



FLEXKB DOCUMENT
Version: HI.001.02132018.65296
Document generated: 02/13/2018



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

FRED HUTCHINSON CANCER RESEARCH CENTER

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.

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.

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.



Fred Hutchinson CRC
Model Purpose

MODEL PURPOSE

SUMMARY

This document provides an overview of the Fred Hutchinson Cancer Research Center (FHCRC) multistage clonal expansion for esophageal adenocarcinoma (MSCE-EAC) model with multiple scales, including the cell, crypt, clonal patch, tissue [normal, Barrett's esophagus (BE), high grade dysplasia (HGD), and esophageal adenocarcinoma (EAC)], individual, and population levels. The model combines an age-dependent gastroesophageal reflux disease (GERD) component with multistage cell kinetic rates that depend on birth cohort to fit US EAC incidence data. Both likelihood-based and detailed multiscale spatial simulation methods are used for analysis and prediction of EAC trends and effects of alternative screening and treatment protocols.

PURPOSE

The purpose of the MSCE-EAC model is to serve as an effective tool for evaluating EAC trends in the US population and the impact of possible interventions on modifying future cancer trends. The model combines rigorous likelihood-based estimation of cell kinetic rates that drive the cancer process with detailed spatial simulation of the growth and extinction of premalignant and malignant clones to evaluate the sensitivities of different biopsy and advanced endoscopic imaging protocols and the potential benefits and harms of radio-frequency ablation or other treatment methods.

FRED HUTCHINSON
CANCER RESEARCH CENTER
A LIFE OF SCIENCE

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



MODEL OVERVIEW

SUMMARY

The MSCE–EAC model provides a mathematical and computational framework for multiscale modeling of the natural history of progression from normal esophageal squamous epithelium to esophageal adenocarcinoma (EAC), and the impact of alternative protocols for biopsy, imaging, and treatment.

PURPOSE

The purpose of the MSCE–EAC model is to provide insight into the biology and natural history of progression and detection of EAC over many length and time scales, beginning with models of fundamental processes represented at the cellular level.

The development of BE is recognized as an early step in progression to EAC, with an enhanced risk for BE among individuals with gastroesophageal reflux disease (GERD) symptoms. The model represents age–dependent development of weekly or more frequent GERD symptoms, with transitions from both GERD and non–GERD pathways to develop BE, two additional mutations or epigenetic changes for the initiation of HGD, with clonal expansion of cells comprising HGD, malignant transformation, and a more rapid clonal expansion process for EAC.

GERD incidence data were utilized to calibrate the model for age–dependent GERD prevalence, and Surveillance Epidemiology and End Results (SEER) incidence data were used for likelihood–based calibration of the remaining parameters of the multiscale EAC progression model.

EAC incidence has increased approximately six–fold in the US since 1975, as reflected in SEER data.¹ These temporal trends were modeled by systematically applying flexible period and cohort trends to the biological parameters of the MSEAC model, and using likelihood methods for model comparison and selection of the best model fit to SEER incidence.

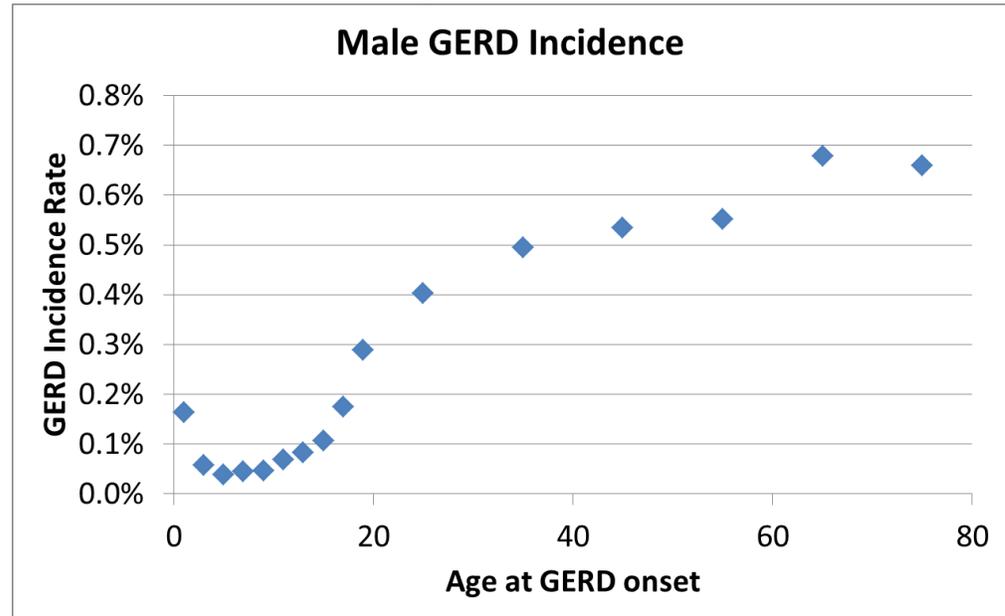
To identify which biological parameters may be influenced by temporal trends, we compared alternative models with period and/or cohort effects influencing GERD development, the transition rate to BE, early mutation steps, growth of premalignant lesions, malignant transformation, and clonal growth of the tumor. The best model fit includes a sigmoidal (birth) cohort trend on both premalignant and malignant clonal expansion (see [Results Overview](#)).

Spatial simulations of the growth of premalignant clones (identified with HGD) and malignant tumors are mapped to represent two–dimensional localization and growth on the BE segment of the esophageal surface (represented as a torus).

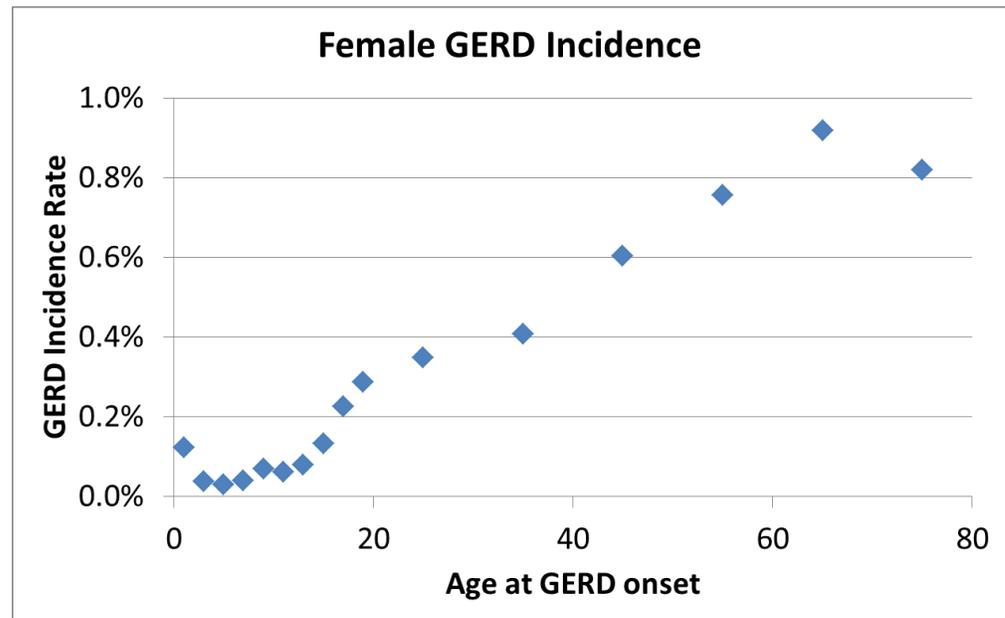
This spatial modeling component of the MSEAC model allows analysis of the probability for biopsy sampling of HGD and preclinical EAC during screening, along with symptomatic cancer detection. This framework is inherently 'multiscale' in that it bridges the cellular scale with the population scale, allowing us to model physically the process of endoscopic screening of BE patients for the presence of premalignant and preclinical malignant lesions prior to the appearance of cancer symptoms and/or a cancer diagnosis.

BACKGROUND

A clinically important component of the MSCE–EAC model is an underlying gender and age–specific model of GERD prevalence, which is generally believed to increase the relative risk for BE. Calibration of the GERD prevalence model utilized data from incident GERD cases in a cohort of 1700 children and adolescents in the Health Improvement Network (THIN) UK primary care database between 2000–2005,² and case–control data on adults with a first diagnosis of GERD in the UK General Practice Research Database (GPRD), including 7451 cases and 10,000 controls.³



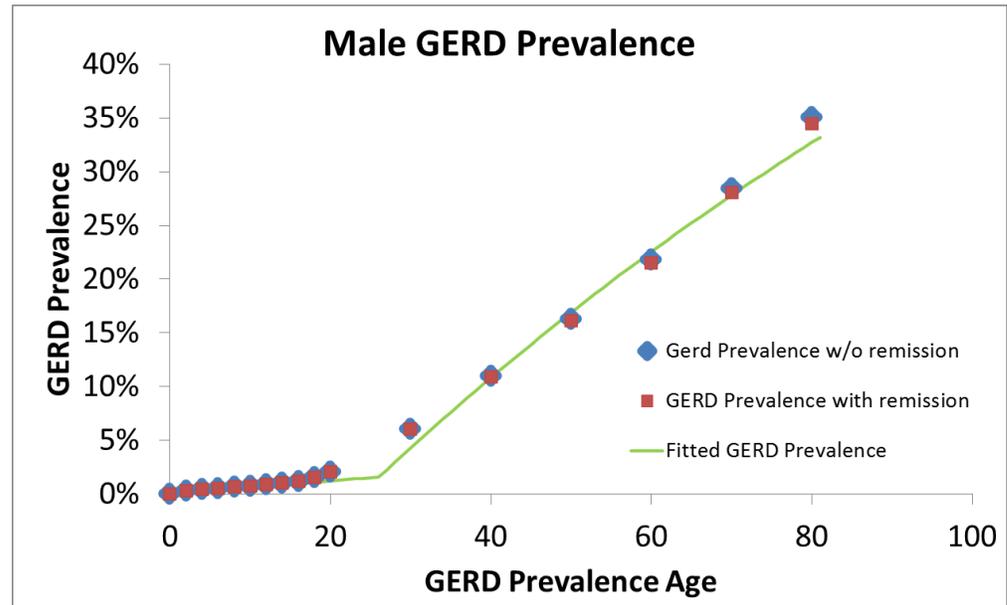
Data from Ruigomez, et al. (1,2)



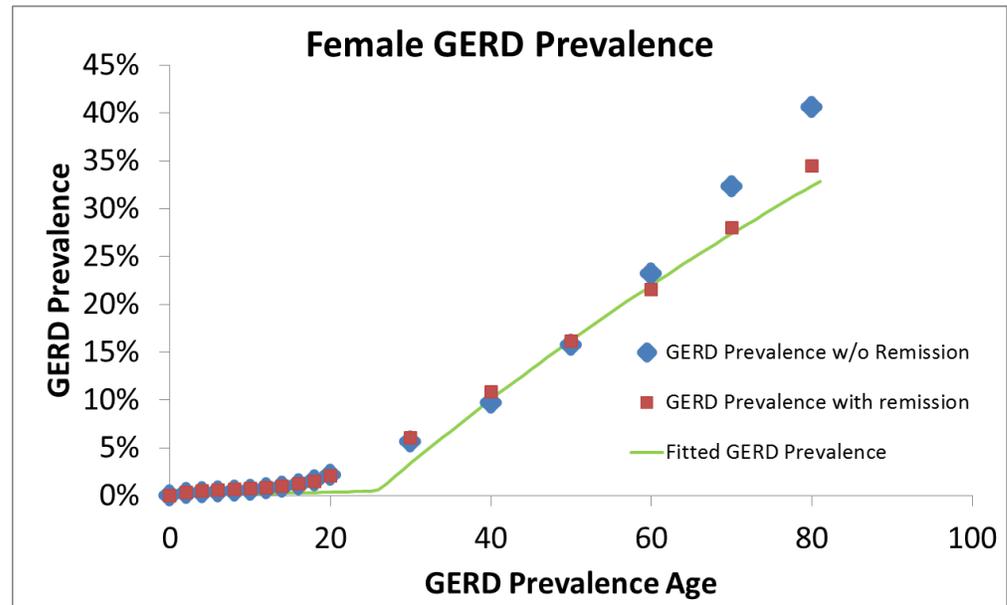
Data from Ruigomez, et al. (1,2)

GERD was defined as heartburn and/or regurgitation experienced at least weekly in these studies.^{2,3} Using this definition, we develop male and female GERD models with GERD prevalence increasing in accordance with the data for age–specific GERD

incidence.^{2,3} However, the models also include a parameter representing reversion rates of GERD symptoms, allowing us to fit age-adjusted GERD prevalence between ages 40 to 85 to an approximate target of 20%, consistent with population-based studies of GERD prevalence.^{4,5,6,7} We then use maximum likelihood methods to fit the data-driven models above and generate simpler 3-parameter gender and age-specific GERD prevalence models that represents an effective childhood/young adult transition rate to GERD, a transition age, and an effective older adult transition rate to GERD. (See green lines in Figures below).



Male GERD prevalence model



Female GERD prevalence model

Epidemiological studies indicate that most individuals with GERD do not develop BE, but that GERD is a significant risk factor for BE but with differing estimates of relative risk (RR) for BE given GERD ranging between 2–15%, and also depending on BE segment length, frequency of GERD symptoms and other factors.^{8,9} A recent

meta-analysis of 14 studies found an odds-ratio of at-least weekly GERD in relation to long segment BE of 4.92, CI=(2.01–12.0), and no association with short segment BE.⁹ Another recent meta-analysis of 5 studies on the association of GERD with EAC found an odds ratio of 4.92, CI= (3.92, 6.22).¹⁰

The model fits to SEER data allow prediction of the background transition rate to BE, and thus BE prevalence is predicted by including GERD and non-GERD pathways, with predictions of 1.5–5% BE prevalence for males for ages ranging between 40 and 85, and between 0.5–1% for females. Population estimates of BE prevalence differ widely,¹¹ but given this uncertainty, the model predictions appear generally consistent with the range of estimates in the studies.

SEER EAC incidence has increased roughly six-fold since 1975.^{1,12} To identify biological parameters that may be influenced by temporal trends, we compared alternative models with period and/or cohort effects influencing GERD development, the transition rate to BE, early mutation steps, growth of premalignant lesions, malignant transformation, and clonal growth of the tumor.

MODEL DESCRIPTION

The development of BE is recognized as an early step in progression to EAC, with an enhanced risk for BE among individuals with gastroesophageal reflux disease (GERD) symptoms. The model represents age-dependent development of weekly or more frequent GERD symptoms, with transitions from both GERD and non-GERD pathways to develop BE, two additional mutations or epigenetic changes for the initiation of HGD, with clonal expansion of cells comprising HGD, malignant transformation, and a more rapid clonal expansion process for EAC. The transition rate ν from normal to BE includes a baseline rate ν_0 for individuals without GERD and a faster rate for individuals with GERD modeled as $\nu_0 * RR$, where RR is the relative risk for BE given GERD. (Calibration of models for GERD and BE prevalences is discussed below).

Age-dependent model of prevalence for GERD and Barrett's esophagus

Let p_G be the probability of GERD at age t , with $p_G(t) = 1 - \exp[-r_1 \min(r_3, t) - r_2 \max(0, t - r_3)]$ being a three-parameter function that we fit to GERD incidence data^{2,3} and age-adjusted GERD prevalence,^{4,5,6,7} as described in the Background section.

We model the age-dependent exponential transition rate $\nu(t)$ for conversion from normal tissue to BE as

$$\nu(t) = \nu_0 [(1 - p_G(t)) - RR \cdot p_G(t)]$$

where RR is the relative risk for GERD given BE.

The density for BE onset times, $f_{BE}(t)$ and cumulative function for BE prevalence $F_{BE}(t)$ are

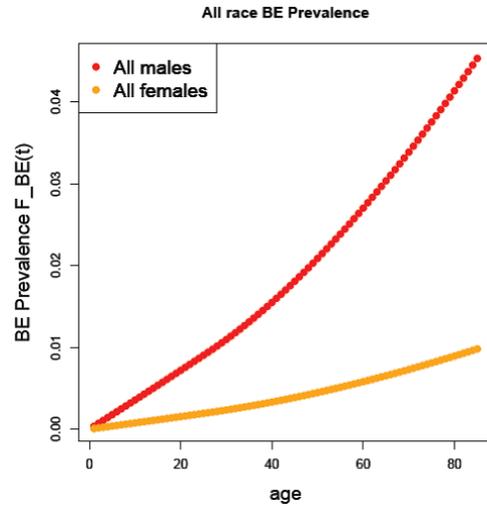
$$f_{BE}(t) = \nu(t) e^{-\int_0^t \nu(u) du},$$

and

$$F_{BE}(t) = 1 - e^{-\int_0^t \nu(u) du}, \text{ respectively.}$$

As described in the Background section, different studies differ in their estimates of

relative risk (RR) for BE given GERD.^{8,9,10} We use a consensus estimate of RR from these studies in assuming a model relative risk of RR=5 for BE given GERD.

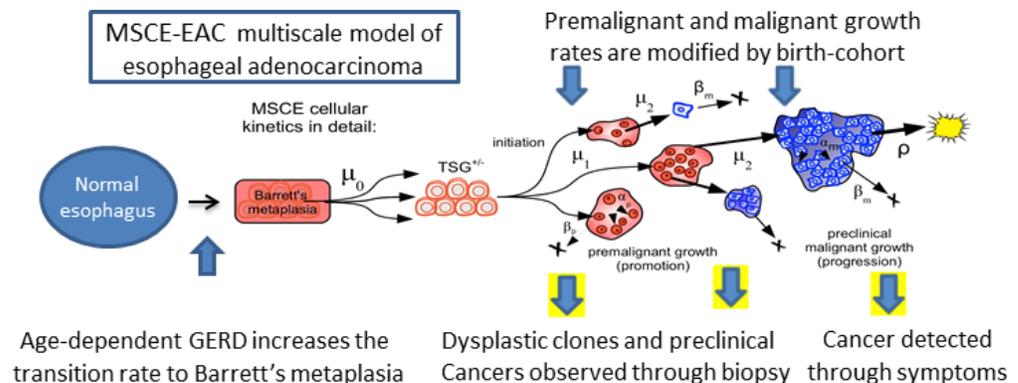


Model BE prevalence based on RR=5

We note that the primary MSCE–EAC model outcome is EAC and the model assumes BE is necessary for EAC. Using the estimate of RR=5, we used maximum likelihood methods to calibrate the MSCE–EAC model to EAC incidence data from SEER to estimate the remaining model parameters.

BE is modeled as a metaplastic tissue with random segment length drawn from a beta distribution $B(16/11, 4)$ ¹³ containing on average 106 BE stem cells. These BE stem cells may undergo mutation or epigenetic modification, with two successive hits occurring during asymmetric cell division (at rates μ_0, μ_1) that are required to inactivate a gatekeeper or tumor suppressor gene (TSG) and generate a premalignant daughter cell with partial loss of tissue homeostasis. (When calibrating to EAC incidence, μ_0 and μ_1 are not separately identifiable, so without loss of generality we set $\mu_0 = \mu_1$). Premalignant cells, which we associate with high grade dysplasia (HGD) may divide (with rate α_p); die, undergo apoptosis or differentiate (at rate β_p); or mutate during asymmetric cell division (at rate μ_2) to generate a malignant cell. Similarly, malignant cells may divide (with rate α_m); undergo apoptosis or differentiate (at rate β_m); or undergo detection through a size–based stochastic observation process based on a per–cell detection rate ρ .

The difference between the birth and death rates is called the net cell proliferation rate (γ) for each cell type. Model parameters are calibrated through maximum likelihood fits to EAC incidence data from nine registries of the Surveillance and End Results (SEER) database between 1975 and 2010.¹



EAC incidence has increased approximately six-fold in the US since 1975.^{1,12} These temporal trends were modeled by systematically applying flexible period and cohort trends to the biological parameters of the MSCE-EAC model, and using likelihood methods for model comparison and selection of the best model fit to SEER incidence. The best fitting model includes a sigmoidal birth cohort effect modifying the growth rates of premalignant and malignant cells, with rates for malignant growth significantly larger than for premalignant growth. The sigmoidal shape for premalignant growth rate is parametrized as shown in the following section.

Growth of premalignant (P–cells) modified by sigmoidal birth–cohort effect

Let b_k represent the birth cohort, indexed by year k .

NOTE: to keep notation simple in the following, we do not add the index k to the division rates α_P or α_M , net cell proliferation rates γ_P or γ_M , or death rates β_P or β_M (where P and M represent premalignant and malignant cells, respectively).

$$\alpha_P = \alpha_{P0} \left(g_1 + \frac{2}{1 + \exp(-g_2(b_k - t_{ref}))} \right) \quad \text{'P cell division'}$$

$$\gamma_P = \gamma_{P0} \left(g_1 + \frac{2}{1 + \exp(-g_2(b_k - t_{ref}))} \right) \quad \text{'Net P-cell proliferation rate'}$$

$$\beta_P = \alpha_P - \gamma_P - \mu_2 \quad \text{'P-cell death (apoptosis) or differentiation rate'}$$

Growth of malignant M–cells modified by sigmoidal birth–cohort effect

$$\alpha_M = \alpha_{M0} \left(g_1 + \frac{2}{1 + \exp(-g_2(b_k - t_{ref}))} \right) \quad \text{'M-cell division rate'}$$

$$\gamma_M = \gamma_{M0} \left(g_1 + \frac{2}{1 + \exp(-g_2(b_k - t_{ref}))} \right) \quad \text{'Net M-cell proliferation rate'}$$

$$\beta_M = \alpha_M - \gamma_M - \rho \quad \text{'M-cell death (apoptosis) or differentiation rate'}$$

The analytic form of the sigmoidal function allows smooth estimation of future trends, with projections until 2030 for males and females shown in the figures below. The figures show that the six-fold increase in incidence can be explained by smaller changes in the net cell proliferation rates of premalignant and malignant clones that increase less than three-fold across birth cohorts spanning a century.

MSCE-EAC Model Differential Equations

$$\frac{dy_1}{dt} = \beta_M - (\alpha_M + \beta_M + \rho) y_1 + \alpha_M y_1^2$$

$$\frac{dy_2}{dt} = 2 \alpha_M y_1 y_2 - (\alpha_M + \beta_M + \rho) y_2$$

$$\frac{dy_3}{dt} = \beta_P + \mu_2 y_1 y_3 - (\alpha_P + \beta_P + \mu_2) y_3 + \alpha_P y_3^2$$

$$\frac{dy_4}{dt} = 2 \alpha_P y_3 y_4 + \mu_2 (y_4 y_1 + y_3 y_2) - (\alpha_P + \beta_P + \mu_2) y_4$$

$$\frac{dy_5}{dt} = \mu_1 y_5 (y_3 - 1)$$

$$\frac{dy_6}{dt} = \mu_1 (y_6 y_3 - y_6 + y_5 y_4)$$

$$\frac{dy_7}{dt} = X \mu_0 y_7 (y_5 - 1)$$



$$\frac{dy_8}{dt} = X \mu_0 (y_8 y_5 - y_8 + y_7 y_6)$$

$$\frac{dy_9}{dt} = \nu y_9 (y_7 - 1) \quad \text{'Survival = } S(t) = y_9 \text{'}$$

$$\frac{dy_{10}}{dt} = -\nu y_8 \quad \text{'Hazard = } h(t) = y_{10} \text{'}$$

MSCE-EAC Model Likelihood

Maximum likelihood methods were used to fit to EAC incidence data from SEER for ages 1 to 84 and calendar years 1975–2010.

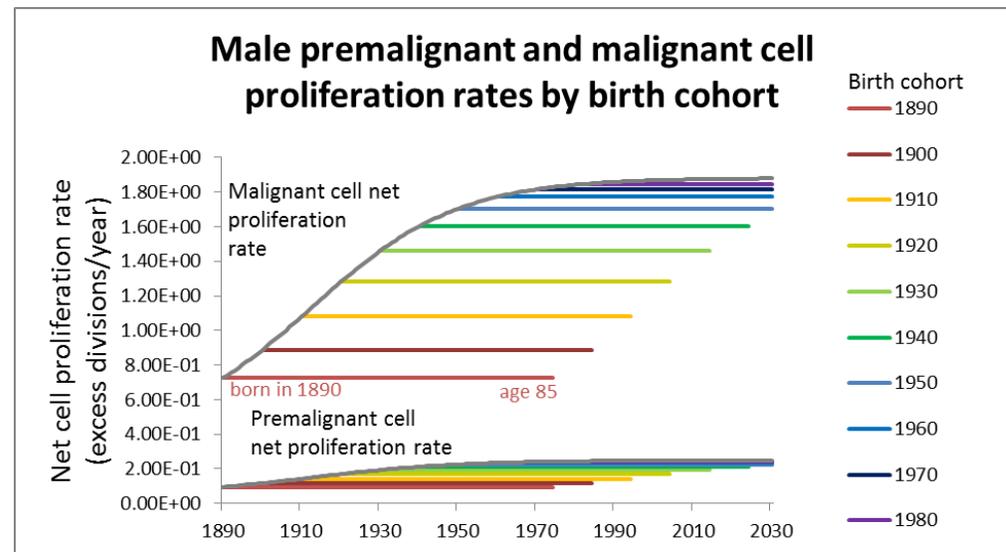
The expected number of EAC cancers at age a_i and period (calendar year) p_j and birth cohort $b_k = c_j - a_i$ is

$\Lambda_{i,j} = PY_{i,j} h(a_i, b_k)$, where $PY_{i,j}$ is the number of person years of age a_i and period p_j , and the birth-cohort specific hazard is $h(a_i, b_k) = h(t)|\{t = a_i, b_k\}$.

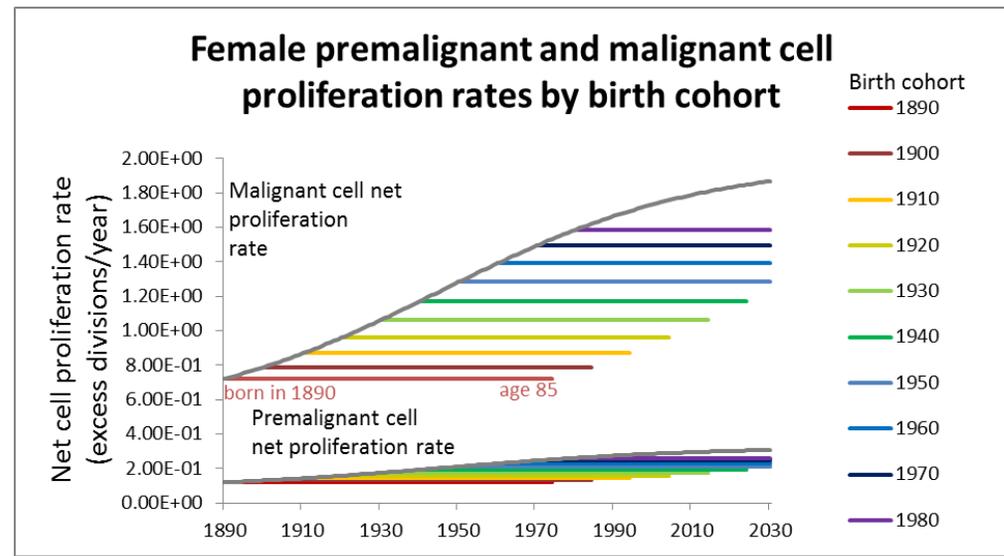
The likelihood is

$$\mathcal{L} = \prod_{i,j} \frac{\Lambda_{i,j}^{O_{i,j}} \exp[-\Lambda_{i,j}]}{O_{i,j}!}, \text{ where } O_{i,j} \text{ is the number of observed EAC cases for age } a_i \text{ and period } p_j.$$

The best model fit includes a sigmoidal (birth) cohort trend on both premalignant and malignant clonal expansion, with rates for malignant growth significantly larger than for premalignant growth.



Of particular note is the observation that the estimated increases in proliferation rates among males are leveling off for recent birth cohorts but continue rising for recent females birth cohorts.



CONTRIBUTORS

Bill Hazelton, Kit Curtius, Georg Luebeck

REFERENCES:

- 1 SEER "Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence – SEER 9 Regs Research Data, Nov 2011 Sub (1973–2010) 2010; "
- 2 Ruigomez, A., Wallander, M. A., Lundborg, P., Johansson, S., Rodriguez, L. A. G. "Gastroesophageal reflux disease in children and adolescents in primary care." in *Scand J Gastroenterol* 2010; 45: 2: 139–146
- 3 Ruigomez, A., Rodriguez, L. A. G., Wallander, M. A., Johansson, S., Graffner, H., Dent, J. "Natural history of gastro-oesophageal reflux disease diagnosed in general practice" in *Alimentary Pharmacology & Therapeutics* 2004; 20: 7: .751–60.
- 4 Talley, N. J., Zinsmeister, A. R., Schleck, C. D., Melton, L. J., 3rd "Dyspepsia and Dyspepsia Subgroups – a Population-Based Study" in *Gastroenterology* 102: 4 Pt 1: 1259–68
- 5 Locke, G. R., 3rd, Talley, N. J., Fett, S. L., Zinsmeister, A. R., Melton, L. J., 3rd "Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota." in *Gastroenterology* 1997; 112: 5: 1448–56
- 6 Locke, G. R., 3rd, Talley, N. J., Fett, S. L., Zinsmeister, A. R., Melton, L. J., 3rd "Risk factors associated with symptoms of gastroesophageal reflux" in *Am J Med* 1999; 106: 6: 642–9
- 7 Dent, J., El-Serag, H. B., Wallander, M. A., Johansson, S "Epidemiology of gastro-oesophageal reflux disease: a systematic review." 2005; 54: 5: 710–7
- 8 Thrift, A. P., Kramer, J. R., Qureshi, Z., Richardson, P. A., El-Serag, H. B. "Age at onset of GERD symptoms predicts risk of Barrett's esophagus." in *American Journal of Gastroenterology* 2013; 108: 6: 915–922
- 9 Taylor, J. B., Rubenstein, J. H. "Meta-analyses of the effect of symptoms of gastroesophageal reflux on the risk of Barrett's esophagus" in *Am J Gastroenterol* 2010; 105: 8: 1729–37
- 10 Rubenstein, J. H., Taylor, J. B. "Meta-analysis: the association of oesophageal adenocarcinoma with symptoms of gastro-oesophageal reflux" in *Aliment Pharmacol Ther* 2010; 32: 10: 1222–7



- 11 Falk, G. W., Jacobson, B. C., Riddell, R. H., Rubenstein, J. H., El-Zimaity, H., Drewes, A. M., Roark, K. S., Sontag, S. J., Schnell, T. G., Leya, J., Chejfec, G., Richter, J. E., Jenkins, G., Goldman, A., Dvorak, K., Nardone, G. "Barrett's esophagus: prevalence–incidence and etiology–origins." in *Ann N Y Acad Sci* 2011; 1232: : 1–17
- 12 Kong, C. Y., Kroep, S., Curtius, K., Hazelton, W. D., Jeon, J., Meza, R., Heberle, C. R., Miller, M. C., Choi, S. E., Lansdorp–Vogelaar, I., van Ballegooijen, M., Feuer, E. J., Inadomi, J. M., Hur, C., Luebeck, E. G. "Exploring the Recent Trend in Esophageal Adenocarcinoma Incidence and Mortality Using Comparative Simulation Modeling." in *Cancer Epidemiol Biomarkers Prev* 2014; 23: 6: 997–1006
- 13 Wikipedia "Beta distribution" in http://en.wikipedia.org/wiki/Beta_distribution 2014;



ASSUMPTION OVERVIEW

SUMMARY

We assume that EAC develops from BE through GERD and non-GERD pathways, with subsequent steps described by a multistage cell-based model. The primary calibration of cell kinetic parameters in the MSCE-EAC model is done through maximum likelihood fits to SEER data. Only specific combinations of cell kinetic parameters are identifiable from incidence data.

BACKGROUND

The model represents the transition from normal esophageal squamous tissue to metaplastic BE tissue as a Poisson process with a rate that is higher for individuals with GERD compared to individuals without GERD.

Following generation of a BE segment in some individuals, a multistage carcinogenesis process is used to represent stochastic cellular processes of sequential mutation, clonal expansion to generate hyperplastic tissue, further mutation to malignant status, clonal growth of the malignant tissue, and detailed observation processes, including biopsy sampling, and treatment, including radio-frequency ablation (RFA).

ASSUMPTION LISTING

1. The transition from normal squamous epithelium to BE tissue is a field effect with a transition rate that depends on GERD status.

The normal to BE transition is modeled as a Poisson process with changing rates that depends on GERD onset age. This approach is used to represent a 'field' effect transition, in which a region (or field) of tissue makes the transition during a short interval of time. This assumption appears consistent with available observational data that suggests that BE tissue segments do not generally change significantly in size subsequent to their first detection.

2. We use a three stage model with two rate-limiting mutations that occur prior to premalignant clonal expansion, and a subsequent mutation that generates the malignant phenotype.

The assumption of two initial mutations is consistent with the tumor suppressor paradigm that requires two genetic or epigenetic changes to enter the premalignant phase of carcinogenesis. This is in contrast to the two-stage clonal expansion (TSCE) model that assumes a single mutation prior to clonal expansion. (The TSCE model has a different asymptotic behavior for the hazard at very old ages that tends to plateau). Earlier testing of TSCE versus three-stage models indicated that the three-stage models provide better fits to the digestive-tract (esophageal, stomach, pancreatic, and colorectal) cancers. Although models with more initial mutations are possible, a mathematical analysis indicates that models with two or more mutations prior to clonal expansion have nearly indistinguishable shapes when the product of the initial mutations remain the same. However, if a larger number of initial mutations is assumed, then the mutation rates are forced to become more rapid, and at some point this becomes biologically implausible. These results provide the basis for use of the three stage model.

3. We assume that cells in BE tissue progress independently along the pathway to cancer through birth, death, and mutation processes



Cells on the pathway to cancer are assumed to arise in the BE tissue and progress independently through a stochastic birth, death, mutation, and observation processes for both premalignant and malignant cells, as specified by a multistage model.

4. We assume each premalignant clone begins as low grade dysplasia (LGD), but as the clone grows it may undergo a stochastic transition to become a high grade dysplasia (HGD) clone.

Any of the premalignant cells in LGD or in HGD may undergo malignant transformation, starting the growth of an Esophageal adenocarcinoma (EAC) clone. The model simulates a transition from LGD to HGD using a transition rate that increases in proportion to the number of premalignant cells in the clone. Thus an individual with BE may have multiple clones of different size, some classified as LGD and others as HGD, and EAC clone(s) may begin within any of the LGD or HGD clones.

5. We assume that the number of stem cells in BE has a fixed value per unit area of BE tissue, and that the first two mutations following BE onset occur at equal rates.

Only certain cell kinetic parameter combinations in the multistage model are identifiable through fits to incidence. Identifiable combinations include the net cell proliferation rate for premalignant cells, the product of the number of stem cells and the first two mutation rates, and the product of the premalignant cell division rate and the mutation rate of these cells to become malignant. Thus some assumptions are required to specify non-identifiable parameters. In particular, we assume equality of the first two mutation rates, and set the number of stem cells to values consistent with the available literature. Cell division rates are only weakly identifiable, and were initially set to biologically plausible values. These assumptions were tested through secondary calibration using detailed spatial simulation methods to fit observations from clinical biopsy outcomes, including number and size of LGD, HGD, and pre-clinical EAC clones.

6. The cancer detection process in the multistage model is represented using a stochastic observation process that assumes each cell contributes independently to the detection probability, collectively leading to a size based detection probability.

This size-based observation process generally seems reasonable, as the detection parameter ρ can be adjusted to reflect the median size of tumor at detection. However, the distributional properties of tumor size at detection may not be identical to that seen in clinical practice.

7. Clones generated through the multistage process are assumed to occur at random within the BE segment.

This assumption is necessary because multistage process itself does not include information on the spatial location of clones, only the number and sizes of clones. Simulated clones generated through the multistage process are placed at random on a 2-D surface sized to reflect the BE segment within the esophagus.

8. We assume that LGD, HGD, or EAC detection occurs when biopsied tissue contains cell counts for LGD, HGD, or EAC that exceed specific threshold values.

The predicted biopsy detection process for LGD, HGD or EAC depends on the physical size of clones and the assumed fraction of biopsy tissue required to make a positive identification of HGD or EAC. Several parameters, including fraction of biopsy tissue required for detection, the number of stem cells per unit area, and the division rate of premalignant cells (a weakly identifiable parameter) were evaluated and compared with clinical reports on the detection frequencies for finding HGD and EAC in BE



patients.

9. We assume a per-cell detection probability for symptomatic EAC diagnosis of

$$\rho = 1.0e^{-9}.$$

This specifies that the median size of an EAC at symptomatic detection will contain approximately 10^9 cancer cells.

PARAMETER OVERVIEW

SUMMARY

Most of the model parameters were estimated through maximum likelihood fits to EAC incidence data from nine registries of the Surveillance, Epidemiology, and End Results (SEER) database by single year of age (20 – 84) and calendar year (1975 – 2010). However, some parameters must be fixed initially to achieve parameter identifiability.

BACKGROUND

Key biological parameter combinations may be deduced from the shape of the cancer incidence curve, described mathematically by a hazard function. The incidence curve may be broken into sections representing an exponential–then–linear character of the multistage hazard function as a function of age at diagnosis.

PARAMETER LISTING OVERVIEW

The slope of the linear phase is $\lambda \equiv \mu_0 X \mu_1 p_\infty$ and the growth parameter of the exponential phase $g_p \equiv \alpha_p - \beta_p - \mu_2$. However, the rates μ_0 and μ_1 cannot be estimated separately because the slope λ depends on their product. Analogous to premalignant growth, the malignant growth parameter $g_M \equiv \alpha_M - \beta_M - \rho$. To identify the malignant conversion rate and detection rate per cell, μ_2 and ρ respectively, it is necessary to fix the cell division rates α_p and α_M . Although the product $\alpha_M \rho$ is mathematically identifiable, we were not able to obtain stable estimates and therefore also fixed the (per cell) cancer detection parameter $\rho = 10^{-9}$, which corresponds to median symptomatic detection of EACs when they contain approximately $n \sim 10^9$ cancer cells (see [Assumption Overview](#)).

We compared multiple models by fixing $g_{M,0}$ and detection rate ρ to different values in order to achieve reasonable mean sojourn times and tumor doubling times that are in line with clinical data. In these Results, the EAC clinical detection rate $\rho = 10^{-9}$ per cell/year, malignant cell proliferation rate $g_{M,0} = 0.75$ per cell/year, and $X = 10^6$ stem cells in an average 5 cm BE segment.

Maximum likelihood methods were used to estimate values for $g_{P,0}, g_1, g_2, g_3$ (reference year), where t_b = birth cohort year, that best explain the temporal trends for EAC incidence in terms of sigmoidal birth cohort trends affecting promotion:

$$g_p = g_{P,0} \left(g_1 + \frac{2}{1 + e^{(-g_2(t_b - g_3))}} \right), \quad g_M = g_{M,0} \left(g_1 + \frac{2}{1 + e^{(-g_2(t_b - g_3))}} \right)$$

$$\alpha_p = \alpha_{P,0} \left(g_1 + \frac{2}{1 + e^{(-g_2(t_b - g_3))}} \right), \quad \alpha_M = \alpha_{M,0} \left(g_1 + \frac{2}{1 + e^{(-g_2(t_b - g_3))}} \right)$$

$$g_p = \alpha_p - \beta_p - \mu_2, \quad g_M = \alpha_M - \beta_M - \rho$$

Value (95% CI)	Males	Females
ν_0	3.65 (3.19 – 4.13) $\times 10^{-4}$	7.48 (4.87 – 10.29) $\times 10^{-5}$
$\mu_0(\mu_1)$	7.99 (6.38 – 9.83) $\times 10^{-4}$	7.05 (6.13 – 12.25) $\times 10^{-4}$
μ_2	4.54 (3.65 – 6.47) $\times 10^{-5}$	6.89 (3.16 – 14.28) $\times 10^{-5}$
$g_{P,0}$	9.91 (9.28 – 10.99) $\times 10^{-2}$	1.23 (1.06 – 1.35) $\times 10^{-1}$
g_1	5.09 (2.75 – 5.90) $\times 10^{-1}$	6.40 (2.16 – 8.44) $\times 10^{-1}$
g_2	5.38 (4.83 – 5.72) $\times 10^{-2}$	2.98 (2.47 – 3.44) $\times 10^{-2}$
g_3	1912.5 (1909.1 – 1914.1)	1945.3 (1923.9 – 1954.4)

Table S1: MSCE-EAC model parameters All parameter estimates have the units of per cell per year. Markov Chain Monte-Carlo 95% confidence intervals provided beside the maximum likelihood estimates.



COMPONENT OVERVIEW

SUMMARY

The MSCE–EAC model includes six components, consisting of a model of symptomatic gastroesophageal disease (sGERD), an analytic multistage clonal expansion model hazard for BE and EAC incidence, a temporal trends component, a hybrid stochastic simulation component, a biopsy screening module, and a radio–frequency ablation (RFA) treatment module.

OVERVIEW

The sGERD component was calibrated to sGERD incidence data and age–adjusted sGERD prevalence data from the US and UK to generate estimates of age–dependent sGERD prevalence by gender. The sGERD prevalence influences the rate of transition to BE, and more importantly, the premalignant clonal expansion rate (net growth rate for high grade dysplasia, or HGD) in the multistage model. The multistage model (MSCE–EAC) hazard was calibrated to EAC incidence in SEER, and describes the biological process of transition to BE in the esophagus, and the multistage carcinogenesis process culminating in EAC. The temporal trends component was used to estimate the maximum likelihood period and birth–cohort trends affecting biological processes, including onset of BE and premalignant promotion, for dates ranging from 1890–2010. The hybrid stochastic simulation component was used to provide realizations of HGD and malignant clones in individuals as they age. The biopsy screening module represents biopsy of HGD and malignant clones through quadrant sampling using forceps, which is referred to as the Seattle Protocol. RFA treatment, which is used in patients diagnosed with BE and HGD, is modeled using a module that allows specific proportions of cells of different types to be eliminated during RFA treatment, and then to assess the results on expected diagnoses of HGD and EAC during subsequent years.

COMPONENT LISTING

sGERD component

We modeled gastroesophageal reflux disease (GERD) symptom prevalence at age t , $p_{sGERD}(t)$, based on data from Ruigomez, et al. for incidence (by 2–year age intervals) of GERD symptoms (that occur weekly or more frequently) among children (n=1700), ¹ and another study by Ruigomez, et al. ² on incidence of weekly GERD symptoms among adults (n=1996) with data provided in 10 year intervals.

We used maximum likelihood methods to fit parameters for a GERD prevalence model separately for males and females, using a transition rate to GERD prevalence based on the GERD incidence data and estimating a back–transition rate (representing recovery from GERD) to fit an assumed 20% target rate for age–adjusted GERD prevalence between ages 40–85. See [GERD Model Component](#) for further detail.

Analytic multistage clonal expansion (MSCE) model hazard component

Analytically construct the hazard function of the MSCE–EAC model which consists of two stochastic processes: the random occurrence of BE and the multistage carcinogenesis process arising in BE. Parameters were estimated using maximum likelihood methods. Mathematically, the MSCE–EAC branching process' probability density function (pdf) may be written as a convolution of the BE conversion density f_{BE} (assumed to be exponential) and the MSCE model density after BE onset (f_{BE-MS}) as



$f_{MSCE}(t) = \int_0^\infty f_{BE}(u)f_{BE-MS}(t-u)u$. Further details are provided in the [Model Overview](#) section.

Temporal trends component

This component was used to estimate the mechanistic role of symptomatic GERD (sGERD) and other factors (OF) in driving the observed U.S. trends, and was accomplished in two phases. Phase 1 focused on identifying important biological mechanisms that are likely driving the observed EAC trends. Phase 2 focused on understanding the mechanistic role of sGERD and OF in acting through the biological processes identified in Phase 1 to drive EAC incidence. Both phases of model development were informed by EAC incidence data from SEER, sGERD incidence data from the UK, and US sGERD prevalence data. Separate multiscale models of EAC incidence were built for all-race men and women. See [Temporal Trends Component](#) for further detail.

Hybrid stochastic simulation component

The simulation begins with generation of individual BE onset times, BE segment lengths for each patient (which determines the number of BE stem cells), and generation of pre-initiated and initiated stem cells using Poisson rate-limiting mutation with rate μ_0 and μ_1 , respectively. Initiated premalignant clones undergo independent birth-death-mutation (b-d-m) processes that we simulate to track cell count and times of malignant transformations. See [Stochastic Simulation Component](#) for further detail.

Biopsy screening module

For simulations following the Seattle biopsy protocol, the BE segment can be visualized as partitioned into identical rectangular sections, which we will call "biopsy quadrants" with a single biopsy in the center of the quadrant. For example, an average BE segment of length 5 cm and 7.5 cm circumference will have 12 biopsy quadrants, 3 levels of length 5/3 cm with 4 quadrant biopsies each. Furthermore, we assume periodic boundary conditions when placing clones in a random quadrant.

To account for different biopsy protocols, incompletely described histological methods, and inter-observer variation of neoplasia grade, we present results from the computational model for different levels of diagnostic sensitivity based on the minimum number of neoplastic (pre-malignant/malignant) crypts within a simulated biopsy specimen required for pathologic diagnosis of dysplasia/malignancy among BE patients without prior diagnosis of EAC.

After a simulated screen of BE patients for detection of LGD, HGD, and preclinical EAC at age t_s , the MSCE-EAC model can be used to further simulate an intervention such as an ablative treatment using radio frequency.

Radio-frequency ablation component

After a simulated screen of BE patients for detection of dysplasia and preclinical EAC at age t_s , the MSCE-EAC model can be used to further simulate an intervention such as an ablative treatment using radio frequency. To replicate current practice with radio frequency ablation (RFA), we first remove the prevalent EAC cases that were screen detected at the index endoscopy and then simulate RFA treatment on positively screened patients with dysplasia. The MSCE-EAC model can then be used to project



the EAC incidence and age-specific prevalence of dysplasia into the future after an ablative treatment. The ablation is assumed to curatively reduce all clonal populations and the number of BE crypts by certain percentages as described in the following. As a simple example, we consider the model's predictions after a single ablative treatment when indicated by the presence of high grade dysplasia on future EAC incidence.

REFERENCES:

- ¹ Ruigomez, A., Wallander, M. A., Lundborg, P., Johansson, S., Rodriguez, L. A. G. "Gastroesophageal reflux disease in children and adolescents in primary care." in *Scand J Gastroenterol* 2010; 45: 2: 139-146
- ² Ruigomez, A., Rodriguez, L. A. G., Wallander, M. A., Johansson, S., Graffner, H., Dent, J. "Natural history of gastro-oesophageal reflux disease diagnosed in general practice" in *Alimentary Pharmacology & Therapeutics* 2004; 20: 7: .751-60.



OUTPUT OVERVIEW

SUMMARY

The model outputs range from estimated trends for incidence and mortality, mechanistic factors driving these trends, simulations of biopsy based screening, and treatment through radio-frequency ablation.

OVERVIEW

Outputs from the model include projections of incidence and mortality for US males and females by birth cohort and calendar year, trends for symptomatic gastroesophageal reflux disease (sGERD) in the US, biological parameters of the multistage process that are driving EAC incidence, mechanistic influences of sGERD and other factors (OF) over time, and detailed stochastic simulations of the multistage clonal expansion process.

OUTPUT LISTING

Incidence and mortality trends

US male and female incidence and mortality trends for all-races and whites for males and females – calibrated to SEER data between 1975 and 2010, and ages 20–84

sGERD trends

Trends for symptomatic gastroesophageal reflux disease (sGERD) in the US – consistent with cross-sectional sGERD incidence and prevalence with longer term trends estimated by fitting to SEER EAC incidence

Biological parameters of the multistage process that are driving EAC incidence

Estimated through maximum likelihood fits to SEER data

Mechanistic influences of sGERD and other factors (OF) over time

Fit through extensive maximum likelihood estimation (MLE) and Markov chain Monte Carlo (MCMC) methods to fit SEER incidence data

Joint distribution of premalignant and malignant clones

Detailed stochastic simulations of the multistage clonal expansion process provide explicit realizations of the joint distribution of premalignant and malignant clones using parameters derived through maximum likelihood fitting to SEER incidence data

Costs and life years gained

A simulator of biopsy screening among males and females under different biopsy screening protocols provides estimates of costs and life years gained through different screening protocols.

Reduction or delay of LGD, HGD, and EAC

The simulator estimates the impact of radio-frequency ablation on reducing or delaying the occurrence of LGD, HGD and EAC.



RESULTS OVERVIEW

SUMMARY

Results from the model include estimates for rates and trends for biological processes occurring during EAC carcinogenesis, sensitivities of biopsy protocols, and the impact of radio-frequency ablation (RFA), including the effects of 'touch-up' treatments when the initial treatment was not fully successful.

OVERVIEW

Results from the model include a study on how trends for symptomatic gastroesophageal reflux disease (sGERD) and other factors (OF) in the US may influence biological parameters of the multistage process to drive EAC incidence trends.¹ Similarly, the methods describing the detailed stochastic simulations of the multistage clonal expansion process, and illustration of explicit realizations of the joint distribution of premalignant and malignant clones and simulation of biopsy screening and sensitivity for detection of HGD and EAC are currently under review in a separate manuscript.

RESULTS LIST

Calibration of EAC incidence and incidence-based mortality to SEER data and projection of trends to year 2030

Maximum likelihood methods were used to calibrate the FHCRC model to GERD incidence data and SEER data. These methods provided excellent fits to the SEER data for US incidence and incidence-based mortality for calendar years 1975–2010 and by single-year birth cohorts. During model development, the FHCRC modeling group compared different models using maximum likelihood methods, finding that the premalignant clonal expansion rate differs significantly by birth cohort. The best fit to the data was found using a sigmoidal birth cohort function influencing the premalignant clonal expansion rate.^{2,3}

The calibrated FHCRC model results indicate that there were 81,069 expected male EAC deaths and 10,375 expected female deaths between 1991–2020. Incidence and mortality trends were projected to year 2030 by utilizing birth-cohort specific parameters, with separate projections for local, regional, and distant staged tumors. These projections suggest that male incidence trends are continuing upward, but show a marked flattening trend, reflecting a decreasing birth-cohort trend for later birth cohorts. Trends for females also continue upward to 2030, but unlike for males, there is no significant flattening of the projected trends. Projections of the FHCRC model to 2030 predict that there will be approximately 81,069 male EAC deaths and 10,375 female deaths between 2011–2030.²

The EAC sojourn time may differ by birth cohort

For the FHCRC model, the EAC sojourn time represents the time between appearance of the first malignant cell that doesn't become extinct and the incidental detection of EAC. (This differs from other CISNET models that estimate the time from smallest clinically detectable lesion to EAC incidence). The birth cohort influence on the clonal expansion rates directly influences the expected clonal extinction probability and the expected EAC sojourn time. The FHCRC model predictions for EAC sojourn time range from ~18 years for the 1900 birth cohort, to < 10 years for recent birth cohorts.

Impact of symptomatic GERD and other factors on explaining EAC trends

Biologically based modeling of the mechanistic impact of symptomatic GERD and other factors in fitting to EAC incidence data and GERD incidence data suggests that at most, ~16% of the observed 6-fold increase in EAC incidence between 1975 and 2009 is attributable to GERD, with the remainder explained by other factors. The modeling suggests that GERD influences the transition to BE, but more importantly, GERD increases the rate of premalignant promotion. The other factors also appear to primarily influence premalignant promotion.¹

Dependence of HGD detection and the probability of missed malignancy on the biopsy sampling sensitivity

Detailed simulation of biopsy sampling according to the Seattle protocol suggests that the probability of detecting HGD depends strongly on the percent of biopsy tissue used for analysis, with sensitivities for HGD ranging from ~2–9% for males with sampling percentages ranging from 10–95%. For females, the probability of HGD detection is lower, ranging from ~1–6% for biopsy sampling percentages ranging from 10–95%. The probability of missed malignancy during biopsy sampling ranges from ~20% with 10% biopsy sampling among males, and ~15% among females; to ~5% for males and ~4% for females at 95% sensitivity.³

The predicted impact of ablation in reducing EAC depends on the cell types ablated and the ablation efficiency

Detailed simulations of the lifetime impact of ablation on cumulative EAC incidence was done assuming different scenarios for the efficacy of ablation, with sensitivity analyses comparing elimination of 50%, 99%, and 100% of all cell types, only HGD cells, or only malignant cells. The results indicate that ablation of 100% of HGD and malignant cells delays the expected incidence curve by approximately seven years, with smaller effects seen for less efficient ablation or ablation of selected cell types.^{4,5}

REFERENCES:

- 1 Hazelton William D., Curtius Kit, Inadomi John M., Vaughan Thomas L., Meza Rafael, Rubenstein Joel H., Hur Chin and Luebeck E. Georg. "The Role of Gastroesophageal Reflux and Other Factors during Progression to Esophageal Adenocarcinoma" in *Cancer Epidemiology, Biomarkers and Prevention* 2015; 24: : 1012-1023
- 2 Kong, C. Y., Kroep, S., Curtius, K., Hazelton, W. D., Jeon, J., Meza, R., Heberle, C. R., Miller, M. C., Choi, S. E., Lansdorp-Vogelaar, I., van Ballegooijen, M., Feuer, E. J., Inadomi, J. M., Hur, C., Luebeck, E. G. "Exploring the Recent Trend in Esophageal Adenocarcinoma Incidence and Mortality Using Comparative Simulation Modeling." in *Cancer Epidemiol Biomarkers Prev* 2014; 23: 6: 997-1006
- 3 Curtius, Kit, Hazelton WD, Jeon J, Luebeck EG "A multiscale model evaluates screening for neoplasia in Barrett's Esophagus." in *PLOS Computational Biology* 2015; 11: 5: e1004272
- 4 Heberle CR, Omidvari AH, Ali A, Kroep S, Kong CY, Inadomi JM, Rubenstein JH, Tramontano AC, Dowling EC, Hazelton WD, Luebeck EG, Lansdorp-Vogelaar I, Hur C "Cost-Effectiveness of Screening Patients with Gastroesophageal Reflux Disease for Barrett's Esophagus With a Minimally Invasive Cell Sampling Device" in *Gastroenterol Hepatol* 2017; S1542-3565: 17: 30197-0
- 5 Kroep S, Heberle CR, Curtius K, Kong CY, Lansdorp-Vogelaar I, Ali A, Wolf WA, Shaheen NJ, Spechler SJ, Rubenstein JH, Nishioka NS, Meltzer SJ, Hazelton WD, van Ballegooijen M, Tramontano AC, Gazelle GS, Luebeck EG, Inadomi JM, Hur C "Impact of Radiofrequency Ablation Treatment of Barrett's Esophagus on Esophageal Adenocarcinoma: A Comparative Modeling Analysis." in *Clin Gastroenterol Hepatol* 2017; 1542-3565: 17: 30019-8



GERD MODEL COMPONENT

SUMMARY

Gastroesophageal Reflux Disease (GERD) Model

OVERVIEW

sGERD component

We modeled gastroesophageal reflux disease (GERD) symptom prevalence at age t , $p_{sGERD}(t)$, based on data from Ruigomez, et al. for incidence (by 2-year age intervals) of GERD symptoms (that occur weekly or more frequently) among children (n=1700),¹ and another study by Ruigomez, et al.² on incidence of weekly GERD symptoms among adults (n=1996) with data provided in 10 year intervals.

DETAIL

We used maximum likelihood methods to fit parameters for a GERD prevalence model separately for males and females, using a transition rate to GERD prevalence based on the GERD incidence data and estimating a back-transition rate (representing recovery from GERD) to fit an assumed 20% target rate for age-adjusted GERD prevalence between ages 40–85. We then found that we could achieve excellent fits to these data by simplified (3 parameter) gender-specific models representing a (slower) transition rate among children, a transition age, and an adult rate for acquiring weekly GERD symptoms (See [Model Overview](#)).

BE prevalence $F_{BE}(t)$ can be estimated, via parameter ν_0 by fitting to SEER data and fixing a value for relative risk $RR \in \{5, 10, 15\}$, given the model for GERD prevalence as described in the main text with the BE conversion rate,

$$\nu(t) = \nu_0 \left((1 - p_{sGERD}(t)) + RR \cdot p_{sGERD}(t) \right)$$

$$\Rightarrow F_{BE}(t) = \Pr[T_{BE} \leq t] = 1 - e^{-\int_0^t \nu(s) ds}$$

REFERENCES:

- ¹ Ruigomez, A., Wallander, M. A., Lundborg, P., Johansson, S., Rodriguez, L. A. G. "Gastroesophageal reflux disease in children and adolescents in primary care." in *Scand J Gastroenterol* 2010; 45: 2: 139-146
- ² Ruigomez, A., Rodriguez, L. A. G., Wallander, M. A., Johansson, S., Graffner, H., Dent, J. "Natural history of gastro-oesophageal reflux disease diagnosed in general practice" in *Alimentary Pharmacology & Therapeutics* 2004; 20: 7: 751-60.

FRED HUTCHINSON
CANCER RESEARCH CENTER
A LIFE OF SCIENCE

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



TEMPORAL TRENDS COMPONENT

SUMMARY

Esophageal Adenocarcinoma incidence has increased over six-fold in the U.S. since 1975.

OVERVIEW

This component was used to estimate the mechanistic role of symptomatic GERD (sGERD) and other factors (OF) in driving the observed U.S. trends, and was accomplished in two phases.

Phase 1 focused on identifying important biological mechanisms that are likely driving the observed EAC trends.

Phase 2 focused on understanding the mechanistic role of sGERD and OF in acting through the biological processes identified in Phase 1 to drive EAC incidence. Both phases of model development were informed by EAC incidence data from SEER, sGERD incidence data from the UK, and US sGERD prevalence data. Separate multiscale models of EAC incidence were built for all-race men and women.

DETAIL

The Phase 1 model family was designed to identify biological mechanisms that may potentially drive the observed EAC incidence trends. In these models, linear or sigmoidal trends for cohort and/or period were applied to one or more biological processes. Thus all individuals of a given age, period, birth cohort, and sex share the same set of biological rates, but these rates may change with birth cohort and calendar year.

The Phase 2 model family extended the Phase 1 models by stratifying the population according to sGERD duration, and then evaluating the mechanistic role of sGERD and OF acting on important biological mechanisms identified in Phase 1. In these Phase 2 models, linear or sigmoidal trends for cohort and/or period were applied to sGERD and OF, which influence biological rates. Individuals of a given age, period, birth cohort, and sex were stratified by decade of sGERD onset, with individuals in each stratum modeled using baseline biological rates before acquisition of sGERD and generally different rates after sGERD onset.



STOCHASTIC SIMULATION COMPONENT

SUMMARY

Carcinogenesis is represented using a multistage clonal expansion model, linking cellular events (division, death, and mutation) to premalignant and malignant clonal growth and cancer detection.

OVERVIEW

The stochastic simulation algorithm (SSA) is a mathematically exact method to follow each event that occurs during a realization of a continuous time Markov chain beginning with a single cell, using cell kinetic parameters fit to SEER incidence data and other sources.

DETAIL

Considering an individual premalignant clone of size X_t at time t , we define the intensity function vector $r(X_t) = (\beta_p X_t, \mu_2 X_t, \alpha_p X_t)$ for death/differentiation, malignant transformation, and birth of new P stem cell, where, over a short period of time s , we expect $r_j(X_t)s + o(s)$ events of type j to occur. Due to the Markovian property of the process, we wait an exponential length of time until the next event occurs with intensity $r_0(X_t) = \sum_{j=1}^3 r_j(X_t) = X_t(\beta_p + \mu_2 + \alpha_p)$. Once an exponential time to next event is chosen, we jump to the neighboring state $X_t + v_j$ with probability $r_j(X_t)/r_0(X_t)$, where v_j is the j^{th} component of the state change vector $v = (-1, 0, 1)$ for the b-d-m process. Fortunately, in the case of the P clone process with constant rates, the probabilities $r_j(X_t)/r_0(X_t)$ are constant with respect to the current state X_t so we may generate a number K of events of the three types with probabilities $(\frac{\beta_p}{\beta_p + \mu_2 + \alpha_p}, \frac{\mu_2}{\beta_p + \mu_2 + \alpha_p}, \frac{\alpha_p}{\beta_p + \mu_2 + \alpha_p})$ and cumulatively sum each $X_t + v^j$ step for the K chosen events to create a state vector $\{ \cdot \mid \mathbb{N} \}$. Then we generate the K exponential waiting times of the process at once from an exponential with mean $\lambda_t = N(\beta_p + \mu_2 + \alpha_p)$ and cumulatively sum these to arrive at a new later time $t_2 > t$.

The SSA works very well when cell count of the P clone is small and the event intensities $r(X_t)$ are fluctuating quickly. In particular, our simulation benefits to use the SSA for the beginning of a P clone's growth from a single cell, when the probability of extinction is high (β_p is only slightly smaller than α_p) and most clones are eliminated after a small number K of initial events. However, the SSA can become excruciatingly slow when a P clone becomes very large, i.e. contains a large number of stem cells. Therefore, rather than simulating every event choice and time, we can employ an accelerated but approximate procedure called the τ -leap method, first introduced by Gillespie and others.^{1,2,3} The goal of this procedure is to advance the cell count by a preselected time increment τ in contrast to the exponential time increments generated in the SSA. To control the loss of accuracy with this approximation, the choice of leap-size τ must satisfy the historically referenced "leap condition" which is large enough that many events occur in that time, but nevertheless small enough that the intensity function value r is likely to change only "infinitesimally" as a consequence of those events. To the extent that this condition is satisfied, the mathematical rationale in replacing Markovian kinetics with Poisson kinetics⁴ states that the number of times each independent event j will occur in the set time length τ can be approximated by a

FRED HUTCHINSON
CANCER RESEARCH CENTER
A LIFE OF SCIENCE

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



Poisson random variable with mean $\omega(t, t + \tau)$ on the interval $(t, t + \tau)$. For the ordinary τ -leap scheme, we assign $\omega(t, t + \tau) = r_j(X_t)\tau$. Thus, we set the intensity of event j equal to the constant $r_j = r_j(X_t)$ and we update the cell count vector $X_{t+\tau} = X_t + \sum_{j=1}^3 n_j v_j$ where n_j are independent Poisson variates with means $r_j\tau$.

When the stochastic simulation of P clones produces a malignant progenitor M cell, an independent birth–death–detection process for an M clone begins also. This can occur during anytime of surveillance and the malignant clones may employ the same algorithm described above. Considering an individual malignant clone of size X_t at time t , we define the intensity function vector $r(X_t) = (\beta_M X_t, \rho X_t, \alpha_M X_t)$ for death/differentiation, EAC detection, and birth of new M stem cell. The times of ρ events may occur between screens and be counted as a spontaneous, interval detection of EAC.

REFERENCES:

- ¹ Gillespie, Daniel T. "Exact stochastic simulation of coupled chemical reactions" in J Phys Chem 1977; 81: : 2340–2361
- ² Gillespie, Daniel T. "Approximate accelerated stochastic simulation of chemically reacting systems" in The Journal of Chemical Physics 2001; 115: 4: 1716
- ³ Cao, Y.; Gillespie, D. T.; Petzold, L. R. "Efficient step size selection for the tau-leaping simulation method" in The Journal of Chemical Physics 2006; 124: 4: 044109
- ⁴ Kurtz, Thomas G. "Approximation of Population Processes" in SIAM 1981; : 75

KEY REFERENCES

- Cao, Y.; Gillespie, D. T.; Petzold, L. R.** (2006) Efficient step size selection for the tau-leaping simulation method in *The Journal of Chemical Physics* 124:4, p 044109
- Curtius, Kit, Hazelton WD, Jeon J, Luebeck EG** (2015) A multiscale model evaluates screening for neoplasia in Barrett's Esophagus. in *PLOS Computational Biology* 11:5, p e1004272
- Dent, J., El-Serag, H. B., Wallander, M. A., Johansson, S** (2005) Epidemiology of gastro-oesophageal reflux disease: a systematic review. 54:5, p 710-7
- Falk, G. W., Jacobson, B. C., Riddell, R. H., Rubenstein, J. H., El-Zimaity, H., Drewes, A. M., Roark, K. S., Sontag, S. J., Schnell, T. G., Leya, J., Chejfec, G., Richter, J. E., Jenkins, G., Goldman, A., Dvorak, K., Nardone, G.** (2011) Barrett's esophagus: prevalence-incidence and etiology-origins. in *Ann N Y Acad Sci* 1232, p 1-17
- Gillespie, Daniel T.** (1977) Exact stochastic simulation of coupled chemical reactions in *J Phys Chem* 81, p 2340–2361
- Gillespie, Daniel T.** (2001) Approximate accelerated stochastic simulation of chemically reacting systems in *The Journal of Chemical Physics* 115:4, p 1716
- Hazelton William D., Curtius Kit, Inadomi John M., Vaughan Thomas L., Meza Rafael, Rubenstein Joel H., Hur Chin and Luebeck E. Georg.** (2015) The Role of Gastroesophageal Reflux and Other Factors during Progression to Esophageal Adenocarcinoma in *Cancer Epidemiology, Biomarkers and Prevention* 24, p 1012-1023
- Heberle CR, Omidvari AH, Ali A, Kroep S, Kong CY, Inadomi JM, Rubenstein JH, Tramontano AC, Dowling EC, Hazelton WD, Luebeck EG, Lansdorp-Vogelaar I, Hur C** (2017) Cost-Effectiveness of Screening Patients with Gastroesophageal Reflux Disease for Barrett's Esophagus With a Minimally Invasive Cell Sampling Device in *Gastroenterol Hepatol* S1542-3565:17, p 30197-0
- Kong, C. Y., Kroep, S., Curtius, K., Hazelton, W. D., Jeon, J., Meza, R., Heberle, C. R., Miller, M. C., Choi, S. E., Lansdorp-Vogelaar, I., van Ballegooijen, M., Feuer, E. J., Inadomi, J. M., Hur, C., Luebeck, E. G.** (2014) Exploring the Recent Trend in Esophageal Adenocarcinoma Incidence and Mortality Using Comparative Simulation Modeling. in *Cancer Epidemiol Biomarkers Prev* 23:6, p 997-1006
- Kroep S, Heberle CR, Curtius K, Kong CY, Lansdorp-Vogelaar I, Ali A, Wolf WA, Shaheen NJ, Spechler SJ, Rubenstein JH, Nishioka NS, Meltzer SJ, Hazelton WD, van Ballegooijen M, Tramontano AC, Gazelle GS, Luebeck EG, Inadomi JM, Hur C** (2017) Impact of Radiofrequency Ablation Treatment of Barrett's Esophagus on Esophageal Adenocarcinoma: A Comparative Modeling Analysis. in *Clin Gastroenterol Hepatol* 1542-3565:17, p 30019-8
- Kurtz, Thomas G.** (1981) Approximation of Population Processes in *SIAM*, p 75
- Locke, G. R., 3rd, Talley, N. J., Fett, S. L., Zinsmeister, A. R., Melton, L. J., 3rd** (1997) Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota. in *Gastroenterology* 112:5, p 1448-56
- Locke, G. R., 3rd, Talley, N. J., Fett, S. L., Zinsmeister, A. R., Melton, L. J., 3rd** (1999) Risk factors associated with symptoms of gastroesophageal reflux in *Am J Med* 106:6, p 642-9

- Rubenstein, J. H., Taylor, J. B.** (2010) Meta-analysis: the association of oesophageal adenocarcinoma with symptoms of gastro-oesophageal reflux in *Aliment Pharmacol Ther* 32:10, p 1222-7
- Ruigomez, A., Rodriguez, L. A. G., Wallander, M. A., Johansson, S., Graffner, H., Dent, J.** (2004) Natural history of gastro-oesophageal reflux disease diagnosed in general practice in *Alimentary Pharmacology & Therapeutics* 20:7, p .751-60.
- Ruigomez, A., Wallander, M. A., Lundborg, P., Johansson, S., Rodriguez, L. A. G.** (2010) Gastroesophageal reflux disease in children and adolescents in primary care. in *Scand J Gastroenterol* 45:2, p 139-146
- SEER** (2010) Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence – SEER 9 Regs Research Data, Nov 2011 Sub (1973–2010)
- Talley, N. J., Zinsmeister, A. R., Schleck, C. D., Melton, L. J., 3rd** Dyspepsia and Dyspepsia Subgroups - a Population-Based Study in *Gastroenterology* 102:4 Pt 1, p 1259-68
- Taylor, J. B., Rubenstein, J. H.** (2010) Meta-analyses of the effect of symptoms of gastroesophageal reflux on the risk of Barrett's esophagus in *Am J Gastroenterol* 105:8, p 1729-37
- Thrift, A. P., Kramer, J. R., Qureshi, Z., Richardson, P. A., El-Serag, H. B.** (2013) Age at onset of GERD symptoms predicts risk of Barrett's esophagus. in *American Journal of Gastroenterology* 108:6, p 915-922
- Wikipedia** (2014) Beta distribution in http://en.wikipedia.org/wiki/Beta_distribution,
-