinical detection directly allows greater flexibility in specification of the stage distribution. Information about size and stage at clinical detection is based on SEER data from 1975–1979 (*i.e.,* prior to diffusion of colorectal cancer screening).

### Survival

Our CRC–survival model is based on SEER data describing survival for cases diagnosed from 1975 to 2003. The CANSURV program (http://srab.cancer.gov/cansurv/) was used to estimate proportional hazard model that were stratified by location (colon or rectum) and AJCC stage with age and sex included as covariates. Models under development that incorporate race will specify that survival also depends on race using the same data (see: Rutter, Johnson, Feuer, et al., 2013).

Other–cause mortality was modeled using all–cause survival probabilities based on product–limit estimates for age and birth–year cohorts from the National Center for Health Statistics Databases (*US Life Tables, 2000*).

# PARAMETER OVERVIEW

## SUMMARY

This document describes calibrated CRC–SPIN parameters.

## BACKGROUND

Parameters are tied to observed data through calibrated using incremental mixture approximate Bayesian computation (IMABC; Rutter, Ozik, DeYoreo, Collier, under review).

Calibration uses targets based on unscreened or minimally screened samples and populations. Model validation more readily incorporates information from screened samples and populations.

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

## PARAMETER LISTING OVERVIEW

***Natural History Model Parameters***

*Adenoma Risk: 7 Parameters* (Adenoma Risk Component)
CRC–SPIN uses a non–homogeneous Poisson process to simulate adenoma occurrence

- Expected baseline log–risk: $\alpha_0$
- Standard deviation of baseline log–risk: $\sigma_\alpha$
- The effect of sex on risk: $\alpha_1$
- The effect of age on risk: $\alpha_{2k}$, $k = 1, \ldots, 4$. CRC–SPIN simulates change in risk for 4 age groups: $[20, 50)$, $[50, 60)$, $[60, 70)$, and $\geq 70$. Calibration results indicate that risk slows and may decline after age 70.

*Adenoma Growth: 4 Parameters* (Transition To Preclinical CRC Component)
CRC–SPIN simulates the time to reach 10mm using a Frèchet (Type 2 Extreme value) distribution for adenoma growth, assuming mutual independence for all parameters:

- $\beta_{1c}, \beta_{1r}$: shape parameters for adenomas in the colon and rectum, respectively
- $\beta_{2c}, \beta_{2r}$: scale parameters for adenomas in the colon and rectum, respectively

*Adenoma Size at Transition to Preclinical CRC: 7 estimated Parameters* (Transition To Preclinical CRC Component)

- Overall intercept, log–size at transition: $\gamma_0$
- Sex effect: $\gamma_1$
- Location effect (colon / rectum): $\gamma_2$
- Interaction between sex and location: $\gamma_3$
- (log) linear effect of age at initiation: $\gamma_4$
- (log) squared effect of age at initiation: $\gamma_5$
- standard deviation of log–size at transition: $\sigma_\gamma$

*Time to Clinical Cancer Component: 3 Parameters* (Transition To Clinical CRC Component)

- Weibull scale parameter: $\mu_1$
- Weibull shape parameter: $\mu_2$
- log–proportional hazards, sojourn time for rectal cancers: $\mu_3$

# Output Overview

## Summary

CRC–SPIN microsimulation model outputs.

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

## Overview

The CRC–SPIN model results in a person– and adenoma–objects that contain the life event histories for the entire simulated population. For each agent, this includes the timing of adenoma occurrence, the timing of transition to preclinical cancer, the timing of transition to clinical cancer, stage and size at clinical detection, survival after detection, and other–cause death date. Summary results are based on post–simulation processing of these life histories. Adenoma and preclincal cancer size can be calculated at any point in time because the adenoma object includes adenoma and cancer growth rates, we can calculate

## Output Listing

Reports are often generated using annual summaries, which are generally aggregated by location (proximal colon, distal colon, rectum), age, sex and year. These summaries include:

- adenoma prevalence
- the average number of adenomas within individuals
- preclinical cancer prevalence
- clinical cancer prevalence
- colorectal cancer mortality
- overall mortality

CRC–SPIN has great flexibility, in terms of the outputs simulated from natural history trajectories. For example, because CRC–SPIN is a 'parallel universe' approach (modelling outcomes for agents both with and without screening), it is possible to calculate the simulated disease–free years attributable to screening.

# RESULTS OVERVIEW

## SUMMARY

Here, we provide a very brief overview of our calibration model and model applications.

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

## OVERVIEW

The Bayesian calibration approaches used for the CRC–SPIN model results in simulated draws from the posterior distribution of model parameters given calibration targets. The CRC–SPIN 1.0 model used a likelihood–based approach that used an approximate Markov Chain Monte Carlo approach. CRC–SPIN 2.x models are calibratedl using Incremental Mixture Approximate Bayesian Calibration (IMABC), a likelihood–free approach. Bayesian calibration has several advantages over frequentist calibration methods, including the ability to simultaneous calibrate the model to multiple targets, incorporation of information via prior distributions, and the ability to simulate draws from the posterior distributions so that they can be used to inform parameter uncertainty and to propagate this uncertainty through the microsimulation model.

The CRC–SPIN 1.0 model has been used to estimate the comparative effectiveness of different screening regimens and has been validated through comparative modeling exercises within CISNET and through external validation to the UK Flexible Sigmoidoscopy study. The CRC–SPIN 2.x model updates this model and is used in publications after September 2018.

## RESULTS LIST

Model results can be found in publications, listed below.

- Zauber AG, Knudsen AB, Rutter CM, Lansdorp–Vogelaar I, Savarino JE, van Ballegooijen M, Kuntz KM. Cost–Effectiveness of CT Colonography to Screen for Colorectal Cancer: Report to the Agency for Healthcare Research and Quality from the Cancer Intervention and Surveillance Modeling Network (CISNET) for MISCAN, SimCRC, and CRC–SPIN Models. January 22, 2009. Available from: https://www.cms.gov/medicare-coverage-database/details/technology-assessments-details.aspx?TAId=58

- Rutter CM, Miglioretti DL, Savarino JE. Bayesian calibration of microsimulation models, Journal of the American Statistical Association, 2009; 104(488):1338–1350. PMCID: PMC2805837.

- Rutter CM, Savarino JE. An evidence–based microsimulation model for colorectal cancer, Cancer Epidemiology Biomarkers and Prevention, 2010; 19(8):1992–2002. PMCID: PMC2919657.

- Knudsen AB, Lansdorp–Vogelaar I, Rutter CM, Savarino JE, van Ballegooijen M, Kuntz KM, Zauber AG. Cost–Effectiveness of CT Colonography Screening for Colorectal Cancer among the Medicare Population, Journal of the National Cancer Institute, 2010; 102:1238–1252. PMCID: PMC2923219.

- Berrington de González A, Kim KP, Knudsen AB, Lansdorp–Vogelaar I, Rutter CM, Smith–Bindman R, Yee J, Kuntz KM, van Ballegooijen M, Zauber AG, Berg CD. Radiation–related cancer risks from CT colonography screening: a risk–benefit analysis, accepted for publication, American Journal of Roentgenology, 2010; 196:816–823. PMCID: PMC3470483.

- Kuntz KM, Lansdorp–Vogelaar I, Rutter CM, Knudsen AB, van Ballegooijen M, Savarino J, Feuer EJ, Zauber AG. A systematic analytical comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression, Medical Decision Making, 2011; 31:530–539. PMCID: PMC3424513.

- van Ballegooijen M, Rutter CM, Knudsen AB, Zauber AG, Savarino J, Lansdorp–Vogelaar I, Feuer EJ, Kuntz KM. Clarifying differences between models for screening. The case of colorectal cancer, Medical Decision Making, 2011; 31:540–549. PMCID: PMC3531980.

- Vanness DJ, Knudsen AB, Lansdorp–Vogelarr I, Rutter CM, Gareen IF, Herman BA, Kuntz KM, Zauber, AG, van Ballegooijen M, Feuer EJ, Chen M, Johnson CD. Comparative Economic Evaluation of Data from the ACRIN National CT Colonography Trial with Three Cancer Intervention and Surveillance Modeling Network Microsimulations, Radiology, 2011; 261:487–498. PMCID: PMC3198218.

- Rutter CM, Miglioretti DL, Savarino JE. Evaluating risk factor assumptions: a simulation–based approach. BMC Medical Informatics and Decision Making, 2011; 11:55. PMCID: PMC3182875.

- Lansdorp–Vogelaar, Gulati R, Mariotto AB, Schechter CB, Heijnsdijk EA, Knudsen AB, van Ravesteijn NT, Wever EM, van Ballegooijen M, Rutter CM, Kuntz KM, Feuer EJ, Etzioni R, de Koning HJ, Zauber* AG, Mandelblatt* JS. Personalizing Age of Screening Cessation Based on Comorbidity – Results of Collaborative Modeling of Breast, Colorectal, and Prostate Cancer, Annals of Internal Medicine, 2014; 162(2):104–12. PMCID: PMC4160041

- Zauber A, Knudsen A, Rutter CM, Lansdorp–Vogelaar I, Kuntz KM. Evaluating the Benefits and Harms of Colorectal Cancer Screening Strategies: A Collaborative Modeling Approach. Technical report for the Agency for Healthcare Research and Quality. AHRQ Publication No. 14–05203–EF–2, October 2015.
http://www.uspreventiveservicestaskforce.org/Page/Document/modeling-report/colorectal-cancer-screening2

- Rutter CM, Lansdorp–Vogelaar I, Knudsen AKB, Marsh T, Kuntz KM, van Ballegooijeen M, Zauber A. Validation of Models used to Inform Colorectal Cancer Screening Guidelines: Accuracy and Implications, Medical Decision Making, 2016; 36:604–614. PMCID: PMC5009464.

- Knudsen AB, Zauber AG, Rutter CM, Naber SK, Doria–Rose VP, Pabiniak C, Johanson C, Fischer SE, Lansdorp–Vogelaar I, Kuntz KM. Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the US Preventive Services Task Force. JAMA. 2016 Jun 21; 315(23):2595–609. PMCID: PMC5493310.

- Rutter CM, Kim JJ, Meester RGS, Sprague BL, Burger EA, Zauber AG, Ergun MA, Campos NG, Doubeni CA, Trentham–Dietz A, Sy S, Alagoz O, Stout NK, Lansdorp–Vogelaar I, Corley DA, Tosteson ANA. Effect of Time to Diagnostic Testing for Breast, Cervical, and Colorectal Cancer Screening Abnormalities on Screening Efficacy: A Modeling Study, Cancer Epidemiol Biomarkers Prev, 2018. PMCID: PMC5809257 [Available on 2019–02–01]

# KEY REFERENCES

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

## BACKGROUND

**Jemal A, Tiwari RC, Murray T, Ghafoor A, Samuels A, Ward E, Feuer EJ, Thun MJ.** Cancer statistics, 2004. CA Cancer J Clin 2004 Jan–2004 Feb 28;54(1):8–29.

**Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, James PD, Mangham CM.** Randomised controlled trial of faecal–occult–blood screening for colorectal cancer. Lancet 1996 Nov 30;348(9040):1472–7.

**Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O.** Randomised study of screening for colorectal cancer with faecal–occult–blood test. Lancet 1996 Nov 30;348(9040):1467–71.

**Towler B, Irwig L, Glasziou P, Kewenter J, Weller D, Silagy C.** A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, hemoccult BMJ 1998;317:559–65.

**Winawer SJ.** A quarter century of colorectal cancer screening: progress and prospects. J Clin Oncol 2001 Sep 15;19(18 Suppl):6S–12S.

**Selby JV, Friedman GD, Quesenberry CP Jr, Weiss NS.** A case–control study of screening sigmoidoscopy and mortality from colorectal cancer. N Engl J Med 1992 Mar 5;326(10):653–7.

**Newcomb PA, Norfleet RG, Storer BE, Surawicz TS, Marcus PM.** Screening sigmoidoscopy and colorectal cancer mortality. J Natl Cancer Inst 1992 Oct 21;84(20):1572–5.

**Winawer SJ, Stewart ET, Zauber AG, Bond JH, Ansel H, Waye JD, Hall D, Hamlin JA, Schapiro M, O'Brien MJ, et al.** A comparison of colonoscopy and double–contrast barium enema for surveillance after polypectomy. National Polyp Study Work Group. N Engl J Med 2000 Jun 15;342(24):1766–72.

## ADENOMA RISK: NON–HOMOGENEOUS POISSON MODEL

**Cox DR, Miller HD.** (1965) The theory of stochastic processes, New York: Chapman and Hall.

## ADENOMA RISK: AUTOPSY STUDIES CONTRIBUTING TO META–ANALYSIS

**Clark JC, Collan Y, Eide TJ, Esteve J, Ewen S, Gibbs NM, Jensen OM, Koskela E. MacLennan R, Simpson JG, Stalsberg H, Zaridze DG.** (1985) Prevalence of polyps in an autopsy series from areas with varying incidence of large–bowel cancer. Int J Cancer, 36, 179–186.

**Blatt LJ.** (1961) Polyps of the colon and rectum: Incidence and distribution. Diseases of the Colon & Rectum, 4, 277–282.

**Chapman I.** (1963) Adenomatous polypi of large intestine: Incidence and distribution. Annals of Surgery, 157, 223–226.

**Hughes LE.** (1968) The incidence of benign and malignant neoplasms of the colon and rectum: A post–mortem study. Australian & New Zealand Journal of Surgery, 38, 30–35.

**Stemmermann GN, Yatani R.** (1973) Diverticulosis and polyps of the large intestine. A necropsy study of Hawaii Japanese. Cancer, 31, 1260–1270.

**Eide TJ, Stalsberg H.** (1978) Polyps of the large intestine in Northern Norway. Cancer, 42, 2839—2848.

**Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM.** (1979) Adenomatous lesions of the large bowel. An autopsy survey. Cancer, 43,1847–1857.

**Vatn MH, Stalsberg H.** (1982) The prevalence of polyps of the large intestine in Oslo: An autopsy study. Cancer, 49, 819–825.

**Williams AR, Balasooriya BAW, Day DW.** (1982) Polyps and cancer of the large bowel: A necropsy study in Liverpool. Gut, 23, 835–842.

**Bombi JA.** (1988) Polyps of the colon in Barcelona, Spain. Cancer, 61, 1472–1476.

**Johannsen LGK, Momsen O, Jacobsen NO.** (1989) Polyps of the large intestine in Aarhus, Demark. An autopsy study. Scandinavian Journal of Gastroenterology, 24, 799–806.

**Jass JR, Young PJ, Robinson EM.** (1992) Predictors of presence, multiplicity, size and dysplasia of colorectal adenomas. A necropsy study in New Zealand. Gut, 33, 1508–1514.

**Szczepanski W, Urban A, Wierzchowski W.** (1992) Colorectal polyps in autopsy material. Part I. Adenomatous polyps. Pat Pol, 43, 79–85.

**Paspatis GA, Papanikolaou N, Zois, E, Michalodimitrakis E.** (2001) Prevalence of polyps and diverticulosis of the large bowel in the Cretan population. An autopsy study. Int J Colorectal Dis, 16, 257–261.

**Rutter CM, Miglioretti DL, Yu O.** (2007) Adenoma risk meta analysis. Statistics in Medicine, 26: 98–109.

## Adenoma Risk: Colonoscopy Studies Contributing to Validation

**DiSario JA, Foutch PG, Mai HD, Pardy K, Manne RK.** (1991) Prevalence and malignant potential of colorectal polyps in asymptomatic, average–risk men. American Journal of Gastroenterology, 86, 941–945.

**Rex DK, Lehman GA, Ulbright TM, Smith JJ, Pound DC, Hawes RH, Helper DJ, Wiersema MJ, Langefeld CD, Li W.** (1993) Colonic neoplasia in asymptomatic persons with negative fecal occult blood tests: Influence of age, gender, and family history. American Journal of Gastroenterology, 88, 825–831.

**Lieberman DA, Weiss DG, Bond JH, Anhen DJ, Garewal H, Chejfec G.** (2000) Use of colonoscopy to screen asymptomatic adults for colorectal cancer. New England Journal of Medicine, 343, 162–168.

**Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, Wong RK, Nugent PA, Mysliwiec PA, Schindler WR.** (2003) Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. New England Journal of Medicine, 349, 2191–2200.

**Ponugoti PL, Rex DK.** (2017). Yield of a second screening colonoscopy 10 years after an initial negative examination in average–risk individuals. Gastrointestinal endoscopy, 85(1), 221–224.

## Adenoma Growth Model

**Tjørve E, Tjørve KM.** A unified approach to the Richards–model family for use in growth analyses: why we need only two model forms. Journal of Theoretical Biology. 2010 Dec 7;267(3):417–25.

## Transition from Adenoma to Preclinical Cancer as a Function of Size

**Nusko G, Mansmann U, Altendorf–Hofmann A, Groitl H, Wittekind C, Hahn EG.** Risk of invasive carcinoma in colorectal adenomas assessed by size and site. Int J Colorectal Dis 1997;12(5):267–71.

**Gillespie PE, Chambers TJ, Chan KW, Doronzo F, Morson BC, Williams CB.** Colonic adenomas—a colonoscopy survey. Gut 1979 Mar;20(3):240–5.

**Shinya H, Wolff WI.** Morphology, anatomic distribution and cancer potential of colonic polyps. Ann Surg 1979 Dec;190(6):679–83.

## CRC Survival

**Rutter CM, Johnson E, Feuer R, Knudsen AB, Kuntz KM, Shrag D.** Secular Trends in Colon and Rectal Cancer Survival, *JNCI,* 2013; 105(23): 1806–13.

## Bayesian Calibration

**Gelman A, Stern HS, Carlin JB, Dunson DB, Vehtari A, Rubin DB.** Bayesian data analysis. 2013, Chapman and Hall/CRC.

**Gilks WR, Richardson S, Spiegelhalter DJ. Eds** Markov Chain Monte Carlo in Practice. London, UK: Chapman & Hall; 1996; c1996.

**Rutter CM, Miglioretti DL, Savarino JE** (2009) Bayesian calibration of microsimulation models. JASA, 104: 1338–1350.

**Rutter CM, Ozik J, DeYoreo M, Collier N.** Microsimulation Model Calibration using Incremental Mixture Approximate Bayesian Computation. arXiv preprint arXiv:1804.02090. 2018. http://arxiv.org/abs/1804.02090.

## Calibration Data

**Surveillance, Epidemiology, and End Results (SEER) Program** (www.seer.cancer.gov), SEER*Stat Database: Incidence – SEER 9 Regs Public–Use, Nov 2003 Sub (1973–2001), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2004, based on the November 2003 submission.

**Strul H, Kariv R, Leshno M, Halak A, Jakubowicz M, Santo M, Umansky M, Shirin H, Degani Y, Revivo M, Halpern Z, Arber N.** The Prevalence Rate and Anatomic Location of Colorectal Adenoma and Cancer Detected by Colonoscopy in Average–Risk Individuals Aged 40–80 Years, American Journal of Gastroenterology 2006: 101: 255–262.

**Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G.** Use of Colonoscopy to Screen Asymptomatic Adults for Colorectal Cancer, NEJM 2000; 343:162–8.

**Pickhardt PJ, Choi R, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, Wong RK, Nugent PA, Mysliwiec PA, Schindler WR.** Computed Tomographic Virtual Colonocopy to Screen for Colorectal Neoplasia in Asymptomatic Adults, NEJM 2003; 349:2191–200.

**Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF.** Risk of Advanced Proximal Neoplasms in Asymptomatic Adults According to the Distal Colorectal Findings, NEJM 2000: 343:169–74.

**Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF.** Results of Screening Colonoscopy Among persons 40 to 49 Years of Age, NEJM 2002; 346:1781–5.

**Church JM.** Clinical Significance of Small Colorectal Polyps, Dis Colon Rectum 2004; 47:481–485.

**Odom SR, Duffy SD, Barone JE, Ghevariya V, McClane SJ.** The Rate of Adenocarcinoma in Endoscopically Removed Colorectal Polyps, The American Surgeon 2005; 71:1024—1026