Optimal timing for cancer screening and adaptive surveillance using mathematical modeling

Kit Curtius^{a,*}, Anup Dewanji^b, William D. Hazelton^c, Joel H. Rubenstein^{d,e}, E. Georg Luebeck^c

^aBarts Cancer Institute, Queen Mary University of London, EC1M 6BQ London, United Kingdom

^bApplied Statistics Unit, Indian Statistical Institute, Kolkata 700108, India

^cProgram in Computational Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

^dDivision of Gastroenterology, University of Michigan Medical School, Ann Arbor, Michigan, USA

^eVeterans Affairs Center for Clinical Management Research, Ann Arbor, MI, USA

Abstract

Cancer screening and early detection efforts have been partially successful in reducing incidence and mortality but many improvements are needed. Although current medical practice is mostly informed by epidemiological studies, the decisions for guidelines are ultimately made *ad hoc*. We propose that quantitative optimization of protocols can potentially increase screening success and reduce overdiagnosis. Mathematical modeling of the stochastic process of cancer evolution can be used to derive and to optimize the timing of clinical screens so that the probability is maximal that a patient is screened within a certain “window of opportunity” for intervention when early cancer development may be observable. Alternative to a strictly empirical approach, or microsimulations of a multitude of possible scenarios, biologically-based mechanistic modeling can be used for predicting when best to screen and begin adaptive surveillance. We introduce a methodology for optimizing screening, assessing potential risks, and quantifying associated costs to healthcare using multiscale models. As a case study in Barrett’s esophagus (BE), we applied our methods for a model of esophageal adenocarcinoma (EAC) that was previously calibrated to US cancer registry data. We found optimal screening ages for patients with symptomatic gastroesophageal reflux disease to be older (58 for men, 64 for women) than what is currently recommended (age > 50 years). These ages are in a cost-effective range to start screening and were independently validated by data used in current guidelines. Our framework captures critical aspects of cancer evolution within BE patients for a more personalized screening design.

Significance: Our study demonstrates how mathematical modeling of cancer evolution can be used to optimize screening regimes. Surveillance regimes could also be improved if they were based on these models.

Graphical Abstract:

*Corresponding author: k.curtius@qmul.ac.uk

Quick Guide to Equations

Ideally, a sensitive clinical screen will be offered to a patient during an opportunistic age window such that some event A (e.g., premalignant disease onset) has likely already occurred but an event B (e.g., cancer detection) has not yet happened. The rates of these events can depend on various risk factors (e.g., sex, ethnicity, environmental exposures) but the derivation of timing is universal. To maximize the probability that a patient is between T_A and T_B years of age at time of screening/surveillance t_s is equivalent to simultaneously (1) maximizing $\Pr[T_A \leq t_s]$ to ensure that event A has already occurred while (2) minimizing $\Pr[T_B \leq t_s, T_A \leq t_s]$ so that screens are not recommended when it is too late for an early intervention. This idea is reflected mathematically in the following relationships, which we will use in the optimal screen design methodology for quantifying ‘screening success’.

- Optimal screening age for some weight w on an adverse outcome, before time of patient death T_D

$$t_s^* = \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot \Pr[T_B \leq t_s, T_A \leq t_s])$$

- Associated risk of adverse outcome for screening at time t_s

$$R_B(t_s) = \Pr[T_B \leq t_s]$$

- Overdiagnosis rate - detecting precursor that will not become cancer by end of follow-up (age) T_f

$$OD(t_s) = \Pr[T_A \leq t_s | T_B > T_f]$$

- Successful diagnosis rate (i.e., early detection of future case of B) by time (age) $T_f > t_s$

$$SD(t_s) = \Pr[T_A \leq t_s | t_s < T_B < T_f]$$

- Over-screening - proportion of overdiagnosed cases in yield of A at screening time t_s

$$OS(t_s) = \Pr[T_C > T_f | T_A \leq t_s]$$

- Cost function $C(t_s, T_f)$ associated with screening time t_s and surveillance until a final age T_f

$$C(t_s, T_f) = OS(t_s) \cdot (T_f - t_s)$$

- Associated risk for adaptive follow-up interval, conditional on screening/surveillance result (SR) at time t_{s_i}

$$R_{B|SR}(t_{s_{i+1}}) = \Pr[T_B \leq t_{s_{i+1}} | SR(t_{s_i})]$$

Introduction

The main rationale for cancer screening is that earlier detection of a disease during a patient’s lifetime offers the opportunity to change its prognosis. Compared with incidental cancers, screening can improve the prognosis for patients with early occult cancers detected before symptoms develop, and perhaps even more importantly through removal of pre-invasive lesions such as adenomas in the colon, cervical intraepithelial neoplasia (CIN) in the cervix, and ductal carcinoma in situ (DCIS) in the breast. In 2018, it was estimated that nearly half of all cancer incidence in the US was attributable to detection by screening [1] but it is difficult to assess how many of these patients were destined to die from other causes before these cancers would have led to a symptomatic diagnosis, thus proving their screens to be ineffective. Current programs suffer from the problem of overdiagnosis of benign lesions and yet, with often dismal and costly consequences, there is still under-diagnosis of dangerous lesions because they were either missed by the screen itself or there was a lack of screening uptake by high risk individuals at the appropriate age [2]. Consequently, improvement in screening success is an important health policy research area, and one primed for quantitative assessment. Simply put, there is a *mathematical balancing act* to solve for human populations from both public health and cost-effectiveness perspectives: the goal is to maximize successful prevention of future lethal cancers (often by removing their precursors detected on a screen) and minimize the likelihood of overdiagnosis.

From a biological perspective, malignant cells develop in the body originally from normal cells at birth through the evolutionary process of carcinogenesis [3]. This process of cancer evolution is stochastic, meaning that random mutations can occur in cells throughout life that may eventually lead to a malignant phenotype selected in a tissue, but not necessarily so. Mathematical models of carcinogenesis, tumour progression, and metastasis are powerful tools to describe this process and, when applied to data, can infer important parameters for evolution such as time of metastatic seeding and selection strength of certain mutations that may confer resistance to treatment [4, 5]. Because such variables are not immediately measurable from data, mathematical models are now being explicitly incorporated into translational aims such as interval planning for adaptive therapy trials in cancer patients [6, 7]. We propose that these quantitative models can be further utilized at an earlier stage for cancer prevention - namely, to infer sojourn (waiting) times for specific stages of cancer evolution that we wish to target, and then to use this information to optimize the efficacy of cancer screening strategies and adaptive surveillance for early detection [8].

Clinical detection of biomarkers on a screen that alert us to phenotypic changes along the path to cancer (e.g., premalignant adenomas found during colorectal cancer (CRC) screening) are a snapshot of *field cancerization*, wherein groups of cells have acquired some but not all of the phenotypes necessary for clinical malignancy [9]. In CRC, the single first malignant cell creates a “Big bang” of tumor initiation, wherein the new malignant clone that develops already has all of the selective advantage it needs to progress, and random mutations accumulated are essentially neutral from an evolutionary perspective [10]. Before this occurs however, we know that there are certain stages of colorectal carcinogenesis that are both common in *type* (*APC* gene inactivation for initiation of adenomas) and in *timing* (adenomas become detectable around age 45 in males and females) [11]. Thus, although no two paths of carcinogenesis are exactly the same in two patients, there are most probable timescales for this stochastic process to take place (as seen in age-specific incidence curves) while still allowing for less likely events to happen (e.g., very young CRC cases).

These stochastic features were captured in the first mathematical descriptions of cancer, which have been studied and applied extensively in combination with biostatistical methods for over 60 years [12, 13]. By analyzing the structure of multistage models, we can also formulate these probability distributions as certain “windows of opportunity” that are crucial to capture in cancer screening planning, such as defining the

likelihood to detect premalignant clones and simultaneously the likelihood that malignant clones would be unlikely to have developed yet. In this study, we derive conclusions from the theory of stochastic processes applied to carcinogenesis to formally optimize the timing of initial cancer screens in a population, and optimize subsequent surveillance adaptively, conditional on a previous exam result.

Here we present our analytical results for a generalized framework and then apply it in a proof-of-concept study for current screening of Barrett’s esophagus (BE), the metaplastic precursor to esophageal adenocarcinoma (EAC). Briefly, because BE itself is essentially asymptomatic, the majority of BE patients remain undiagnosed and thus most EAC cases are diagnosed at an advanced stage. This is unfortunate because, (1) mortality associated with EAC is very high (majority of diagnosed patients die within a year), and (2) prior diagnosis of BE is positively associated with improved EAC survival in patients [14]. Therefore US and UK gastroenterologists have been focused on positively identifying BE patients on screening endoscopy, and then recommending these patients to undergo lifelong endoscopic surveillance in order to detect any dysplasia or early cancer that may be removed, thus preventing future lethal EAC. However the clinical reality is that only 0.2-0.5% of people with BE develop EAC each year [15]. Thus, the majority of diagnosed BE patients will attain little benefit per follow-up exam and over-surveillance of BE patients with no evidence of premalignant high grade dysplasia (HGD) poses a costly problem due to lifelong surveillance recommendations (as frequently as every 3 years by current guidelines). Beyond cost-effectiveness for a small number of microsimulated screening ages [16], the *optimal* starting age for BE screening in at-risk subpopulations is unknown.

Some screening recommendations have been justified by microsimulation models [11] but most screening guidelines consider current epidemiological data alone and are ultimately proposed by and decided by experts. The clinical consequences of these decisions are recorded as population outcomes, and then further research determines how effective these recommendations proved to be. When appropriate randomized control trial data is lacking, as is commonly the case in cancer screening [17], we propose that the initial choice of screening design could first be determined using an optimization algorithm with a machine learning and/or model-based approach. To address this problem, we derive probability equations to be used for optimal screening/surveillance timing and risk estimation, and present results for BE screening ages using a calibrated mathematical model that reproduces US cancer incidence data. Our predictions are in a cost-effective range and were independently validated with endoscopic screening data used for current BE guideline rationale.

Materials and Methods

Below we derive the mathematical framework to determine the optimal timing of screening and surveillance of ‘at-risk’ patients using stochastic models of carcinogenesis. Herein, *cancer screening* refers to initial testing for the presence of a specified premalignant/malignant change, and subsequent tests that are offered after a positive screen diagnosis are defined as *surveillance*. If the age at completion of lead time (i.e., the length of time between the early detection of a preclinical state and the future clinical, symptomatic detection) surpasses patient lifetime, this patient will be considered an overdiagnosed case for that cancer. Published estimates of clinical endpoints like cancer incidence and the prevalence of precursors are reported on a population level, thus providing measures for mean dwell times in certain preclinical disease states that can be measured from population screening. For cancers with detectable precursors, we seek the age for which the probability is maximal that 1) individuals in a population harbor a screen-detectable premalignant clone, and 2) this timing occurs before incidental cancers have developed (Figure 1A).

Biological processes not revealed in population data alone give rise to considerable inter-patient heterogeneity during the progression from normal tissue to incident cancer, and thus optimal timing for follow-up is

Figure 1: Optimizing screening and surveillance using branching process models of cancer evolution.

A) The goal of early detection and prevention is to perform an initial screen on individuals within a larger population after premalignant changes reside in cells but before incidental cancer occurs in order to intervene. After an initial screen at time t_s^* , adaptive recommendations for patient-specific follow-up surveillance can iteratively account for the heterogeneity in a screened population. B) Somatic evolution of stem cell lineages since birth leads to stochastic trajectories (certain symmetric division examples shown on tree nodes) that accumulate mutations and may be selected for advantageous phenotypes. For example, within an at-risk patient's esophagus, normal squamous epithelium may transform to a columnar BE segment with rate $\nu(t)$. Stochastic appearance of premalignant *P* cells can occur (e.g., after 'two-hit' gene inactivation of *TP53* in a cancer-promoting microenvironment), which can then undergo clonal expansion described by birth-death-mutation processes. Malignant cells *M* that are initiated, in turn, may undergo clonal expansions or go extinct. Clinical detection of *M* clones that do not go extinct may occur through a size-based detection process with rate ρ . C) Criteria for optimal screening strategies can be formally defined using probability distributions in (A) for the timescales that characterize certain stages of carcinogenesis (see Text for details).

expected to be highly heterogeneous within the population based on individual screening outcome. To account for this heterogeneity along with the temporal effects of screening (e.g., different prognoses resulting from

screening earlier versus later), mathematical models can employ a stochastic approach to incorporate premalignant and malignant clonal expansions during clonal evolution explicitly (Figure 1B) [18, 19, 20, 21, 22]. This offers the advantage that the onset of (detectable) precursor clones can be defined as random variables and clonal population sizes can be tracked over time so that dwell times, risk of future cancer, and adaptive surveillance intervals may be formally derived (Figure 1C).

We first use basic tenets of probability theory to quantify these timescales in equations that can be applied for cancers that have measurable precursors to target (such as adenomas in the colon). The methods are organized in four parts: (1) General equations for optimal screening and application to multistage models of cancer, (2) Three strategies for optimal screening, (3) Metrics for efficacy assessment, and (4) Adaptive surveillance for 2 potential (negative/positive) screening outcomes. Then we apply the framework to determine the optimal timing of BE screening and surveillance in Results.

Optimal screening framework

Optimal design options for weighting adverse outcomes. For some event A and later event B , maximizing the probability that a patient is between ages T_A and T_B at time of screening t_s is equivalent to maximizing $\Pr[T_A \leq t_s]$ to ensure that event A has already occurred while simultaneously minimizing $\Pr[T_B \leq t_s, T_A \leq t_s]$ to ensure that event B has not yet occurred by time t_s . This statement is reflected mathematically as,

$$\Pr[T_A \leq t_s \leq T_B] = \Pr[T_A \leq t_s] - \Pr[T_B \leq t_s, T_A \leq t_s]. \quad (1)$$

If event B implies event A , the second term on the right hand side of this equation simplifies to $\Pr[T_B \leq t_s, T_A \leq t_s] = \Pr[T_B \leq t_s]$, the absolute risk of event B by time t_s .

It may not be a justifiable public health goal to equally weigh the criteria that both event A has occurred and event B has not occurred, as is the case for maximizing the probability in Eq. (1). In general for events A and B , we define the probability of a “successful screen” with some specified positive weight w ,

$$\Pr[T_A \leq t_s] - w \cdot \Pr[T_B \leq t_s, T_A \leq t_s]. \quad (2)$$

When $w = 1$, as in Eq. (1), both events contribute equally during optimization that a screen time occurs within the specific disease state window. As an alternative to equal weighting, Dewanji et al. [23] considered a simple illness-death model and maximized the probability that time t_s be in the window between events A and B (where event B implies event A) with an extra penalty term for event B (which in our notation corresponds to the case $w = 2$).

If $w > 1$, maximization of Eq. (2) will yield earlier optimal screening ages t_s^* because there will be greater weight to avoid event B before screening, while older screening ages will be selected as optimal for $w < 1$. Cancer screening policy makers must decide how much weight should be given for each potential screening age t_s , i.e., how heavily to guard against events B in the general population, such as cancer, balanced against the prospect of futile screens when patients have not yet undergone event A , such as premalignant onset.

Incorporating realistic time constraints for screening. Next we incorporate realistic cutoff times due to life expectancy to formally condition that we do not optimize screening ages that are unrealistically old, i.e., when the patient is more probable to have already died than to have developed the disease we wish to detect early enough to intervene. Here we include life table estimates for all cause mortality survival probabilities in the US [24] as an independent random variable T_D in order to calculate $\Pr[T_D \geq t_s]$, which is a close approximation when cause-specific cancer death rates in the general population are low, as is the case for

esophageal adenocarcinoma. Thus, this method allows for random truncation due to death from other causes. For the general weighting $w > 0$ for events A and B , we compute the criteria defined as,

$$\Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot \Pr[T_B \leq t_s, T_A \leq t_s]). \quad (3)$$

In these methods we will derive optimal screen times t_s^* by maximizing Eq. (3) and refer to these t_s^* as the screening times that satisfy the *optimality criterion* for a given value of weighting factor w ,

$$t_s^* = \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot \Pr[T_B \leq t_s, T_A \leq t_s]). \quad (4)$$

Quantifying risk for screening times. To perform a constrained optimization problem with additional risks incorporated, we quantify a penalty for certain screening times by introducing an equation for the risk $R_B(t_s^*)$ that some later event B has already occurred given some optimal screening time t_s^* ,

$$R_B(t_s^*) = \Pr[T_B \leq t_s^*].$$

For cancer screening guideline decision-making, we can increase the weight w in Eq. (3) to find optimal time t_s^* such that $R_B(t_s^*) < \epsilon$, where ϵ is a threshold parameter for the allowable amount of risk of event B . Alternatively, if the risk of possibly allowing an undesirable event B before a proposed time t_s^* is too small to be of considerable worry, we reduce the weight w assigned to event B and thus raise the risk of event B to an appropriate, or tolerable, level.

Applying design to multistage carcinogenesis models. We derive optimal screening scenarios in a case study of Barrett's esophagus (BE) using the multistage clonal expansion for esophageal adenocarcinoma (MSCE-EAC) model that has been employed within the Cancer Intervention and Surveillance Modeling Network (CISNET), a National Cancer Institute consortium, for cost-effectiveness studies of EAC prevention (see [25, 26, 27, 28] and Supplementary Figure S1 for model details). The random variables described in the next section will refer to stages of progression to cancer appropriate for BE but could be analogously formulated and expanded for similar multistage models of screening in colon, lung, breast, and others.

Random variables for screening outcomes. The random variables of interest that represent four clinical points of the multistage process are:

$$\begin{aligned} T_A &= \text{time of abnormal precursor } \mathcal{A} \text{ onset (e.g., BE onset, ulcerative colitis onset, etc.)} \\ T_P^d(t) &= \begin{cases} 0 & \text{if no detectable premalignant clone exists at time } t \\ 1 & \text{if there exists at least one detectable premalignant clone at time } t \end{cases} \\ T_M^d(t) &= \begin{cases} 0 & \text{if no detectable malignant clone exists at time } t \\ 1 & \text{if there exists at least one detectable malignant clone at time } t \end{cases} \\ T_C &= \text{time of cancer (e.g., EAC) spontaneous (symptomatic) detection} \end{aligned}$$

As applied in [25, 27], we derive the forthcoming formulas using a general precursor $\mathcal{A}$ onset distribution $f_{\mathcal{A}}(t)$,

$$f_{\mathcal{A}}(t) = \nu(t) e^{-\int_0^t \nu(u) du}, \quad (5)$$

where $\nu(t)$ is the rate of precursor development. As has also been applied recently in case studies of gastric,

lung, and oral cancers [29], we allow $\nu(t)$ to be time-dependent to allow straightforward incorporation of etiological agent prevalence that affect initiation rates in a population. For BE, we model the density $f_{BE}(t)$ with rate $\nu(t)$ given as a function of the prevalence of symptomatic gastroesophageal reflux disease (GERD) at age t ,

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

where $p_{sGERD}(t)$ is the prevalence of GERD symptoms at age t , RR is the relative risk of BE initiation due to exposure to GERD symptoms, and ν_0 is a positive constant.

Unlike the one-time occurrence for T_A and T_C , the detectability of premalignancy ($T_P^d(t)$) and malignancy ($T_M^d(t)$) require specific definitions for appropriate binary random variables. If we designate these times to be the time of initiation or malignant transformation of a cell regardless of its fate (i.e., the ancestor to a clonal progeny that eventually goes extinct or that survives), we may derive analytical probabilities with first-passage time for continuous random variables T_P and T_M , respectively, using master equations. However, it is more clinically relevant to define these random variables as binary outcomes for detectable clones at the time of screening t_s . In the derivations of our screening windows, the probability of detection at time t_s of a premalignant clone that began with a single ancestor P cell initiated at time s is denoted as $p_2(s, t_s) = \Pr[T_P^d(t_s) = 1 | P(s) = 1]$, and analogously the probability of detection at time t_s of a malignant clone that began with a single ancestor M cell initiated at time s , $p_1(s, t_s) = \Pr[T_M^d(t_s) = 1 | M(s) = 1]$.

In the Results below, we assign these binary detection variables the value 1 if a clone born at time s is non-extinct at time t_s (size of clone is ≥ 1) and 0 otherwise. In the case of EAC, this “perfect sensitivity” definition will be useful as progressively better high-resolution imaging technologies and minimally invasive sampling devices are tested in trials in BE [16, 30] and used in future clinical practice. Lastly, we present the two survival functions that will appear throughout the mathematical derivations: $\Pr[T_C \geq t_s | T_A = \tau] = S_{MSCE}(t_s - \tau)$, and $\Pr[T_C \geq t_s] = S_C(t_s)$, where $S_{MSCE}(t)$ is the survival function for the multistage clonal expansion model of interest after precursor A onset and $S_C(t)$ is the survival function for the full model ($S_{EAC}(t)$ for the MSCE-EAC model, see Supplementary Methods 1).

Strategies for optimizing initial screening ages

Strategy 1: Optimization of precancerous yield before cancer. This strategy for initial screening success aims to maximize the probability that an individual has developed the precursor of interest before the time of screening but has not yet developed clinical cancer (Figure 1C, beige),

$$\Pr[T_A \leq t_s \leq T_C, T_D \geq t_s]. \quad (6)$$

Thus to satisfy the optimality criterion in this scenario for weighting parameter w described by Eq. (3), we solve for the optimal screening age t_s^* (Eq. (4)) as follows

$$\begin{aligned} t_s^* &= \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot \Pr[T_C \leq t_s]) \\ &= \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot (1 - S_C(t_s))). \end{aligned} \quad (7)$$

Strategy 2: Optimization of precancerous yield before dysplasia. This strategy for initial screening success aims to maximize the probability that an individual has developed the precursor before the time of screening but has not yet developed a detectable premalignant clone (e.g., dysplasia, Figure 1C, light pink),

$$\Pr[T_A \leq t_s, T_P^d(t_s) = 0, T_D \geq t_s].$$

Then solving the optimality criterion for weighting parameter w on dysplasia development, we solve for this optimal screening age t_s^* as follows,

$$\begin{aligned} t_s^* &= \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot \Pr[T_P^d(t_s) = 1]) \\ &= \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot \left(\Pr[T_A \leq t_s] - w \cdot \int_0^{t_s} \Pr[T_P^d(t_s) = 1 | T_A = \tau] \cdot f_A(\tau) d\tau \right). \end{aligned} \quad (8)$$

We utilize the filtered Poisson process (FPP) approach to analytically solve for $\Pr[T_P^d(t_s) = 1 | T_A = \tau]$ in the integrand (see Supplementary Methods 2 for derivation). Thus together with $\mathcal{A}$ onset distribution in Eq. (5) and mortality risk, we have the analytical solution for optimal screening age in Eq. (8).

Strategy 3: Optimization of dysplasia yield before cancer. As a final example, we aim to maximize the probability that an individual has developed detectable dysplasia before the time of screening but has not yet developed a synchronous screen-detectable malignancy (Figure 1C, dark pink),

$$\Pr[T_P^d(t_s) = 1, T_M^d(t_s) = 0, T_D \geq t_s].$$

Thus the optimal screening time t_s^* for Strategy 3 for specified w is given by,

$$t_s^* = \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_P^d(t_s) = 1] - w \cdot \Pr[T_P^d(t_s) = 1, T_M^d(t_s) = 1]). \quad (9)$$

Integrating over all possible onset times for precursor $\mathcal{A}$, we can also derive the analytical solution for this optimal screen time (full derivation in Supplementary Methods 3).

Metrics for assessing screening efficacy

Successful screening of future cancers cases. For screening effectiveness results, we define a successful diagnosis function $SD(t_s)$ for the probability of successful screening for future cancers before an age T_f ,

$$\begin{aligned} SD(t_s) &= \Pr[T_A \leq t_s | t_s < T_C < T_f] = \frac{1}{S_C(t_s) - S_C(T_f)} \int_0^{t_s} \Pr[t_s < T_C < B | T_A = \tau] \cdot f_A(\tau) d\tau \\ &= \frac{1}{S_C(t_s) - S_C(T_f)} \int_0^{t_s} (S_{MSE}(t_s - \tau) - S_{MSE}(T_f - \tau)) \cdot f_A(\tau) d\tau. \end{aligned}$$

Thus $SD(t_s)$ is the proportion of patients who get cancer before T_f , who biologically have $\mathcal{A}$ (undiagnosed) before screening age t_s . Here we distinguish T_D , a random variable for patient-specific death, from T_f , a known constant time point (e.g., age 80) representing a set length of patient follow-up for comparison.

Overdiagnosis at screening age. Alternatively, $OD(t_s)$ is the probability of overdiagnosis (OD), given by

$$\begin{aligned} OD(t_s) &= \Pr[T_A \leq t_s | T_C > T_f] = \frac{\Pr[T_A \leq t_s, T_C > T_f]}{\Pr[T_C > T_f]} \\ &= \frac{1}{S_C(T_f)} \int_0^{t_s} S_{MSE}(T_f - \tau) \cdot f_A(\tau) d\tau. \end{aligned}$$

Thus $OD(t_s)$ is the proportion of patients who will not get cancer before age T_f but who have $\mathcal{A}$ at time t_s .

Proportion of over-screening in a positive-screen population. Another important quantity is the proportion of positive screens at the time t_s who will undergo needless, costly surveillance. We can reformulate this proportion as a function with terms previously computed from $OD(t_s)$ above and Strategy 1 (Eq. (6)),

$$OS(t_s) = \Pr[T_C > T_f | T_A \leq t_s < T_C] = \frac{\Pr[T_A \leq t_s < T_f < T_C]}{\Pr[T_A < t_s < T_C]}.$$

Proportion of successful screening in a positive-screen population. Lastly, the fraction of those who will be a successful screen (SS), for whom surveillance may be life-saving due to early detection of small cancers that will form prior to symptomatic detection, can be considered the positive predictive value (PPV) of the screen. We can reformulate this proportion as the complementary conditional event of $OS(t_s)$ above,

$$SS(t_s) = \Pr[t_s < T_C < T_f | T_A \leq t_s < T_C] = 1 - OS(t_s). \quad (10)$$

Comparing costs of different screening ages. To quantify the effects of different screening ages, we define a cost function $C(t_s, T_f)$ for screening time t_s and surveillance until an age T_f . We aim to minimize the cost function which incorporates needless surveillance of the over-screened population from age t_s to age T_f ,

$$C(t_s, T_f) = \underbrace{OS(t_s) \cdot (T_f - t_s)}_{\text{lifetime surveillance of overdiagnosed patients}}. \quad (11)$$

Here we assume that over-treatment/over-surveillance costs are likely proportional to the fraction of over-screened individuals multiplied by a set number of life-years after diagnosis. Assuming that within each strata there is equal risk of premalignancy to treat during surveillance, this weighting determined by number of surveillance years is reasonable in our cost metric.

Optimizing re-screening and surveillance in an adaptive way

After a screen, current clinical guidelines recommend fixed follow-up examination intervals based on a crude analysis of existing evidence, often without incorporating known risk factors of patient progression (e.g., patient age, gender). In our proposed method for adaptive surveillance, mechanistic modeling of the carcinogenesis process is assessed *conditional* on patient demographic/clinical features and detected stage of progression (or lack thereof) at screening time t_{s1} to provide a more refined surveillance recommendation.

Outcome 1 (negative for precursor): when to re-screen. After a normal screen, an open question is whether the patient should be asked to return for repeat screen at a later time. Analogous to Strategy 1, we could aim to screen this individual again at a later age t_{s2} when he/she is most probable to have developed the precursor by this time but not yet clinical cancer,

$$\Pr[T_A \leq t_{s2} < T_C, T_D \geq t_{s2} | T_A > t_{s1}, T_D > t_{s1}].$$

We derive the optimal age (see Supplementary Methods 4 for full solution) for a next screening time as,

$$t_s^* = \arg \max_{t_s} \Pr[T_D \geq t_{s2} | T_D > t_{s1}] \cdot (\Pr[T_A \leq t_{s2} | T_A > t_{s1}] - w \cdot \Pr[T_C \leq t_{s2} | T_A > t_{s1}]). \quad (12)$$

Outcome 2 (positive for precursor): associated cancer risk. After diagnosis, we can iteratively condition on the screening/surveillance result at time t_{s_i} , $SR(t_{s_i})$, and derive an optimal interval for the next surveillance given some chosen outcome of interest to obtain $t_{s_{i+1}}^*$. For Outcome 2 with no dysplasia/neoplasia detected, the subsequent risk for cancer developing before a suggested next surveillance exam at time t_{s_2} is,

$$R_{C|SR}(t_{s_2}) = \Pr[T_C \leq t_{s_2} | SR(t_{s_1})] = \Pr[T_C \leq t_{s_2} | T_A < t_{s_1}, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0].$$

In general, a physician will often have no way of knowing how long a patient has harbored undetected cancer precursors in the body, only that onset occurred sometime before screening ($T_A < t_{s_1}$). But if we could measure this precursor onset time τ , then his/her associated cancer risk for next surveillance age t_{s_2} would be more specific than defined above and would be defined (see Supplementary Methods 4 for derivation) as,

$$R_{C|SR}(t_{s_2}) = \Pr[T_C \leq t_{s_2} | T_A = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0]. \quad (13)$$

Outcome 2 (positive for precursor): Strategy 1

In ‘adaptive’ surveillance, we consider when the ideal next surveillance exam should happen conditional on the outcome of a screen or last surveillance exam. Assuming we have an estimate τ of a patient’s precursor onset time, we solve for optimal next-surveillance age $t_{s_2}^*$ (see Supplementary Methods 5 for full solution),

$$t_{s_2}^* = \arg \max_{t_{s_2}} \Pr[T_M^d(t_{s_2}) = 0, T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0].$$

Previously calibrated parameter values reproduce population cancer incidence. The optimal screen design methodology presented thus far is applicable to multistage models, in particular branching process disease models with defined stages of clonal expansions occurring during carcinogenesis. In a case study for BE screening presented here, we apply the optimality criteria defined above for the MSCE-EAC model specifically.

For this model, the cell-level description is linked to the population scale by means of the model hazard function, defined as the instantaneous rate of detecting cancers among individuals who have not been previously diagnosed with cancer [27]. Briefly, the hazard function can be derived from the backward Kolmogorov equations for the stochastic multistage process described above and solved numerically via a system of coupled ordinary differential equations (Supplementary Methods 1). Thus, one may infer rates of cellular processes from cancer incidence data (see [25, 27] for model results using EAC incidence data from the Surveillance, Epidemiology, and End Results (SEER 9) registry 1975-2010, with predicted trends to 2030). For the quantitative Results in the use case below, we evaluate the equations defined in the above methods using values for evolutionary variables previously estimated in [27], provided in Supplementary Table S1.

Results

Recommendations for risk-stratified Barrett’s esophagus (BE) screening and surveillance regimes have been proposed to optimize beneficial use of healthcare resources, but these have been limited in strata choices, variable across organizations, and based on conclusions from observational data alone [31]. For example, one current US guideline from the American College of Gastroenterology (ACG) strongly recommends BE screening for men with chronic and/or frequent (more than weekly) gastroesophageal reflux disease (GERD) symptoms and two or more risk factors for BE or EAC. These risk factors include: age greater than 50 years, Caucasian race, presence of central obesity, current or past history of smoking, and a confirmed family history of BE or EAC (in a first-degree relative) [32]. Alternatively in the UK, the British Society of Gastroenterology recommends endoscopic BE screening for any gender of patients with chronic GERD symptoms and multiple

Figure 2: Contour plot of optimality criteria for single screen Strategy 1. The optimality criteria are plotted for given screening age t_s between ages 10 and 119 and weighting parameters $w \in [0, 20]$. The red line on the 2D contour plot denotes the optimal screening ages for all races combined, born in 1950.

risk factors (at least three among: age 50 years or older, white race, male sex, obesity)[33]. After BE diagnosis, all of the guidelines recommend a certain uniform timing of surveillance and treatment regimes that are determined by presence or absence of detectable dysplasia (no other risk factors considered formally, including patient age), which were proposed by experts and deemed cost-effective [28].

Instead, we applied our framework in the case of BE to obtain optimal screening and surveillance design from theoretical predictions *a priori*. In our notation for this framework (Materials and Methods), we specifically refer to initial precursor development $T_A = T_{BE}$, the detection of premalignant (high grade dysplasia) and preclinical malignant clones $T_P^d(t_s)$ and $T_M^d(t_s)$, respectively, and clinical cancer occurrence $T_C = T_{EAC}$.

Optimal ages to initialize screening in Barrett’s esophagus. We first used the model to optimize the choice of recommended age to start screening, rather than assuming age 50 as the threshold to begin screening across risk groups as guidelines suggest currently (in all those that specify an age [31]). We applied Strategy 1 (Materials and Methods) for a single screen to maximize the probability that an individual has developed

Table 1: Optimal screening ages for BE in GERD and general populations

Risk group	Sex	Optimal age t_s^* (95% CI)	SD	OD	PPV	Cancer risk t_s^*
General	Male	64 (63 - 66)	99.5%	2.4%	16.8%	0.17%
GERD		58 (56 - 60)	99.6%	7.1%	28.6%	0.37%
General	Female	69 (68 - 71)	100 %	0.6%	9.6%	0.04%
GERD		64 (59 - 67)	100 %	1.9%	17.9%	0.11%

Model estimates of optimal screening ages t_s^* and associated metrics for Strategy 1 ($w = 1$, see Materials and Methods) including successful diagnoses (SD) of future EAC cases, overdiagnoses (OD) of non-EAC cases, positive predictive value (PPV), and the risk of cancer occurring before the screening age recommendations for US persons, all races, born in 1950, stratified by sex and GERD symptom status.

BE before the time of screening but has not yet developed clinical EAC, thus obtaining optimal screening ages t_s^* defined in Eq. (7) for weight parameters w .

In line with current risk factors, we computed optimality criteria for timing of individuals screened for BE stratified by age, sex, and GERD status (i.e., general population or symptomatic GERD population). Figure 2 depicts the contour plots of these optimality criteria for individuals born in year 1950 along with the optimal ages for each w (red lines). We sought ages by whole years for simplicity, in the way that guidelines are formulated currently. For $w = 1$ (i.e., equal weighting of positive screen and safeguarding from cancer before screening), the optimal screening times were (A) $t_s^* = 64$ years old for males, (B) $t_s^* = 58$ years old for males with GERD symptoms (indicated by red diamonds), (C) $t_s^* = 69$ years old for females, and (D) $t_s^* = 64$ years old for females with GERD symptoms (indicated by red circles). In consistency with previous efficacy and cost-effectiveness analyses of BE screening [16, 25, 27, 28], we present results applied to the 1950 birth cohort as a ‘base case’ study and note that screening optimization predictions are robust through modern birth cohorts in a 50 year range (Supplementary Figure S2). This is consistent with previous BE prevalence estimates at index endoscopy that found only a small secular age-specific trend in index BE prevalence in white male GERD screenees but no trend in women across the calendar year period years 2000-2006 [34].

Table 1 provides the optimal ages computed for stratified populations along with associated metrics of efficacy. Optimal screening ages t_s^* defined in Eq. (7) depend on a choice of parameter w , to protect against the adverse outcome of EAC development before age of screening. For males in the general population, the range of optimal screening times was 48-68 years of age (mean age 53 years) with $w \in [0, 20]$ for Strategy 1.

Weighting of associated costs affects screening optimization. From a preventative viewpoint, screening recommendations may be aimed to reduce the risk that patients developed high grade dysplasia (HGD) prior to screening to avoid invasive radiofrequency ablative treatment after index endoscopy. To determine how optimal screening ages shift if we aim to screen patients who initially have BE but do not yet have existing HGD clones, we applied Strategy 2 in our optimization to obtain optimal screening times (see Eq. (8)).

Figure 3 depicts the contour plots of the optimality criteria for individuals born in 1950 stratified by age, sex, and GERD status along with the optimal ages for each w (red lines). For $w = 1$ (i.e., equal weighting of positive screen and safeguarding from HGD before screening), the optimal screening times were $t_s^* = 59$ years old for males and $t_s^* = 65$ years old for females, which are 5 and 4 years earlier for males and females, respectively, than the optimal ages computed for the general population in Strategy 1. Interestingly, for the symptomatic GERD population, there is no optimal time obtained for $w = 1$ because it is unfeasible to screen early enough such that those with GERD and BE do not have any small HGD clones present. However, to lessen the cost of having small HGD clones at time of screening in our optimization, we set $w = 0.95$ and

Figure 3: Contour plot of optimality criteria for single screen Strategy 2. The optimality criteria are plotted for given screening age t_s between ages 10 and 119 and weighting parameters $w \in [0, 1]$. Note, as the weighting factor increases the optimal screen age becomes earlier. The red line on the 2D contour plot denotes the optimal screening ages for all races combined, born in 1950.

found the same optimal ages for GERD patients as were found using Strategy 1.

Optimal screening ages can improve ad hoc guidelines by decreasing over-surveillance. We compared our optimal screening ages with current BE screening guidelines in the US and UK [32, 33] that recommend screening high risk male patients with GERD symptoms starting at age 50 and older. This choice of starting age was based on results from the US Clinical Outcomes Research Initiative (CORI study that quantified age-specific prevalence of BE in patients (mainly white race) with no prior EAC at time of first endoscopy, stratified by sex and presence of GERD symptoms as indicated in the endoscopy report [34] (Figure 4A).

First we directly compared the CORI prevalence with our model by computing the analogous conditional probability to represent the fraction of first-time diagnosed BE cases by age within the total 155,641 cancer-free patient cohort who underwent first endoscopy between 2000-2006,

Figure 4: Validating predicted screening ages and computing optimal ages with Strategy 1. A) Model predictions of the conditional prevalence (solid lines, Eq. (14)) recapitulate CORI data for BE prevalence detected on index endoscopy in cancer-free individuals (dashed lines). B) Strategy 1 (single age results using Eq. (7) with $w = 1$, 95% confidence intervals shaded) incorporates both protection from previous cancer formation (missed cases) and all-cause mortality pre-screening time into optimal screening decision-making (green diamonds).

$$\Pr[T_A \leq t_s | t_s \leq T_{EAC}, T_D \geq t_s] = \frac{\Pr[T_A \leq t_s \leq T_{EAC} | T_D \geq t_s]}{\Pr[T_{EAC} \geq t_s]} = \frac{\Pr[T_A \leq t_s \leq T_{EAC}]}{S_{EAC}(t_s)}, \quad (14)$$

where this numerator was computed previously in Strategy 1 (see Eq. (6)).

This data independently validated the similar age trends predicted by our model (which was fit to SEER incidence, irrespective of CORI) for these strata born in 1950 (Figure 4A for $w = 1$). In only the case for the highest risk group, our estimates of prevalence were found to be slightly greater than the published estimates likely because the model calculates the true yield of all BE cases detectable by a highly sensitive screen rather than diagnosed BE cases among patients in the CORI study, which may have also suffered from a ‘harvesting’ effect for older age groups that contained less patients [34]. Further, we independently confirmed the two main Results from Rubenstein et al. that: (1) the yield of upper endoscopy for the diagnosis of BE rapidly increases among men with GERD until approximately age 55, and then the prevalence begins to level-off, and (2) women with GERD were at no increased risk compared with men without GERD (red and green lines, respectively) and thus, if screening were considered, it should be recommended at the same age (or not at all, if lacking additional risk factors). With the ability to compute not only this prevalence analytically but furthermore an optimal criterion with Strategy 1, we can maximize the BE yield at any continuous age t subject to a *joint* probability of being cancer-free and alive (Figure 4B for $w = 1$). For previous guideline formulations, such clinical decisions had to be made essentially ad hoc or *conditional* with cohort data grouped into decade age-groups [34] for proposals of most effective screening age for BE yield.

With our mathematical formulation, we can also quantitatively compare costs associated with chosen screening initialization ages. Using symptomatic GERD males as our example, we compared the initial screen age of $t_s = 50$ (current practice) versus our optimal age of $t_s^* = 58$. The costs associated with over-surveillance defined in Eq. (11) in detecting cancers up to age $T_f = 80$ in these two cases was,

$$C(50, 80) = 19.40 \quad \text{vs.} \quad C(58, 80) = 15.71$$

Thus, the overall costs associated with screening at optimal age $t_s^* = 58$ were estimated to be less than current costs, mainly due to savings gained from eliminating 8 years of futile surveillance in a large number of middle-aged patients. For both screening ages 50 and 58 in GERD men, there was less than 0.5% risk of screening too early (i.e., very few men who develop EAC will have developed BE after those screening ages and progressed undetected without surveillance). Also, screening younger ages did not always imply increased numbers of future cancer cases detected, the main goal for successful screening. The probability of a missed EAC case occurring in a male GERD patient between these two ages was minimal,

$$\Pr[50 < T_{EAC} < 58] = S_{EAC}(50) - S_{EAC}(58) = 0.28\%.$$

The current cost burden for EAC cancer control is then likely attributable moreso to years of futile surveillance than costs associated with missed cases expected to arise before later screening age recommendations.

Less than 3% of normal patients will be found with BE on re-screen. For a first outcome, suppose a patient is screened for suspected BE and the screen at time t_{s_1} returns negative for BE (Outcome 1, Materials and Methods). An important decision for clinicians is whether to suggest that this patient come back for a one-time follow-up second screen at a later age and if so, at what age they should return. Current guidelines do not advise re-screening unless a patient is healing from esophagitis wounding to ensure underlying BE was not missed initially [32]. This recommendation was based on a CORI study that found 2.3% of patients with Outcome 1 had BE on repeat examination, implying that BE is rarely missed and develops early in patients on surveillance [35]. With our methods we can quantify: (1) when is the best age to offer a second screen, and (2) how likely it would be that BE would be found at that age.

We optimized when to screen this individual again at a later age t_{s_2} when he/she is most probable to have developed BE by this time but not yet clinical cancer. Given a range of prior screening times $t_{s_1} \in \{45, 50, 58, 64, 69\}$, Figure 5 shows results for optimal subsequent screening times t_{s_2} , where $t_{s_1} < t_{s_2} < 120$. The black diamonds depict the optimal screen times $t_{s_2}^*$ (see Eq. (12)) for this simulated cohort (US population, all races, born in 1950). For our choices of t_{s_1} in ascending order, we found that optimal follow-up screens were (A) $t_{s_2}^* = \{73, 74, 76, 77, 78\}$ for all males combined, (B) $t_{s_2}^* = \{72, 73, 75, 76, 77\}$ for GERD males, (C) $t_{s_2}^* = \{77, 78, 79, 81, 82\}$ for all females combined, and (D) $t_{s_2}^* = \{76, 77, 79, 80, 81\}$ for GERD females. Thus, even though initial screening times may be over 20 years apart, the optimal range of a follow-up screen at age $t_{s_2}^*$ is only within a 5 year window.

Our prevalence figures for these are very close to those found in the CORI study for both genders [35], further validating our straightforward approach to assess public health questions that typically depend on grouped, conditional empirical population data (alive, cancer-free) for prevalences rather than optimally scheduled predictions of screening outcomes and yields based on continuous age. Importantly, we found that the initial age of screening (rather than time since initial screening) affects the future probability of BE yield; such information is not currently considered when clinicians decide whether certain GERD patients should return to be re-screened. For example we found that, if initial screening took place for GERD males at age 58, there would be a 1% chance of a positive re-screen rather than 2 – 3% when screening first at age 50. Importantly, even this 1% will be mostly overdiagnoses as *de novo* BE is much less likely to have time to develop

Figure 5: Optimality criteria for Outcome 1 (negative for BE): Strategy 1. The optimality criteria for each choice of next screening time t_{s2} , given prior screening age $t_{s1} \in \{45, 50, 58, 64, 69\}$ are denoted for weighting $w = 1$ (including typical case for current practice, $t_{s1} = 50$, shown in dashed lines). Optimal re-screen times t_{s2}^* shown for all races, born in 1950 (diamonds), and for the risk group-specific optimal t_{s1}^* (dotted lines), were (A) $t_{s2}^* = 77$ for all males, (B) $t_{s2}^* = 75$ for GERD males, (C) $t_{s2}^* = 82$ for all females, and (D) $t_{s2}^* = 80$ for GERD females.

to EAC within the remainder of patient lifespan, thus further devaluing a re-screen after initial negative result.

Accurate cancer risk prediction depends on premalignant molecular age. In Outcome 2, a patient is found to have BE but no neoplasia (HGD/malignancy) at screening time t_{s1} . Current guidelines suggest these patients return for surveillance endoscopies every 3-5 years for the rest of their lives [32, 33] because they still carry risk of EAC that may in fact increase with time regardless of persisting without a diagnosis of dysplasia [36]. Furthermore, non-dysplastic BE patients with “missed” HGD/EAC were significantly older than those who progressed later [37], suggesting the important role age plays as a significant risk factor indicating missed EAC and the need for surveillance of BE patients who have unknowingly harbored BE for longer than younger patients. This prompts the question whether patient-specific BE “dwell time” influences future neoplastic risk and thus also the efficacy of timing of surveillance schedules.

If we measured BE onset time τ for a patient with Outcome 2 at time t_{s1} (e.g., via methylation-based molecular clocks [26]), then his/her associated EAC risk for given next screening age t_{s2} is given by Eq. (13). Figure 6 depicts EAC risk for US 1950 birth cohort, with prior screening ages $t_{s1} \in \{45, 50, 58, 64\}$ for males (A) and $t_{s1} \in \{45, 50, 64, 69\}$ for females (B). As shown in Figure 6, the model predicted that in general the

Figure 6: Associated EAC Risk for Outcome 2 (positive for BE). The EAC risk in Eq. (13) is plotted for each choice of next screening time t_{s2} until age 85, given prior screening age $t_{s1} \in \{45, 50, 58, 64\}$ (denoted by line type) for males (A) and $t_{s1} \in \{45, 50, 64, 69\}$ for females (B), and a BE onset age $\tau \in \{20, 30, 40\}$ (denoted by color).

associated EAC risk for those that developed BE earlier in life is greater than the risk for those BE patients who developed BE later in life. For the screen performed at time $t_{s1} = 50$ (dashed lines) in both sexes, the BE patient who developed BE at age 20 (purple line) has a 2.4 predicted relative EAC risk versus the BE patient who developed BE at age 40 (yellow line). Currently in clinical practice however, these two patients would likely be treated equally because they were diagnosed at the same age. Our findings reiterate that heterogeneity in BE onset ages may translate to large differences in EAC risk [26].

Sensitivity of predictions. Our multistage model was calibrated to population incidence data [27] and has been previously tested to quantify the effects of perturbations in model parameters in Supplementary Table S1 such as initiation and growth rates of dysplastic clones that reflect published premalignant prevalence within BE (see details in [25, 38, 39]). Similarly, in these analyses, we tested our model for risk-stratified populations and found that our optimal timing predictions are robust to perturbations in parameters, as calculated for optimal ages in a bootstrap analysis of 1,000 re-sampling iterations of the Markov chain Monte Carlo (MCMC) posterior distributions for each calibrated parameter (Figure 4B, shaded regions).

We also considered 5 additional modern birth cohorts beyond the 1950 base case and found that: (1) Optimal screening ages remain unchanged for males (GERD and general populations) born between 1950 and 2000, and (2) similar results to the base case were found for females such that optimal screening times t_s^* varied between these cohorts by only 3 years in females with GERD and 2 years for the general female population (Supplementary Figure S2). For optimal choice of adaptive surveillance based on screening outcomes however, we found that results will be sensitive to certain patient-specific parameters such as BE onset age and thus these will be important to include in order to perform more personalized surveillance.

Data and code availability

All analytical distributions to recreate our Results and apply to other data are available in the Text and Supplementary Methods. Parameters used as input data for the case of BE screening are provided in Supplementary Table S1. Code to solve equations was developed in R (version 3.6.1). This framework will be made publicly available upon publication (github.com/yosoykit/OptimalTimingScreening).

Discussion

Mathematical models of cancer evolution are being widely developed to understand timing of events and dynamics of treatment resistance. In this study, we applied such models to optimize screening and surveillance. Three examples of modeling techniques that could be formally assessed for screening timing include Markov models for natural history of disease [40, 41, 42], biologically-based models that incorporate dynamic processes like clonal expansions and biomarker shedding [18, 19, 22, 25, 43, 44], and biological event timing models that infer ordered genetic events [45]. While utilizing different data types, these all include distinct stages of latency periods separated by rate-limiting events on a patient's forecasted trajectory. Our mathematical formalism provides one means of targeting these windows to answer early detection questions with models. Our approach also provides the flexibility of choosing the 'window of opportunity' and the specified weight w for the adverse event to suit the purpose of the health policy goal.

Several previous studies address optimal screening design via biomathematical modeling [18, 19, 23, 46], but this study targeted cancer precursors specifically and allowed for this flexible weighting of events to determine the specific age to screen (and whether to screen) in an analytical framework that does not require simulation. Microsimulation studies using models have also been used to inform policy-making decisions on clinical recommendations/modalities, for example in colorectal cancer [47] and lung cancer [40] screening, by the US Preventive Services Task Force (USPSTF). The model we present in our study has been previously used in this way to evaluate certain screening and surveillance scheduling of Barrett's esophagus (BE) to prevent esophageal adenocarcinoma (EAC) [28], but our optimal design framework developed here contributes unique strengths based on theory, and can be used to complement and enhance these current methods of screening design assessment. In particular, while microsimulation analyses can evaluate the benefits and harms of many hypothetical scenarios, this process can be time-intensive, prone to sensitivities in parameter choices, and ultimately only test a finite number of options [48]. In contrast, we showed that the probabilities of successful population screening can be analytically solved in certain cases and that these equations 1) are straightforward to optimize in a single computation, 2) reveal the structure of parameters and model assumptions that have the greatest implications on the sensitivity of the results obtained, and 3) predict future trends to obtain results for younger birth cohorts in a logical, data-driven way.

In the case of BE, screening has been recommended to begin in patients with multiple risk factors including increasing age (the specific threshold is 50 years of age), but the clinical evidence for these recommendations have been low to moderate in quality [31]. In order to justify the proposed recommendations, cost-effectiveness studies have used decision analytic models to quantify expected costs accrued versus life-years gained via simulations of initial screening and subsequent surveillance including endoscopic eradication therapy of detected neoplasia. While these models have each typically tested a small number of screening cohorts, we considered the full range of screening ages and found that the *optimal ages* to begin screening patients with symptomatic gastroesophageal reflux disease (GERD) were age 58 for men and 64 for women. These ages are still within cost-effective ranges found for men with GERD in previous modeling analyses

that compared a few possible screening ages (including 50, 60, and 70) in sensitivity analyses [16, 49], particularly when using minimally invasive, cheaper screening modalities even in women [16]. With our framework we also reasoned that earlier screening ages are more costly, although the cost function defined serves as a simple example for how optimal control functions in cancer control programs could quantify trade-offs of futile surveillance with underdiagnoses, but extensions could also include quality-adjusted life years (QALYs).

Using our framework in a similar way, we also determined that if BE screening were recommended in the general population, the optimal ages to recommend would be age 64 for men and age 69 for women (for equal weighting given to BE onset and EAC risk in optimization). These findings for the ages of optimal BE yield are supported by prospective studies' results that it is more likely to find BE during screening at ages closer to 60 than 50 [50], specifically in two large studies that found the same mean age at BE diagnosis of 61 in men [51], and 61 in both genders together [52]. Lastly, women found with BE were more likely to be older, between ages 61 to 70 at diagnosis, than women found without BE [53, 54]. Taken together, these studies confirm our findings that it may be justified biologically to begin screening at these later ages rather than worrying a large portion of the population starting at age 50. These proposed older ages to initialize screening in GERD and general population could be specifically tested in future cost-effectiveness analyses.

Our model describes a number of key transitions in cell biology, each defined by a specific rate parameter. Direct measurement of these parameters is usually lacking, and indeed there remains uncertainty about what key rate-limiting transitions in cancer evolution are essential. Nevertheless, our model is broadly resistant to this uncertainty as we found in the tightly estimated range for optimal screening ages in our sensitivity analyses, because the main timescales and windows of opportunity to target in our strategies are robust to perturbations in cell-scale variables to ensure that we recapitulate expected cancer incidence. First, the model's stages were determined via rigorous likelihood ratio testing when fit to SEER cancer registry data, and its structure is further supported by whole-genome sequencing data pointing to the existence of boundary-defining mutations (or combinations of mutations) in BE that create windows for curative interventions [55]. We then must make certain assumptions - for example that the first two 'hits' of initiation of high grade dysplasia in BE (e.g., inactivation of each allele of *TP53*) are equally likely because we can only infer the product of the two rates with this method. Although our application does not explicitly include molecular data, we note that these parameters could be more uniquely measured using multi-region samples of clones within patient BE segments over time, especially as resources like Pre-Cancer Genome Atlases are becoming available to the public [56].

With respect to effective surveillance, we did find that molecular age of precursors (beyond patient chronological age at time of screening) was a critical variable for accurate cancer risk prediction given screening outcome. Specifically, the effect of BE dwell time (time since onset to clinical diagnosis) can be built into a model for optimizing screening and we urge the importance of finding innovative means (such as molecular clocks with patient-specific data) to measure these "hidden" events that currently cannot be clinically ascertained by traditional methods [57]. Because the framework is predicated on a biological description of the natural history of cancer, the inferences drawn are predictive and will need to be validated prospectively even though the underlying mechanistic model framework was validated using available screening data. This approach complements the development of adaptive treatment strategies (ATS) and dynamic treatment regimen (DTR) that rely on large amounts of patient-specific health records and machine learning [58]. Although these new developments are promising and are likely to yield important insights considering the wealth of clinical records for common diseases, they are difficult to extend to rare or understudied diseases and therefore are prone to biases. In contrast, models borrow strength from incidence patterns and are easily interpretable

and applicable to balance risks with benefits when there is an intention to treat or intervene.

While it is difficult to justify delaying screening from a public health perspective, the problem of over-diagnosis is only becoming larger as fervor for early detection assays is growing. Our study’s predictions are reasonable considering the scarcity of epidemiological data to suggest earlier screening is ideal. We show that using a mechanism-based model of cancer could help relieve the costs of unnecessary screening and surveillance, though further validation of associated risks and sensitivities is needed. Overall such approaches are currently under-utilized in public health policies, but there is clear potential to improve guideline formation through deeper mathematical examination of data for cancers where screening can be life-saving. As “learning cancer screening programs” are being proposed to test many potential strategies in populations via randomized screening arms [17], we suggest that incorporating a model-based approach from the outset could aid in an optimal design plan (such as screening age choice in certain risk strata) in order to improve health outcomes faster. We believe that innovative teamwork between physicians, mathematical modelers, and policy-makers can improve screening design thereby reducing future cancer mortality.

1. Acknowledgements

This research was supported by the National Cancer Institute (www.cancer.gov) under grants U01 CA152926 (CISNET), U01 CA182940 (BG-U01), U54CA163059 (BETRNET), and U01CA199336 (CISNET), and UKRI Rutherford Fund Fellowship (KC). The authors also thankfully acknowledge the contributions of raw data on Barrett’s Esophagus prevalence from the National Endoscopic Database of the Clinical Outcomes Research Initiative (CORI).

References

- [1] R. C. Wender, O. W. Brawley, S. A. Fedewa, T. Gansler, R. A. Smith, A blueprint for cancer screening and early detection: Advancing screening’s contribution to cancer control, *CA Cancer J Clin* 69 (1) (2019) 50–79.
- [2] R. A. Smith, K. S. Andrews, D. Brooks, S. A. Fedewa, D. Manassaram-Baptiste, D. Saslow, et al., Cancer screening in the United States, 2018: a review of current American Cancer Society guidelines and current issues in cancer screening, *CA Cancer J Clin* 68 (4) (2018) 297–316.
- [3] M. Greaves, C. C. Maley, Clonal evolution in cancer, *Nature* 481 (7381) (2012) 306–211.
- [4] Z. Hu, J. Ding, Z. Ma, R. Sun, J. A. Seoane, J. S. Shaffer, et al., Quantitative evidence for early metastatic seeding in colorectal cancer, *Nat Genet* 51 (7) (2019) 1113–1122.
- [5] M. J. Williams, B. Werner, T. Heide, C. Curtis, C. P. Barnes, A. Sottoriva, T. A. Graham, Quantification of subclonal selection in cancer from bulk sequencing data, *Nat Genet* 50 (6) (2018) 895–903.
- [6] J. Zhang, J. J. Cunningham, J. S. Brown, R. A. Gatenby, Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer, *Nat Commun* 8 (1) (2017) 1–9.
- [7] D. S. Park, M. Robertson-Tessi, K. A. Luddy, P. K. Maini, M. B. Bonsall, R. A. Gatenby, A. R. Anderson, The Goldilocks window of personalized chemotherapy: Getting the immune response just right, *Cancer Res* 79 (20) (2019) 5302–5315.

- [8] R. C. Rockne, A. Hawkins-Daarud, K. R. Swanson, J. P. Sluka, J. A. Glazier, P. Macklin, et al., The 2019 mathematical oncology roadmap, *Phys Biol* 16 (4) (2019) 041005.
- [9] K. Curtius, N. A. Wright, T. A. Graham, An evolutionary perspective on field cancerization, *Nat Rev Cancer* 18 (1) (2018) 19–32.
- [10] A. Sottoriva, H. Kang, Z. Ma, T. A. Graham, M. P. Salomon, J. Zhao, et al., A Big Bang model of human colorectal tumor growth, *Nature Genet* 47 (3) (2015) 209–216.
- [11] U. S. Preventive, K. Bibbins-Domingo, D. Grossman, S. Curry, K. Davidson, J. Epling, et al., Screening for colorectal cancer: US Preventive Services Task Force recommendation statement, *JAMA* 315 (2016) 2564–2575.
- [12] S. H. Moolgavkar, Commentary: Multistage carcinogenesis and epidemiological studies of cancer, *Int J Epidemiol* 45 (3) (2016) 645–649.
- [13] E. Luebeck, S. Moolgavkar, Multistage carcinogenesis and the incidence of colorectal cancer, *Proc Natl Acad Sci U S A* 99 (2002) 15095–15100.
- [14] T. N. Wenker, M. C. Tan, Y. Liu, H. B. El-Serag, A. P. Thrift, Prior diagnosis of Barrett’s esophagus is infrequent, but associated with improved esophageal adenocarcinoma survival, *Dig Dis Sci* 63 (11) (2018) 3112–3119.
- [15] G. Prasad, A. Bansal, P. Sharma, K. Wang, Predictors of progression in Barrett’s esophagus: current knowledge and future directions, *Am J Gastroenterol* 105 (2010) 1490–1502.
- [16] C. R. Heberle, A.-H. Omidvari, A. Ali, S. Kroep, C. Y. Kong, J. M. Inadomi, et al, Cost effectiveness of screening patients with gastroesophageal reflux disease for Barrett’s esophagus with a minimally invasive cell sampling device, *Clin Gastroenterol Hepatol* 15 (9) (2017) 1397–1404.
- [17] M. Kalager, M. Bretthauer, Improving cancer screening programs, *Science* 367 (6474) (2020) 143–144.
- [18] J. Jeon, R. Meza, S. Moolgavkar, G. Luebeck, Evaluation of screening strategies for pre-malignant lesions using a biomathematical approach, *Math Biosci* 213 (2008) 56–70.
- [19] L. Hanin, L. Pavlova, Optimal screening schedules for prevention of metastatic cancer, *Stat Med* 32 (2013) 206–219.
- [20] M. D. Ryser, W. T. Lee, N. E. Ready, K. Z. Leder, J. Foo, Quantifying the dynamics of field cancerization in tobacco-related head and neck cancer: a multi-scale modeling approach, *Cancer Res* 76 (24) (2016) 7078–7088.
- [21] A. Dhawan, T. A. Graham, A. G. Fletcher, A computational modeling approach for deriving biomarkers to predict cancer risk in premalignant disease, *Cancer Prev Res* 9 (4) (2016) 283–295.
- [22] B. M. Lang, J. Kuipers, B. Misselwitz, N. Beerenwinkel, Predicting colorectal cancer risk from adenoma detection via a two-type branching process model, *PLoS Comput Biol* 16 (2) (2020) e1007552
- [23] A. Dewanji, A. Biswas, An optimal design for simple illness-death model, *J Stat Plan Infer* 96 (2001) 289–300.

- [24] E. Arias, United States life tables, 2008., National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System 61 (3) (2012) 1–63.
- [25] K. Curtius, W. Hazelton, J. Jeon, E. Luebeck, A multiscale model evaluates screening for neoplasia in Barrett’s esophagus, PLoS Comput Biol 11 (5) (2015) e1004272.
- [26] K. Curtius, C.-J. Wong, W. D. Hazelton, A. M. Kaz, A. Chak, J. E. Willis, et. al, A molecular clock infers heterogeneous tissue age among patients with Barrett’s esophagus, PLoS Comput Biol 12 (5) (2016) e1004919.
- [27] C. Y. Kong, S. Kroep, K. Curtius, W. D. Hazelton, J. Jeon, R. Meza, et al, Exploring the recent trend in esophageal adenocarcinoma incidence and mortality using comparative simulation modeling, Cancer Epidemiol Biomarkers Prev 23 (6) (2014) 997–1006.
- [28] S. Kroep, C. R. Heberle, K. Curtius, C. Y. Kong, I. Lansdorp-Vogelaar, A. Ali, et al, Radiofrequency ablation of Barrett’s esophagus reduces esophageal adenocarcinoma incidence and mortality in a comparative modeling analysis, Clin Gastroenterol Hepatol 15 (9) (2017) 1471–1474.
- [29] A. F. Brouwer, M. C. Eisenberg, R. Meza, Case studies of gastric, lung, and oral cancer connect etiologic agent prevalence to cancer incidence, Cancer Res 78 (12) (2018) 3386–3396.
- [30] H. R. Moinova, T. LaFramboise, J. D. Lutterbaugh, A. K. Chandar, J. Dumot, A. Faulx, et al, Identifying DNA methylation biomarkers for non-endoscopic detection of Barrett’s esophagus, Sci Transl Med 10 (424) (2018) eaao5848.
- [31] A. Patel, C. P. Gyawali, Screening for Barrett’s esophagus: Balancing clinical value and cost-effectiveness, J Neurogastroenterol Motil 25 (2) (2019) 181–188.
- [32] N. J. Shaheen, G. W. Falk, P. G. Iyer, L. B. Gerson, ACG clinical guideline: Diagnosis and management of Barrett’s esophagus, Am J Gastroenterol 111 (2015) 30–50.
- [33] R. C. Fitzgerald, M. di Pietro, K. Ragunath, Y. Ang, J.-Y. Kang, P. Watson, et al., British Society of Gastroenterology guidelines on the diagnosis and management of Barrett’s oesophagus, Gut 63 (1) (2014) 7–42.
- [34] J. H. Rubenstein, N. Mattek, G. Eisen, Age- and sex-specific yield of Barrett’s esophagus by endoscopy indication, Gastrointest Endosc 71 (1) (2010) 21 – 27.
- [35] S. Rodriguez, N. Mattek, D. Lieberman, B. Fennerty, G. Eisen, Barrett’s esophagus on repeat endoscopy: should we look more than once?, Am J Gastroenterol 103 (8) (2008) 1892–1897.
- [36] R. Krishnamoorthi, G. P. Ramos, N. Crews, M. Johnson, R. Dierkhising, Q. Shi, et al., Persistence of nondysplastic Barrett’s esophagus is not protective against progression to adenocarcinoma, Clin Gastroenterol Hepatol 15 (6) (2017) 950–952.
- [37] M. van Putten, B. T. Johnston, L. J. Murray, A. T. Gavin, D. T. McManus, S. Bhat, et. al, ‘Missed’ oesophageal adenocarcinoma and high-grade dysplasia in Barrett’s oesophagus patients: A large population-based study, United European Gastroenterol J 6 (4) (2018) 519–528.
- [38] E. Luebeck, K. Curtius, J. Jeon, W. Hazelton, Impact of tumor progression on cancer incidence curves, Cancer Res 73 (3) (2013) 1086–1096.

- [39] W. D. Hazelton, K. Curtius, J. M. Inadomi, T. L. Vaughan, R. Meza, J. H. Rubenstein, et al, The role of gastroesophageal reflux and other factors during progression to esophageal adenocarcinoma, *Cancer Epidemiol Biomarkers Prev* (2015) 1012–1023.
- [40] H. J. de Koning, R. Meza, S. K. Plevritis, K. Ten Haaf, V. N. Munshi, J. Jeon, et al, Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the US Preventive Services Task Force, *Ann Intern* 160 (5) (2014) 311–320.
- [41] M. D. Ryser, M. Worni, E. L. Turner, J. R. Marks, R. Durrett, E. S. Hwang, Outcomes of active surveillance for ductal carcinoma in situ: a computational risk analysis, *J Natl Cancer Inst* 108 (5) (2016) djv372.
- [42] P. M. Altrock, J. Ferlic, T. Galla, M. H. Tomasson, F. Michor, Computational model of progression to multiple myeloma identifies optimum screening strategies, *JCO Clin Cancer Inform* (2) (2018) 1–12.
- [43] S. S. Hori, A. M. Lutz, R. Paulmurugan, S. S. Gambhir, A model-based personalized cancer screening strategy for detecting early-stage tumors using blood-borne biomarkers, *Cancer Res* (2017) 2570–2584.
- [44] K. Lahouel, L. Younes, L. Danilova, F. M. Giardiello, R. H. Hruban, J. Groopman, et al, Revisiting the tumorigenesis timeline with a data-driven generative model, *roc Natl Acad Sci U S A* 117 (2) (2020) 857–864.
- [45] T. J. Mitchell, S. Turajlic, A. Rowan, D. Nicol, J. H. Farmery, T. O’Brien, et al., Timing the landmark events in the evolution of clear cell renal cell cancer: Tracerx renal, *Cell* 173 (3) (2018) 611–623.
- [46] G. Parmigiani, S. Skates, M. Zelen, Modeling and optimization in early detection programs with a single exam, *Biometrics* 58 (1) (2002) 30–36.
- [47] A. B. Knudsen, A. G. Zauber, C. M. Rutter, S. K. Naber, V. P. Doria-Rose, C. Pabiniak, C. Johanson, S. E. Fischer, I. Lansdorp-Vogelaar, K. M. Kuntz, Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the US Preventive Services Task Force, *JAMA* 315 (23) (2016) 2595–2609.
- [48] J. M. Inadomi, N. Saxena, Screening and Surveillance for Barrett’s Esophagus: Is It Cost-Effective? *Dig Dis Sci*, 63 (8) (2019) 2094–2104.
- [49] T. Benaglia, L. D. Sharples, R. C. Fitzgerald, G. Lyratzopoulos, Health benefits and cost effectiveness of endoscopic and nonendoscopic cytosponge screening for Barrett’s esophagus. *Gastroenterol* 144 (1) (2013) 62–73.
- [50] B. Qumseya, A. Bukannan, S. Gendy, Y. Ahemd, S. Sultan, P. Bain, et al, Systematic review and meta-analysis of prevalence and risk factors for Barrett’s esophagus, *Gastrointest Endosc* 90 (5) (2019) 707–717.
- [51] J. H. Rubenstein, H. Morgenstern, H. Appelman, J. Scheiman, P. Schoenfeld, L. F. McMahon Jr, et al, Prediction of Barrett’s esophagus among men, *Am J Gastroenterol* 108 (3) (2013) 353–362.
- [52] D. K. Rex, O. W. Cummings, M. Shaw, M. D. Cumings, R. K. Wong, R. S. D. Dunne, et al, Screening for Barrett’s esophagus in colonoscopy patients with and without heartburn, *Gastroenterol* 125 (6) (2003) 1670–1677.

- 677 [53] J. Ronkainen, P. Aro, T. Storskrubb, S. E. Johansson, T. Lind, E. Bolling-Sternevald, et al, Prevalence
678 of Barrett's esophagus in the general population: an endoscopic study, *Gastroenterol* 129(6) (2005)
679 825–831.
- 680 [54] L. B. Gerson, S. Banerjee, Screening for Barrett's esophagus in asymptomatic women, *Gastrointestinal*
681 *Endosc*, 70(5) (2009) 867–873.
- 682 [55] J. M. Weaver, C. S. Ross-Innes, N. Shannon, A. G. Lynch, T. Forshew, M. Barbera, , et al, Ordering of
683 mutations in preinvasive disease stages of esophageal carcinogenesis, *Nature Genet* 46 (8) (2014) 837–843.
- 684 [56] J. D. Campbell, S. A. Mazzilli, M. E. Reid, S. S. Dhillon, S. Platero, J. Beane, et al, The case for a
685 Pre-Cancer Genome Atlas (PCGA), *Cancer Prev Res* 9 (2) (2016) 119–124.
- 686 [57] M. Yu, W. D. Hazelton, G. E. Luebeck, W. M. Grady, Epigenetic aging: more than just a clock when
687 it comes to cancer, *Cancer Res* 80 (3) (2019) 367–374.
- 688 [58] G. Chen, Y. Liu, D. Shen, K. MR, Composite large margin classifiers with latent subclasses for hetero-
689 geneous biomedical data, *Stat Anal Data Min* 9 (2016) 75–88.
- 690 [59] S. Moolgavkar, A. Dewanji, D. Venzon, A stochastic two-stage model for cancer risk assessment. ii. the
691 number and size of premalignant clones. the hazard function and the probability of tumor, *Risk Anal*
692 8 (3) (1988) 383–392.
- 693 [60] M. Little, Are two mutations sufficient to cause cancer? Some generalizations of the two-mutation model
694 of carcinogenesis of Moolgavkar, Venzon, and Knudson, and of the multistage model of Armitage and
695 Doll, *Biometrics* 51 (1995) 1278–1291.
- 696 [61] C. Crump, R. Subramaniam, C. Van Landingham, A numerical solution to the nonhomogeneous two-
697 stage MVK model of cancer., *Risk Anal* 25 (2005) 921–926.
- 698 [62] S. Karlin, H. E. Taylor, A second course in stochastic processes, Elsevier, 1981.
- 699 [63] E. Parzen, Stochastic Processes, SIAM, Philadelphia, PA, 1999.

700 **Supplementary Figures**

Figure S1: Optimizing screening and surveillance using the multistage clonal expansion model for EAC (MSCE-EAC) A) The goal of early detection and prevention is to perform an initial screen on individuals within a larger population after premalignant changes reside in cells but before incidental cancer occurs in order to intervene. B) After an initial screen at time t_{s1} , most surveillance recommendations are “one-size-fits-all” based solely on a population average (green) of a single variable- the diagnosis at initial screen (such as detection of Barrett’s esophagus (BE)). More refined adaptive recommendations for a follow-up surveillance at patient-specific time t_{s2} can iteratively account for the heterogeneity in a screened population (i.e., other factors that contribute to a patient being low risk (blue) versus high risk (red)). C) Within the MSCE-EAC multi-type branching process framework in a single patient’s esophagus, normal squamous epithelium may transform to BE cells X at a random onset time with rate $\nu(t)$, followed by a ‘two-hit’ tumor initiation process with Poisson rates μ_0 (creating P^* progeny), and μ_1 , which leads to the stochastic appearance of premalignant progenitor P cells. Premalignant cells undergo a first clonal expansion described by a birth-death-mutation process with cell division rate α_P , cell death-or-differentiation rate β_P , and malignant transformation rate μ_2 . Malignant cells M , in turn, may undergo a second clonal expansion by a birth-death-detection process with cell division and death rates α_M and β_M , respectively. Clinical detection of M cells occurs through a size-based detection process with parameter ρ . D) An illustration of MSCE-model stochastic trajectories. TSG, tumor suppressor gene; -/+, inactivation of one wild-type allele; -/-, inactivation of both wild-type alleles, e.g., loss/mutation of *TP53* in BE.

720

721 **Figure S2: Sensitivity analysis for optimal screening ages for birth cohorts 1960 - 2000.** We
722 found that the optimal ages for men of both risk strata remain unchanged from the base case results (birth
723 cohort 1950, Figure 4B) with increasing birth cohort. Specifically, for men with GERD we found that
724 optimal screening age was $t_s^* = 58$ and for all men combined was $t_s^* = 64$ for all 5 additional birth cohorts
725 tested. For women, we found slight differences due to the flatter optimality criterion computed: optimal
726 initial screening ages were $t_s^* = \{62, 61, 60, 59, 59\}$ for women with GERD, and $t_s^* = \{69, 68, 69, 67, 68\}$ for all
727 women combined, for birth cohorts 1960, 1970, 1980, 1990, and 2000, respectively, denoted by color.

728 Supplementary Tables

Table S1: MSCE-EAC model biological parameters (rates per cell per year).

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)

*Variables denoted with g are parameters of sigmoidal functions specific to a birth cohort and gender to obtain premalignant and malignant clonal proliferation rates, respectively, so that solving $g_P = \alpha_P - \beta_P - \mu_2$ and $g_M = \alpha_M - \beta_M - \rho$ obtains the parameters in Fig S1 (see [25] for explicit details).

729 Supplementary Methods

730 1. Multistage clonal expansion (MSCE) model survival functions

We first introduce the notation for the following random variables of the multi-type branching process

- $BE(t)$ = Bernoulli random variable for BE conversion by time t
- $X(t)$ = number of BE stem cells in a tissue at time t
- $P^*(t)$ = number of pre-initiated cells at time t
- $P(t)$ = number of premalignant (initiated) cells at time t
- $M(t)$ = number of malignant (preclinical) cells (prior to detection) at time t
- $C(t)$ = number of cancer cells (after detection) at time t
- $D(t)$ = Bernoulli random variable for clinical detection by time t

Let us consider the probability generating function (pgf) Ψ for the entire process starting at $\tau = 0$, ie. when an individual is born

$$\Psi(y_{BE}, y_1, y_2, y_3, z; t) = \sum_{i,j,k,l,n} y_{BE}^i y_1^j y_2^k y_3^l z^n P(i, j, k, l, n; t),$$

$$P(i, j, k, l, n; t) = \Pr[BE(t) = i, P^*(t) = j, P(t) = k, M(t) = l, D(t) = n | BE(0) = 0, P^*(0) = 0, P(0) = 0, M(0) = 0, D(0) = 0]$$

where, explicitly, $i, n = \{0, 1\}$ and $BE(t), D(t)$ are the following indicator functions corresponding to BE conversion and EAC clinical detection, respectively

$$BE(t) = \begin{cases} 0 & \text{if BE has not developed by time } t \\ 1 & \text{if BE conversion has taken place by time } t \end{cases}$$

$$D(t) = \begin{cases} 0 & \text{if no cancer detected clinically by time } t \\ 1 & \text{if a malignant cell is detected by time } t. \text{ ie } C(\tau) > 0 \text{ for some } \tau \leq t \end{cases}$$

The Chapman-Kolmogorov equations governing the transition probabilities for this multistage process include contributions from the initial Armitage-Doll type transition to BE, the two Poisson transitions to initiation, and the two birth-death-migration processes, all of which have been derived previously [13, 59, 60]. We begin with a method for solving for these generating functions using the Kolmogorov backward equations. Beginning with an active BE segment (BE), a single pre-initiated (P^*), premalignant (P), or malignant (M) cell at time τ only, we define the following generating functions $\Phi_{BE}, \Phi_{P^*}, \Phi_P$, or Φ_M , respectively,

$$\begin{aligned} \Phi_M(y_3, z; \tau, t) &= E[y_3^{M(t)} z^{D(t)} | M(\tau) = 1, D(\tau) = 0] \\ &= \sum_{k,l} y_3^k z^l \Pr[M(t) = k, D(t) = l | M(\tau) = 1, D(\tau) = 0] \end{aligned} \quad (15)$$

$$\begin{aligned} \Phi_P(y_2, y_3, z; \tau, t) &= E[y_2^{P(t)} y_3^{M(t)} z^{D(t)} | P(\tau) = 1, M(\tau) = 0, D(\tau) = 0] \\ &= \sum_{j,k,l} y_2^j y_3^k z^l \Pr[P(t) = j, M(t) = k, D(t) = l | P(\tau) = 1, M(\tau) = 0, D(\tau) = 0] \end{aligned} \quad (16)$$

$$\begin{aligned} \Phi_{P^*}(y_1, y_2, y_3, z; \tau, t) &= E[y_1^{P^*(t)} y_2^{P(t)} y_3^{M(t)} z^{D(t)} | P^*(\tau) = 1, P(\tau) = 0, M(\tau) = 0, D(\tau) = 0] \\ &= \sum_{i,j,k,l} y_1^i y_2^j y_3^k z^l \Pr[P^*(t) = i, P(t) = j, M(t) = k, D(t) = l | P^*(\tau) = 1, P(\tau) = 0, M(\tau) = 0, D(\tau) = 0] \end{aligned} \quad (17)$$

$$\Phi_{BE}(y_{BE}, y_1, y_2, y_3, z; \tau, t) = E[y_{BE}^{BE(t)} y_1^{P^*(t)} y_2^{P(t)} y_3^{M(t)} z^{D(t)} | BE(\tau) = 1, P^*(\tau) = 0, P(\tau) = 0, M(\tau) = 0, D(\tau) = 0] \quad (18)$$

731

$$= \sum_{i,j,k,l,n} y_{BE}^i y_1^j y_2^k y_3^l z^n \Pr[BE(t) = i, P^*(t) = j, P(t) = k, M(t) = l, D(t) = n | BE(\tau) = 1, P^*(\tau) = 0, P(\tau) = 0, M(\tau) = 0, D(\tau) = 0]$$

The generating functions satisfy the following Kolmogorov backward equations

$$\frac{\partial \Phi_M(y_3, z; \tau, t)}{\partial \tau} = -\alpha_M \Phi_M^2(y_3, z; \tau, t) - \beta_M \quad (19)$$

$$\begin{aligned} & -z\rho\Phi_M(y_3, z; \tau, t) + [\alpha_M + \beta_M + \rho]\Phi_M(y_3, z; \tau, t) \\ \frac{\partial \Phi_P(y_2, y_3, z; \tau, t)}{\partial \tau} &= -\alpha_P \Phi_P^2(y_2, y_3, z; \tau, t) - \beta_P \quad (20) \\ & + [\alpha_P + \beta_P + \mu_2]\Phi_P(y_2, y_3, z; \tau, t) - \mu_2\Phi_P(y_2, y_3, z; \tau, t)\Phi_M(y_3, z; \tau, t) \end{aligned}$$

$$\frac{\partial \Phi_{P^*}(y_1, y_2, y_3, z; \tau, t)}{\partial \tau} = -\mu_1\Phi_{P^*}(y_1, y_2, y_3, z; \tau, t)[\Phi_P(y_2, y_3, z; \tau, t) - 1] \quad (21)$$

$$\frac{\partial \Phi_{BE}(y_{BE}, y_1, y_2, y_3, z; \tau, t)}{\partial \tau} = -\mu_0 X \Phi_{BE}(y_{BE}, y_1, y_2, y_3, z; \tau, t)[\Phi_{P^*}(y_1, y_2, y_3, z; \tau, t) - 1] \quad (22)$$

$$\frac{\partial \Psi(y_{BE}, y_1, y_2, y_3, z; \tau, t)}{\partial \tau} = \nu(\tau)[\Psi(y_{BE}, y_1, y_2, y_3, z; \tau, t) - \Phi_{BE}(y_{BE}, y_1, y_2, y_3, z; \tau, t)] \quad (23)$$

To connect the cellular level description to the population level, we first solve for the overall survival function (for EAC cancer detection), starting at time 0, which in our notation is

$$\begin{aligned} S_{EAC}(t) &= 1 - P_{EAC}(t) = \Pr[D(t) = 0 | BE(0) = 0, P^*(0) = 0, P(0) = 0, M(0) = 0, D(0) = 0] \\ &= \Psi(1, 1, 1, 1, 0; 0, t) \end{aligned}$$

732 where $P_{EAC}(t)$ is the probability of a cancer detection at time t ,

$$P_{EAC}(t) = \Pr[D(t) = 1 | BE(0) = 0, P^*(0) = 0, P(0) = 0, M(0) = 0, D(0) = 0]$$

733 We will here denote $\Phi_M(1, 0; \tau, t) \equiv \Phi_M(\tau, t)$, $\Phi_P(1, 1, 0; \tau, t) \equiv \Phi_P(\tau, t)$, $\Phi_{P^*}(1, 1, 1, 0; \tau, t) \equiv \Phi_{P^*}(\tau, t)$,
734 $\Phi_{BE}(1, 1, 1, 1, 0; \tau, t) \equiv \Phi_{BE}(\tau, t)$, and $\Psi(1, 1, 1, 1, 0; \tau, t) \equiv \Psi(\tau, t)$. A dot designates a first derivative with
735 respect to t . The hazard function, i.e., the rate at which cancer is detected in individuals who have not been
736 diagnosed before, is given by

$$h_{EAC}(t) = -\frac{\dot{S}_{EAC}(t)}{S_{EAC}(t)} = -\frac{\dot{\Psi}(0, t)}{\Psi(0, t)} = -\frac{d}{dt} \ln[\Psi(0, t)] \quad (24)$$

For fixed t , this boundary value system of coupled PDEs can be converted into an initial value problem (IVP) with the change of variables $u = t - \tau$, where u is the “running” time. This redefinition and equations hereafter follow the method used by Crump et al. [61]. Define the following variables for the new IVP: $Y_1(u, t) = \Phi_M(\tau, t)$, $Y_2(u, t) = \dot{\Phi}_M(\tau, t)$, $Y_3(u, t) = \Phi_P(\tau, t)$, $Y_4(u, t) = \dot{\Phi}_P(\tau, t)$, $Y_5(u, t) = \Phi_{P^*}(\tau, t)$, $Y_6(u, t) = \dot{\Phi}_{P^*}(\tau, t)$, $Y_7(u, t) = \Phi_{BE}(\tau, t)$, $Y_8(u, t) = \dot{\Phi}_{BE}(\tau, t)$, $Y_9(u, t) = \Psi(\tau, t)$, $Y_{10}(u, t) = \dot{\Psi}(\tau, t)$ with corresponding initial conditions $Y_1(0, t) = Y_3(0, t) = Y_5(0, t) = Y_7(0, t) = Y_9(0, t) = 1$, $Y_4(0, t) = Y_6(0, t) = Y_8(0, t) =$

$Y_{10}(0, t) = 0$, and $Y_2(0, t) = -\rho$. Then the equations to solve for our IVP are the following

$$\frac{dY_1(u, t)}{du} = \beta_M - (\alpha_M + \beta_M + \rho)Y_1(u, t) + \alpha_M Y_1^2(u, t) \quad (25)$$

$$\frac{dY_2(u, t)}{du} = 2\alpha_M Y_1(u, t)Y_2(u, t) - (\alpha_M + \beta_M + \rho)Y_2(u, t) \quad (26)$$

$$\frac{dY_3(u, t)}{du} = \beta_P + \mu_2 Y_1(u, t)Y_3(u, t) - (\alpha_P + \beta_P + \mu_2)Y_3(u, t) + \alpha_P Y_3^2(u, t) \quad (27)$$

$$\frac{dY_4(u, t)}{du} = 2\alpha_P Y_3(u, t)Y_4(u, t) + \mu_2(Y_4(u, t)Y_1(u, t) + Y_3(u, t)Y_2(u, t)) - (\alpha_P + \beta_P + \mu_2)Y_4(u, t) \quad (28)$$

$$\frac{dY_5(u, t)}{du} = \mu_1 Y_5(u, t)(Y_3(u, t) - 1) \quad (29)$$

$$\frac{dY_6(u, t)}{du} = \mu_1(Y_6(u, t)Y_3(u, t) - Y_6(u, t) + Y_5(u, t)Y_4(u, t)) \quad (30)$$

$$\frac{dY_7(u, t)}{du} = \mu_0 X Y_7(u, t)(Y_5(u, t) - 1) \quad (31)$$

$$\frac{dY_8(u, t)}{du} = \mu_0 X(Y_8(u, t)Y_5(u, t) - Y_8(u, t) + Y_7(u, t)Y_6(u, t)) \quad (32)$$

$$\frac{dY_9(u, t)}{du} = \nu(u)(Y_7(u, t) - Y_9(u, t)) \quad (33)$$

$$\frac{dY_{10}(u, t)}{du} = \nu(u)(Y_{10}(u, t) - Y_8(u, t)). \quad (34)$$

These 10 coupled ODEs can be solved numerically to obtain the desired survival functions,

$$S_{MSE}(t) = Y_7(t, t), \quad S_{EAC}(t) = Y_9(t, t). \quad (35)$$

In the Results, we assigned binary detection variables the value 1 if a clone born at time s is non-extinct at time t_s (size of clone is ≥ 1) and 0 otherwise. In the case of EAC, this “perfect sensitivity” definition will be useful as progressively better high-resolution imaging technologies and minimally invasive sampling devices are tested in trials in BE [16, 30] and used in future clinical practice. For detectability based on this non-extinction criterion at time of screening, we adopt the probability of detection with a binary outcome variable for malignant clones with an unconditional birth-death-mutation size distribution (see Jeon et al. [18]). The forms of these are given by,

$$p_1(s, t_s) = 1 - \beta_M \zeta_M(s, t_s) \quad (36)$$

where

$$\zeta_M(s, t_s) = \frac{e^{(\alpha_M - \beta_M)(t_s - s)} - 1}{\alpha_M e^{(\alpha_M - \beta_M)(t_s - s)} - \beta_M}.$$

As $t_s \rightarrow \infty$, $p_1(s, t_s) \rightarrow 1 - \beta_M / \alpha_M$. We obtain the analogous definition for $p_2(s, t_s)$ with P parameters for the detection probability of premalignant clones -

$$p_2(s, t_s) = 1 - \beta_P \zeta_P(s, t_s) \quad (37)$$

748 where

$$\zeta_P(s, t_s) = \frac{e^{(\alpha_P - \beta_P)(t_s - s)} - 1}{\alpha_P e^{(\alpha_P - \beta_P)(t_s - s)} - \beta_P}.$$

749 As $t_s \rightarrow \infty$, $p_2(s, t_s) \rightarrow 1 - \beta_P/\alpha_P$, which is the asymptotic probability of non-extinction.

750 2. Initial screen strategy 2

We aim to derive the optimal screening age t_s^* (Eq. (8) in Main Text)

$$\begin{aligned} t_s^* &= \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot (\Pr[T_A \leq t_s] - w \cdot \Pr[T_P^d(t_s) = 1]) \\ &= \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot \left(\Pr[T_A \leq t_s] - w \cdot \int_0^{t_s} \Pr[T_P^d(t_s) = 1 | T_A = \tau] f_A(\tau) d\tau \right). \end{aligned}$$

Applying useful theory in multi-type branching processes for multistage models, we utilize the filtered Poisson process (FPP) approach [62, 63] to analytically solve for $\Pr[T_P^d(t_s) = 1 | T_A = \tau]$. Let $N(\sigma, t_s)$ be the number of initiated P cells from a P^* cell born at time σ that result in a detectable HGD clone at time t_s

$$N(\sigma, t_s) = \sum_l N_l(\sigma, s_l, t_s),$$

where $N_l(\sigma, t_s, s_l) = 1$ if initiation at time s_l produces a detectable clone at time t_s , and $s_l \leq t_s$, with probability $p_2(s_l, t_s) = \Pr[T_P^d(t_s) = 1 | P(s_l) = 1]$. Then $N(\sigma, t_s)$ is a FPP such that

$$N(\sigma, t_s) \sim \text{Poisson} \left(\int_{\sigma}^{t_s} \mu_1 p_2(s, t_s) ds \right),$$

where $p_2(s_l, t_s)$ is defined in Eq. (37). Next let $M(\tau, t_s)$ be the number of preinitiated cells after BE onset at time τ that result in a detectable P clone at time t_s . Then we have

$$M(\tau, t_s) = \sum_i M_i(\tau, \sigma_i, t_s),$$

where $M_i(\tau, \sigma_i, t_s)$ is binary and is equal to 1 if the preinitiated cell originating at time σ_i after BE onset at time τ results in a detectable P clone at time t_s , with $\tau \leq \sigma_i \leq t_s$. Note that $M_i(\tau, \sigma_i, t_s) = 1$ is equivalent to $N(\sigma_i, t_s) \geq 1$ which has the probability $p_3(\sigma_i, t_s) = 1 - \exp \left[- \int_{\sigma_i}^{t_s} \mu_1 p_2(s, t_s) ds \right]$ with $p_2(s, t_s)$ given by Eq. (37). Then $M(\tau, t_s)$ is a FPP such that

$$M(\tau, t_s) \sim \text{Poisson} \left(\int_{\tau}^{t_s} \mu_0 X p_3(\sigma, t_s) d\sigma \right)$$

Thus

$$\begin{aligned} \Pr[T_P^d(t_s) = 1 | T_A = \tau] &= 1 - \Pr[T_P^d(t_s) = 0 | T_A = \tau] \\ &= 1 - \exp \left[- \int_{\tau}^{t_s} \mu_0 X \left(1 - \exp \left[- \int_{\sigma}^{t_s} \mu_1 p_2(s, t_s) ds \right] \right) d\sigma \right]. \end{aligned} \quad (38)$$

751

Thus together with $\mathcal{A}$ onset distribution in Eq. (5) and mortality risk, we have the analytical solution for optimal screening age in Strategy 2, given in Eq. (8).

3. Initial screen strategy 3

Here we derive the analytical solution for the optimal screening time for Strategy 3, given some weighting w to avoid risk of co-existing malignant clones (Eq. (9) in Main Text),

$$t_s^* = \arg \max_{t_s} \Pr[T_D \geq t_s] \cdot \left(\int_0^{t_s} \Pr[T_P^d(t_s) = 1 | T_A = \tau] \cdot f_A(\tau) d\tau - w \cdot \int_0^{t_s} \Pr[T_P^d(t_s) = 1, T_M^d(t_s) = 1 | T_A = \tau] \cdot f_A(\tau) d\tau \right). \quad (39)$$

The first integral in this difference was derived in Strategy 2 (see Eq. (38)). We solve for the probability in the integrand of the second integral with detectability given in Eq. (36) and Eq. (37) for malignant clones and HGD clones, respectively. Thus we seek an explicit formula for the integrand,

$$\Pr[T_P^d(t_s) = 1, T_M^d(t_s) = 1 | T_{BE} = \tau]. \quad (40)$$

To derive the above probability, we must define C_1, C_2, C_3 , mutually exclusive scenarios of counts of initiations of different types. Scenario C_1 corresponds to the scenario that the P clone that is non-extinct at time t_s is the ancestor of the non-extinct M clone at time t_s . This is most commonly the case in the stochastic realizations since the bigger P clones that remain non-extinct at time of screening have the highest probability to have had μ_2 events (malignant transformation) occur since their initiation. Scenario C_2 accounts for a non-extinct P clone at time t_s that, if it created any malignant progeny before time of screening, those malignancies went extinct. Scenario C_3 accounts for a non-extinct M clone at time t_s whose ancestral P clone went extinct before time t_s . Thus the probability we aim to derive accounts for either event C_1 or C_2 and C_3 together. For each of the three scenarios, we will use filtered Poisson processes techniques and abuse notation slightly by redefining FPP random variables N_1 and N_2 differently but analogously for each scenario, to avoid verbosity. We start with deriving the probability that event C_1 occurs (at least once) with $u > s$,

$$C_1 : \quad \text{Let } N_1(u, t_s, s) = 1 \text{ if } T_P^d(t_s, s) = 1, T_M^d(u, t_s, s) = 1 | P(s) = 1, M(u) = 1 \quad (41)$$

with probability $p_2^d(s, t_s) \cdot p_1^d(u, t_s) = (1 - \beta_P \zeta_P(s, t_s)) \cdot (1 - \beta_M \zeta_M(u, t_s))$. Then we have that

$$N_1(t_s, s, \sigma) = \sum_u N_1(u, t_s, s, \sigma) \sim \text{Poisson} \left(p_2^d(s, t_s) \cdot \int_s^{t_s} \mu_2 p_1^d(u, t_s) du \right) \quad (42)$$

and

$$N_2(\sigma, t_s) = \sum_l N_1(t_s, s_l, \sigma) \quad (43)$$

so that $N_2(\sigma, t_s)$ is a FPP and the probability generating function (PGF) of $N_2(\sigma, t_s)$ is

$$\Psi(x; \sigma, t_s) = \exp \left[\int_\sigma^{t_s} \mu_1 \cdot \{ \Psi(x; s, t_s) - 1 \} ds \right] \quad (44)$$

where $\Psi(x; s, t_s)$ is the PGF of $N_1(t_s, s, \sigma) = e^{-\lambda(1-x)}$, where $\lambda = p_2^d(s, t_s) \cdot \int_s^{t_s} \mu_2 p_1^d(u, t_s) du$. Therefore

$$\Pr[C_1 \geq 1] = \Pr[N_2(\sigma, t_s) \geq 1] = 1 - \Pr[N_2(\sigma, t_s) = 0] = 1 - \Psi(x = 0, t_s) \quad (45)$$

$$= 1 - \exp \left[\int_\sigma^{t_s} \mu_1 \cdot \{e^{-\lambda} - 1\} ds \right] \quad (46)$$

Similarly,

$$C_2 : \quad \text{Let } N_1(u, t_s, s) = 1 \text{ if } T_P^d(t_s, s) = 1, T_M^d(u, t_s, s) = 0 | P(s) = 1, M(u) = 1 \quad (47)$$

with probability $p_2^d(s, t_s) \cdot (1 - p_1^d(u, t_s)) = (1 - \beta_P \zeta_P(s, t_s)) \cdot \beta_M \zeta_M(u, t_s)$. Then we have that

$$N_1(t_s, s, \sigma) = \sum_u N_2(u, t_s, s, \sigma) \sim \text{Poisson} \left(p_2^d(s, t_s) \cdot \int_s^{t_s} \mu_2 (1 - p_1^d(u, t_s)) du \right) \quad (48)$$

and

$$N_2(\sigma, t_s) = \sum_l N_1(t_s, s_l, \sigma) \quad (49)$$

so that $N_2(\sigma, t_s)$ is a FPP and the PGF of $N_2(\sigma, t_s)$ is

$$\Psi_2(x; \sigma, t_s) = \exp \left[\int_\sigma^{t_s} \mu_1 \cdot \{\Psi_2(x; s, t_s) - 1\} ds \right], \quad (50)$$

where $\Psi_2(x; s, t_s)$ is the PGF of $N_1(t_s, s, \sigma) = e^{-\lambda_2(1-x)}$, where $\lambda_2 = p_2^d(s, t_s) \cdot \int_s^{t_s} \mu_2 (1 - p_1^d(u, t_s)) du$. Therefore

$$\Pr[C_2 \geq 1] = \Pr[N_2(\sigma, t_{s_2}) \geq 1] = 1 - \Pr[N_2(\sigma, t_{s_2}) = 0] = 1 - \Psi_2(x = 0, t_{s_2}) \quad (51)$$

$$= 1 - \exp \left[\int_\sigma^{t_s} \mu_1 \cdot \{e^{-\lambda_2} - 1\} ds \right]. \quad (52)$$

Lastly, we define

$$C_3 : \quad \text{Let } N_1(u, t_s, s) = 1 \text{ if } T_P^d(t_s, s) = 0, T_M^d(u, t_s, s) = 1 | P(s) = 1, M(u) = 1 \quad (53)$$

with probability $p_2^d(s, u) \cdot (1 - p_2^d(u, t_s)) \cdot p_1^d(u, t_s) = (1 - \beta_P \zeta_P(s, u)) \cdot \beta_P \zeta_P(u, t_s) \cdot (1 - \beta_M \zeta_M(u, t_s))$. Then we have that

$$N_1(t_s, s, \sigma) = \sum_u N_1(u, t_s, s, \sigma) \sim \text{Poisson} \left(\int_s^{t_s} \mu_2 \cdot p_2(s, u) \cdot (1 - p_2^d(u, t_s)) \cdot p_1^d(u, t_s) du \right) \quad (54)$$

and

$$N_2(\sigma, t_s) = \sum_l N_1(t_s, s_l, \sigma), \quad (55)$$

so that $N_2(\sigma, t_s)$ is a FPP and the PGF of $N_2(\sigma, t_s)$ is

$$\Psi_3(x; \sigma, t_s) = \exp \left[\int_{\sigma}^{t_s} \mu_1 \cdot \{\Psi_3(x; s, t_s) - 1\} ds \right], \quad (56)$$

where $\Psi_3(x; s, t_s)$ is the PGF of $N_1(t_s, s, \sigma) = e^{-\lambda_3(1-x)}$, where $\lambda_3 = \int_s^{t_s} \mu_2 \cdot p_2(s, u) \cdot (1 - p_2^d(u, t_s)) \cdot p_1^d(u, t_s) du$. Therefore

$$\Pr[C_3 \geq 1] = \Pr[N_2(\sigma, t_s) \geq 1] = 1 - \Pr[N_2(\sigma, t_s) = 0] = \Psi_3(x = 0, t_s) \quad (57)$$

$$= 1 - \exp \left[\int_{\sigma}^{t_s} \mu_1 \cdot \{e^{-\lambda_3} - 1\} ds \right]. \quad (58)$$

Next, let $M(\tau, t_s)$ be the number of preinitiated cells after BE onset at time τ that result in $C_1 \cup (C_2 \cap C_3)$

$$M(\tau, t_s) = \sum_i M_i(\tau, \sigma_i, t_s), \quad (59)$$

where $M_i(\tau, \sigma_i, t_s) = 1 \iff C_1 \geq 1 \cup (C_2 \geq 1, C_3 \geq 1)$. Due to the mutual exclusivity of these events, we have that $M_i(\tau, \sigma_i, t_s) = 1$ with probability

$$p_3(\sigma_i, t_s) = \Pr[C_1 \geq 1 \cup (C_2 \geq 1 \cap C_3 \geq 1)] \quad (60)$$

$$= \Pr[C_1 \geq 1] + \Pr[C_2 \geq 1] \cdot \Pr[C_3 \geq 1] - \Pr[C_1 \geq 1] \cdot \Pr[C_2 \geq 1] \cdot \Pr[C_3 \geq 1] \quad (61)$$

Then $M(\tau, t_s)$ is a FPP such that

$$M(\tau, t_s) \sim \text{Poisson} \left(\int_{\tau}^{t_s} \mu_2 \cdot p_3(\sigma, t_s) d\sigma \right) \quad (62)$$

Finally, we solve for our integrand of interest from Eq. (40)

$$\Pr[T_M^d(t_s) = 1, T_P^d(t_s) = 1 | T_{BE} = \tau] \quad (63)$$

$$= 1 - \exp \left[- \int_{\tau}^{t_s} \mu_0 X \cdot p_3(\sigma, t_s) d\sigma \right]. \quad (64)$$

With this, we have analytically solved for all components of the optimality criterion for Strategy 3 from Eq. (39).

4. Adaptive surveillance and associated-risk

Continuing from the Main Text we derive the following for optimal time for surveillance given Outcome 1 (negative for precursor),

$$t_s^* = \arg \max_{t_s} \Pr[T_D \geq t_{s2} | T_D > t_{s1}] \cdot (\Pr[T_A \leq t_{s2} | T_A > t_{s1}] - w \cdot \Pr[T_C \leq t_{s2} | T_A > t_{s1}]) \quad (65)$$

$$= \arg \max_{t_s} \Pr[T_D \geq t_{s2} | T_D > t_{s1}] \cdot \left(\Pr[T_A \leq t_{s2} | T_A > t_{s1}] - w \cdot \frac{1}{\Pr[T_A > t_{s1}]} \int_{t_{s1}}^{t_{s2}} (1 - S_{MSE}(t_{s2} - \tau)) f_A(\tau) d\tau \right). \quad (66)$$

772 Then for Outcome 2 (positive for precursor) we defined a risk associated with a surveillance timing as,

$$R_{C|SR}(t_{s_2}) = \Pr[T_C \leq t_{s_2} | T_A = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0]. \quad (67)$$

Here we derive this risk associated with next surveillance time given a certain screening result, $R_{C|SR}(t_{s_2})$, as general lifetime risk without other cause mortality for use in optimization/cost-effectiveness analysis. Once again, for the scenario with a perfectly sensitive screen that can detect single cells, we may derive this probability using a filtered Poisson approach analogous to the approach used in Strategy 2 for a single screen. We redefine (for sake of retaining simpler notation) $N(\sigma, t_{s_2})$ to be the number of initiated P cells from a P^* cell born at time σ that result in a clinically detected EAC clone at time t_{s_2}

$$N(\sigma, t_{s_2}) = \sum_l N_l(\sigma, s_l, t_{s_2}),$$

where $N_l(\sigma, s_l, t_{s_2}) = 1$ if initiation at time s_l produces EAC at time t_{s_2} , and $t_{s_1} < s_l < t_{s_2}$, with probability $p_C(s_l, t_{s_2}) = \Pr[T_C < t_{s_2} | P(s_l) = 1]$. Then $N(\sigma, t_{s_2})$ is a FPP such that

$$N(\sigma, t_{s_2}) \sim \text{Poisson} \left(\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 p_C(s, t_{s_2}) ds \right).$$

Clinical EAC cumulative distribution, $p_C(s_l, t_{s_2})$, can be solved via the Kolmogorov equations of the MSCE-EAC branching process from the PDE in Eq. (20). Specifically we have

$$p_C(s, t_{s_2}) = \Pr[T_C < t_{s_2} | P(s) = 1] = 1 - \Pr[T_C \geq t_{s_2} | P(s) = 1] \quad (68)$$

$$= 1 - \Phi_P(1, 1, 0; s, t_{s_2}) = 1 - Y_3(s, t_{s_2}) \quad (69)$$

where $\Phi_P(1, 1, 0; s, t_{s_2})$ is the EAC survival from a single P cell (see Eq. (16)), $Y_3(u, t_{s_2})$ is the associated ODE (see Supplementary Methods 1) and $u = t_{s_2} - s$. Next let $M(\tau, t_{s_2})$ be the number of preinitiated cells after BE onset at time τ that result in detectable EAC at time t_{s_2} . Then we have

$$M(\tau, t_{s_2}) = \sum_i M_i(\tau, \sigma_i, t_{s_2}),$$

773 where $M_i(\tau, \sigma_i, t_{s_2})$ is, as before, binary and is equal to 1 if the preinitiated cell originating at time σ_i after
 774 BE onset at time τ results in a detectable EAC at time t_{s_2} , with $\tau < t_{s_1} < t_{s_2}$. Note that $M_i(\tau, \sigma_i, t_{s_2}) = 1$
 775 is equivalent to $N(\sigma_i, t_{s_2}) \geq 1$ which has the probability $p_3(\sigma_i, t_{s_2}) = 1 - \exp \left[- \int_{\max(\sigma_i, t_{s_1})}^{t_{s_2}} \mu_1 p_C(s, t_{s_2}) ds \right]$.

Then $M(\tau, t_{s_2})$ is a FPP such that

$$M(\tau, t_{s_2}) \sim \text{Poisson} \left(\int_{\tau}^{t_{s_2}} \mu_0 X p_3(\sigma, t_{s_2}) d\sigma \right).$$

776 Thus

$$R_{C|SR}(t_{s_2}) = \Pr[T_C \leq t_{s_2} | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \quad (70)$$

777

$$= 1 - \Pr[T_C > t_{s_2} | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \quad (71)$$

$$= 1 - \exp \left[- \int_{\tau}^{t_{s_2}} \mu_0 X \left(1 - \exp \left[- \int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 p_C(s, t_{s_2}) ds \right] \right) d\sigma \right]. \quad (72)$$

5. Adaptive screen outcome 2, strategy 1

As a final example of adaptive surveillance, an important clinical question to consider is, “Should we offer a one-time surveillance and when would that optimal time be?” Assuming we have an estimate τ of a patient’s precursor onset time, we solve for optimal next-surveillance age $t_{s_2}^*$,

$$t_{s_2}^* = \arg \max_{t_{s_2}} \Pr[T_M^d(t_{s_2}) = 0, T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \quad (73)$$

$$= \arg \max_{t_{s_2}} \Pr[T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \\ - w \cdot \Pr[T_M^d(t_{s_2}) = 1, T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0]. \quad (74)$$

Again, we can increase or decrease w depending on how much weight we want to place on keeping the risk of developing malignancy under a certain value, say ϵ . For ease of notation in this strategy, we do not include the other cause mortality term $\Pr[T_D \geq t_{s_2} | T_D > t_{s_1}]$ that can be multiplied to the entire expression of optimality criterion provided in Eq. (74) as we have done in previous strategies.

Let us first derive a formula for the first term of this difference based on the MSCE-EAC model construction using again a filtered Poisson process approach,

$$\Pr[T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0]. \quad (75)$$

Let $N(\sigma, t_{s_1}, t_{s_2})$ be the number of initiated cells from a P^* cell born at time σ that result in a detectable clone at time t_{s_2} , not born before t_{s_1}

$$N(\sigma, t_{s_1}, t_{s_2}) = \sum_l N_l(\sigma, t_{s_1}, t_{s_2}, s_l), \quad (76)$$

where $N_l(\sigma, t_{s_1}, t_{s_2}, s_l) = 1$ if initiation at time s_l produces a detectable clone at time t_{s_2} , $s_l > t_{s_1}$ with probability $p_2(s_l, t_{s_2}) = \Pr[T_P^d(t_{s_2}) = 1, T_P^d(t_{s_1}) = 0 | P(s_l) = 1] > 0$ if and only if $s_l > t_{s_1}$. Then $N(\sigma, t_{s_1}, t_{s_2})$ is a FPP such that

$$N(\sigma, t_{s_1}, t_{s_2}) \sim \text{Poisson} \left(\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 p_2(s, t_{s_2}) ds \right). \quad (77)$$

Assuming detectability of P cells based on non-extinction at time of screening, we have the following probability (see Eq. (37)),

$$p_2(s, t_{s_2}) = 1 - \beta_P \zeta_P(s, t_{s_2}), \quad t_{s_1} < s < t_{s_2} \quad (78)$$

where

$$\zeta_P(s, t_{s_2}) = \frac{e^{(\alpha_P - \beta_P)(t_{s_2} - s)} - 1}{\alpha_P e^{(\alpha_P - \beta_P)(t_{s_2} - s)} - \beta_P}. \quad (79)$$

The asymptotic solution for this probability is $p_2(s, t_{s_2}) = 1 - \beta_P / \alpha_P$. We will assume the same for malignant clones, with detection probability provided in Eq. (36).

Next let $M(\tau, t_{s_2})$ be the number of preinitiated cells after BE onset at time τ that result in a detectable P clone at time t_{s_2} . Then we have that

$$M(\tau, t_{s_2}) = \sum_i M_i(\tau, \sigma_i, t_{s_2}), \quad (80)$$

where $M_i(\tau, \sigma_i, t_{s_2})$ is, as before, binary and is equal to 1 if the preinitiated cell originating at time σ_i after BE onset at time τ results in a detectable EAC at time t_{s_2} , with $\tau < t_{s_1} < t_{s_2}$. Note that $M_i(\tau, \sigma_i, t_{s_2}) = 1$ is equivalent to $N(\sigma_i, t_{s_2}) \geq 1$ which has the probability $p_3(\sigma_i, t_{s_2}) = 1 - \exp \left[- \int_{\max(\sigma_i, t_{s_1})}^{t_{s_2}} p_2(s, t_{s_2}) ds \right]$.

Then $M(\tau, t_{s_2})$ is a FPP such that

$$M(\tau, t_{s_2}) \sim \text{Poisson} \left(\int_{\tau}^{t_{s_2}} \mu_0 X p_3(\sigma, t_{s_2}) d\sigma \right) \quad (81)$$

Thus

$$\Pr[T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \quad (82)$$

$$= 1 - \exp \left[- \int_{\tau}^{t_{s_2}} \mu_0 X \left(1 - \exp \left[- \int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 p_2(s, t_{s_2}) ds \right] \right) d\sigma \right]. \quad (83)$$

For the second term in Eq. (74), we will now derive

$$\Pr[T_M^d(t_{s_2}) = 1, T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \quad (84)$$

We define similar C_1, C_2, C_3 mutually exclusive scenarios of counts of initiations of different types as was utilized in initial screen Strategy 3. For each of the three scenarios, we will use filtered Poisson processes techniques and abuse notation slightly by redefining FPP random variables N_1 and N_2 variables differently but analogously for each scenario, to avoid verbosity. We start with deriving the probability that event C_1 occurs (at least once),

$$C_1 : \quad \text{Let } N_1(u, t_{s_2}, s) = 1 \text{ if } T_P^d(t_{s_2}, s) = 1, T_M^d(u, t_{s_2}, s) = 1 | P(s) = 1, M(u) = 1 \quad (85)$$

with probability $p_2^d(s, t_{s_2}) \cdot p_1^d(u, t_{s_2}) = (1 - \beta_P \zeta_P(s, t_{s_2})) \cdot (1 - \beta_M \zeta_M(u, t_{s_2}))$. Then we have that

$$N_1(t_{s_2}, s, \sigma) = \sum_u N_1(u, t_{s_2}, s, \sigma) \sim \text{Poisson} \left(p_2^d(s, t_{s_2}) \cdot \int_s^{t_{s_2}} \mu_2 p_1^d(u, t_{s_2}) du \right) \quad (86)$$

and

$$N_2(\sigma, t_{s_2}) = \sum_l N_1(t_{s_2}, s_l, \sigma) \quad (87)$$

so that $N_2(\sigma, t_{s_2})$ is a FPP and the PGF of $N_2(\sigma, t_{s_2})$ is

$$\Psi(x; \sigma, t_{s_1}, t_{s_2}) = \exp \left[\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 \cdot \{\Psi(x; s, t_{s_2}) - 1\} ds \right], \quad (88)$$

where $\Psi(x; s, t_{s_2})$ is the PGF of $N_1(t_{s_2}, s, \sigma) = e^{-\lambda(1-x)}$ where $\lambda = p_2^d(s, t_{s_2}) \cdot \int_s^{t_{s_2}} \mu_2 p_1^d(u, t_{s_2}) du$. Therefore

$$\Pr[C_1 \geq 1] = \Pr[N_2(\sigma, t_{s_2}) \geq 1] = 1 - \Pr[N_2(\sigma, t_{s_2}) = 0] = 1 - \Psi(x = 0, t_{s_2}) \quad (89)$$

$$= 1 - \exp \left[\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 \cdot \{e^{-\lambda} - 1\} ds \right] \quad (90)$$

Similarly,

$$C_2 : \quad \text{Let } N_1(u, t_{s_2}, s) = 1 \text{ if } T_P^d(t_{s_2}, s) = 1, T_M^d(u, t_{s_2}, s) = 0 | P(s) = 1, M(u) = 1 \quad (91)$$

with probability $p_2^d(s, t_{s_2}) \cdot (1 - p_1^d(u, t_{s_2})) = (1 - \beta_P \zeta_P(s, t_{s_2})) \cdot \beta_M \zeta_M(u, t_{s_2})$. Then we have that

$$N_1(t_{s_2}, s, \sigma) = \sum_u N_2(u, t_{s_2}, s, \sigma) \sim \text{Poisson} \left(p_2^d(s, t_{s_2}) \cdot \int_s^{t_{s_2}} \mu_2 (1 - p_1^d(u, t_{s_2})) du \right) \quad (92)$$

and

$$N_2(\sigma, t_{s_2}) = \sum_l N_1(t_{s_2}, s_l, \sigma) \quad (93)$$

so that $N_2(\sigma, t_{s_2})$ is a FPP and the PGF of $N_2(\sigma, t_{s_2})$ is

$$\Psi_2(x; \sigma, t_{s_1}, t_{s_2}) = \exp \left[\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 \cdot \{\Psi_2(x; s, t_{s_2}) - 1\} ds \right], \quad (94)$$

where $\Psi_2(x; s, t_{s_2})$ is the PGF of $N_1(t_{s_2}, s, \sigma) = e^{-\lambda_2(1-x)}$ where $\lambda_2 = p_2^d(s, t_{s_2}) \cdot \int_s^{t_{s_2}} \mu_2 (1 - p_1^d(u, t_{s_2})) du$. Therefore

$$\Pr[C_2 \geq 1] = \Pr[N_2(\sigma, t_{s_2}) \geq 1] = 1 - \Pr[N_2(\sigma, t_{s_2}) = 0] = 1 - \Psi_2(x = 0, t_{s_2}) \quad (95)$$

$$= 1 - \exp \left[\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 \cdot \{e^{-\lambda_2} - 1\} ds \right]. \quad (96)$$

Lastly, we define

$$C_3 : \quad \text{Let } N_1(u, t_{s_2}, s) = 1 \text{ if } T_P^d(t_{s_2}, s) = 0, T_M^d(u, t_{s_2}, s) = 1 | P(s) = 1, M(u) = 1 \quad (97)$$

with probability $p_2^d(s, u) \cdot (1 - p_2^d(u, t_{s_2})) \cdot p_1^d(u, t_{s_2}) = (1 - \beta_P \zeta_P(s, u)) \cdot \beta_P \zeta_P(u, t_{s_2}) \cdot (1 - \beta_M \zeta_M(u, t_{s_2}))$. Then we have that

$$N_1(t_{s_2}, s, \sigma) = \sum_u N_1(u, t_{s_2}, s, \sigma) \sim \text{Poisson} \left(\int_s^{t_{s_2}} \mu_2 \cdot p_2(s, u) \cdot (1 - p_2^d(u, t_{s_2})) \cdot p_1^d(u, t_{s_2}) du \right) \quad (98)$$

and

$$N_2(\sigma, t_{s_2}) = \sum_l N_1(t_{s_2}, s_l, \sigma) \quad (99)$$

so that $N_2(\sigma, t_{s_2})$ is a FPP and the PGF of $N_2(\sigma, t_{s_2})$ is

$$\Psi_3(x; \sigma, t_{s_1}, t_{s_2}) = \exp \left[\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 \cdot \{\Psi_3(x; s, t_{s_2}) - 1\} ds \right] \quad (100)$$

where $\Psi_3(x; s, t_{s_2})$ is the PGF of $N_1(t_{s_2}, s, \sigma) = e^{-\lambda_3(1-x)}$ where $\lambda_3 = \int_s^{t_{s_2}} \mu_2 \cdot p_2(s, u) \cdot (1 - p_2^d(u, t_{s_2})) \cdot p_1^d(u, t_{s_2}) du$. Therefore

$$\Pr[C_3 \geq 1] = \Pr[N_2(\sigma, t_{s_2}) \geq 1] = 1 - \Pr[N_2(\sigma, t_{s_2}) = 0] = \Psi_3(x = 0, t_{s_2}) \quad (101)$$

$$= 1 - \exp \left[\int_{\max(\sigma, t_{s_1})}^{t_{s_2}} \mu_1 \cdot \{e^{-\lambda_3} - 1\} ds \right]. \quad (102)$$

Next, let $M(\tau, t_{s_2})$ be the number of preinitiated cells after BE onset at time τ that result in $C_1 \cup (C_2 \cap C_3)$

$$M(\tau, t_{s_2}) = \sum_i M_i(\tau, \sigma_i, t_{s_2}) \quad (103)$$

where $M_i(\tau, \sigma_i, t_{s_2}) = 1 \iff C_1 \geq 1 \cup C_2 \geq 1, C_3 \geq 1$. Due to the mutual exclusivity of these events, we have that $M_i(\tau, \sigma_i, t_{s_2}) = 1$ with probability

$$p_3(\sigma_i, t_{s_2}) = \Pr[C_1 \geq 1 \cup (C_2 \geq 1 \cap C_3 \geq 1)] \quad (104)$$

$$= \Pr[C_1 \geq 1] + \Pr[C_2 \geq 1] \cdot \Pr[C_3 \geq 1] - \Pr[C_1 \geq 1] \cdot \Pr[C_2 \geq 1] \cdot \Pr[C_3 \geq 1] \quad (105)$$

Then $M(\tau, t_{s_2})$ is a FPP such that

$$M(\tau, t_{s_2}) \sim \text{Poisson} \left(\int_{\tau}^{t_{s_2}} \mu_2 \cdot p_3(\sigma, t_{s_2}) d\sigma \right) \quad (106)$$

With these, we have analytically solved for our expression of interest from Eq. (84),

$$\Pr[T_M^d(t_{s_2}) = 1, T_P^d(t_{s_2}) = 1 | T_{BE} = \tau, T_P^d(t_{s_1}) = 0, T_M^d(t_{s_1}) = 0] \quad (107)$$

$$= 1 - \exp \left[- \int_{\tau}^{t_{s_2}} \mu_0 X \cdot p_3(\sigma, t_{s_2}) d\sigma \right], \quad (108)$$

and thus have both components to compute the optimality criterion for Outcome 2: Strategy 1 from Eq. (74).