Recurrence-time model adjusted for time-varying incidence and sensitivity

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Sanda Cristina Oancea

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Prof. Dr. Timothy R. Church

July 2011

Acknowledgements

I would first like to thank my advisor Professor Dr. Timothy R. Church, for his wisdom, great knowledge, patience and understanding throughout the challenging process of writing this thesis. His belief in me and positive words were the energy that pushed me to challenge myself and become a better researcher and person every day. Second, I would like to thank Professor Dr. Chap Le, whose example of life tenacity and overcoming huge obstacles motivated me to keep going even when I felt that it was too tough at times. Third I would like to thank Professor Dr. Bruce Alexander and Associate Professor Dr. George Maldonado for imparting their knowledge and guiding me throughout my studies. Fourth I would like to thank Associate Professor Dr. Bonnie Westra, Dr. Martin Oken and all those in the School of Nursing and in the Health Studies Section whose work inspired this thesis.

Finally I would like to thank my family who always encouraged me and believed in me.

Dedication

I dedicate my thesis to all those who never let me fall, motivated me and were bringing many rays of sunshine in my life.

Abstract

Screening is being widely implemented as a public health measure for many progressive diseases such as colon cancer, lung cancer and prostate and ovarian cancer, just to name a few.

For programs of screening for cancer, a reduction in disease-specific mortality in the screened group is the main proof of benefit. There has been a long existing debate as to how to deal with biases that occur during the process of data analysis, such as lead-time bias, length bias and overdiagnosis. The focus of this work was initially to estimate and account for lead-time bias. In developing the model, it expanded to include estimated time-dependent sensitivity and age-specific preclinical incidence.

To avoid biases, some investigators have used randomized trials of cancer screening, and compared mortality rates from the cancer under investigation between the screened and the control group, instead of comparing survival rates. Since a lot of data are collected in screening programs done outside the randomized trials, this work will focus on screening programs in “one-arm” studies, those for which randomization either does not exist or may be unethical.

After preliminary discussions in Chapter I regarding screening models, including the concepts of sensitivity, specificity, positive and negative predictive values, and lead time,

Chapter II presents an extension of recurrence-time models of periodic screening to accommodate time-varying incidence and sensitivity. Chapter III presents an application of the developed model to colon cancer screening data and Chapter IV is a summary and a conclusion of the entire work.

Table of Contents

List of Tables

List of Main Notations

List of Figures

Chapter I: Preliminaries

I.1 Introduction: Components of Screening for Disease

In 1951 the United States Commission of Chronic Illness defined screening as ‘‘the presumptive identification of unrecognized disease or defect by the application of tests, examinations, or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment’’ (CCI, 1957).

According to the definition of the US National Center for Health Statistics (http://www.cdc.gov/nchs/), a chronic disease is one that lasts 3 months or more. Chronic diseases tend to become more common with age. In developed countries the leading chronic diseases include, among others, cancer such as cervical, breast and colorectal. Screening for chronic diseases in asymptomatic individuals can provide public health benefits if treatment delivered in earlier disease is more effective than in later stages.

Screening tests can be divided according to the type of abnormality they aim to detect (Cuzick, 1999). The goal of many tests is early detection of the cancer itself. Such tests can never reduce the incidence of disease, but do hope to reduce the mortality and

morbidity associated with advanced disease. For example, women over the age of 50 years who have been screened by mammography have a clear reduction in breast cancer mortality and this is due to the fact that, at least in some women, early detection permits the surgical removal of the lump before it has metastasized, so that the chances of cure are high (Lee C.H., 2002). But there are also limitations of screening tests aimed at early detection, since some breast cancers metastasize at a very early stage, before they are detectable by mammography. For these women earlier detection of their cancers by mammography is of little benefit.

Other screening modalities aim to detect pre-cancerous lesions, like for example, cervical cytology and sigmoidoscopy. Here the goal is to detect precursor lesions before they become invasive, so that removal carries an almost 100% cure rate. In this case a major problem is to know which lesions are likely to become cancerous if left untreated, so as to avoid over-treatment of benign lesions with no malignant potential.

A third form of screening is genetic testing. This field is continuously developing, as more cancer associated genes are found, and tests for mutations become simpler. Genetic screening will lead to very different sorts of programs, which will involve whole families, require intensive follow-up of individuals testing positive, and the use of chemo-preventive agents or prophylactic surgical removal of the organs at risk.

Other tests have morbidity due to either the test itself (*e.g.*, perforation by a colonoscope) or the diagnostic testing for the disease that follows (*e.g.*, lung biopsy). Therefore side effects need to be carefully considered. Consent needs to be obtained from potential participants in screening after they are informed of the benefits and harms (Lee, 1993).

Cancer screening programs have been implemented in many countries all over the world, especially for cancer of the uterine cervix, breast and colorectum. The main issue is whether the test reduces the mortality or morbidity to a degree commensurate with the risks and costs of the testing. Mass screening may not only reduce the mortality rate of disease but also it may reduce the incidence rate and the medication cost of the disease. Cost is a factor that should be taken into consideration in screening. The direct cost includes charges of screening test, more diagnostic procedures and follow-up. The indirect cost includes expense of time and work, management of program etc. Also, there is a psychological and biological impact of screening that should be taken into account, such as anxiety for positive results, risk of complication, harm and pain that the screening test may bring, etc.

Eddy (Eddy, 1980) stated that the effectiveness of screening needs to be compared to the costs. The benefit of screening has to outweigh the costs. If the cost is too high or no lives are saved, then screening is definitely not effective. The value of a screening test is closely related to the chance that an individual has a specific type of cancer, and this depends on many factors, such as the incidence and the prevalence of the specific cancer

type under investigation, and the age of an individual being screened. There are many more factors involved, some of which are dependent on the specifics of the cancer type, but the ones that most commonly occur in all of them are the ones mentioned before.

The models of natural history of the disease presented to date are based on a *progressive disease model*, which assumes individuals are in a healthy state until they enter the preclinical disease state and all individuals in this state eventually emerge with clinical symptoms if untreated (assuming the individual lived long enough). This last is the key assumption. An individual moves from one state to the next, never moving backwards. The period from the inception of detectable disease to the point the disease causes symptoms that prompt diagnosis ("clinical manifestation") is called the preclinical state and its duration is called the sojourn-time. To be successful, it is necessary, but not sufficient that an early detection program be capable of identifying subjects that are in this state. The state that follows the preclinical state, when the tumor is symptomatic, is called the clinical state.

The problem with the progressive disease model is that it does not accurately describe the natural history of lesions detected by screening for many cancers. It may be appropriate for cancers where the screening test detects invasive disease, such as breast cancer. But it may be inappropriate for cancers where the screening test detects preinvasive, even precancerous lesions, such as cervical or large bowel cancer. Once a lesion becomes invasive, it almost never regresses without treatment and it is assumed all invasive lesions

arise from a preinvasive lesion. On the other hand, preinvasive lesions may revert to normal tissue, some may persist, and some may progress to invasive disease (Brookmeyer, Day, 1987).

Before going into more detail about the different screening models, let us first establish some basic statistical terminology and notation.

I.2 Statistical Terminology and Notation

Incidence: the rate of new events per person during a specified time period. An "event" is an occurrence of any phenomenon of disease or health that can be discretely characterized. Therefore,

$$\text{Incidence} = \frac{\#\text{ of new events}}{\#\text{ in defined population}}$$ during a specified time period

Prevalence: includes both, new cases (events) and existing cases within a specific time period. Therefore,

$$\text{Prevalence} = \frac{\#\text{ of existing cases}}{\#\text{ in defined population}}$$ at a specified time period

Table 1: Cross classification of individuals by disease status and test result

	Positive test	Negative test
Disease present	# of True Positives (TP)	# of False Negatives (FN)
Disease absent	# of False Positives (FP)	# of True Negatives (TN)

In the evaluation of screening tests, interval cases (those that are not detected at screening but diagnosed by other means during the interval between screenings) are considered false negatives.

Sensitivity: the probability that the test is positive when given to a patient with the disease.

Sensitivity = Pr [Test is positive | Patient has the disease] = TP/ (TP + FN)

The false-negative rate we define to be one minus the sensitivity.

In screening, sensitivity may be overestimated because some detected cases may be pseudo-disease (preclinical disease that would not have produced any signs or symptoms before the individual would have died from other causes) and some false-negative cases may not be discovered during the interval follow-up (Begg, 1988). Some of the measures related to sensitivity are the prevalence of detected cases at first screening and the prevalence ratio (ratio of the prevalence of detected cases at first screening to the incidence of cases in a comparable control group) (Day et al., 1989). All these sensitivity measures are directly related to the lead time, which will be defined later.

Specificity: the probability that the test will be negative among patients who do not have the disease.

Specificity = Pr [Test is negative | Patient is healthy] = TN / (TN + FP)

The false-positive rate we definite to be one minus the specificity.

Positive Predictive Value (PPV): the probability that the patient has the disease given that he tested positive.

PPV = Pr [Patient has the disease | Test is positive] = TP / (TP + FP)

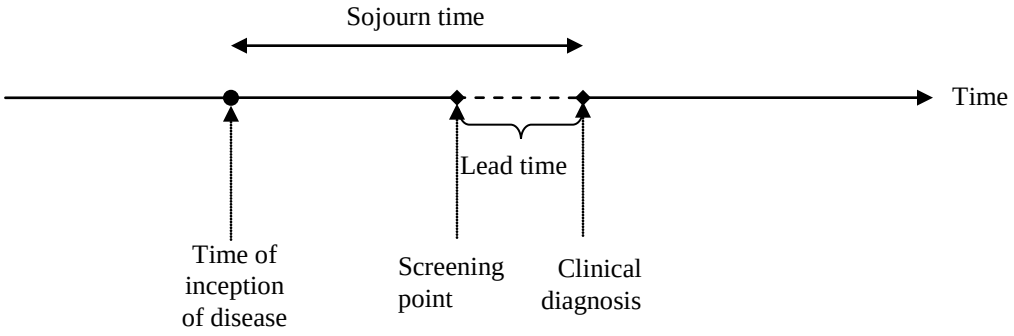
Negative Predictive Value (NPV): the probability that the patient will not have the disease given that he tested negative.

NPV = Pr [Patient is healthy | Test is negative] = TN / (TN + FN)

The period from the inception of detectable disease (in our case cancer) to the point the disease causes symptoms ("clinical manifestation") is called the preclinical state and its duration is called the sojourn-time. The early detection program must be capable of identifying subjects that are in this state. The state that follows the preclinical state, when the tumor is symptomatic, is called the clinical state.

Lead time is the amount of time by which the diagnosis has been advanced by screening (Figure 1). The problem with lead time is that it is part of the observed survival from diagnosis of screen-detected cases, biasing survival compared to cases that are not screen-detected. The lead-time distribution depends on both the sojourn-time distribution and the sensitivity of the screening test. The probability of being screen-detected depends on the screening regimen offered in terms of frequency, sensitivity and the rate of participation in the screening program.

Figure 1: Relationship between the duration of preclinical disease, screening point and lead time



I.3 Screening Evaluation Methods

The process of screening is very important to understand. Obuchowski et al. (Obuchowski et al., 2001) give the criteria for effective screening. They can be summarized as follows:

1) The disease has serious consequences, such as mortality or severe and prolonged morbidity.

2) The screened population has a high enough prevalence of detectable preclinical phase to justify the cost of screening by producing adequate PPV.

3) The screening test detects little overdiagnosis (sometimes also called pseudo-disease), that is disease that appears on the screening test, but undetected would not negatively affect the individual's life.

4) The screening test has adequately high accuracy for detecting the disease in its preclinical phase, such that the PPV and NPV are high.

5) For screening to be effective, the natural history of the disease must have a critical point (e.g., for many cancers the critical point occurs when the primary tumor metastasizes) before which therapy is more effective, and screening must advance the time of diagnosis from after the critical point to before it. Cole and Morrison (Cole et al., 1980) suggest offering screening at an optimal age and optimally spaced to affect this.

6) Screening should be applied to asymptomatic people.

7) The screening test has to be affordable and available to the target population.

Methods of screening evaluation have adapted over time, evolving from simple surveys to case-control studies (Day, 1989; Cronin et al., 1998; Sobue, 2000; Duffy et al., 2002) and randomized trials (Morabia et al., 2004).

Many investigators have moved to using randomized trials of cancer screening, and compared mortality rates from the cancer under investigation between the screened and the control group, instead of comparing case-survival rates. The advantage of randomized controlled trials (RCT) is that we can expect that all patient covariates, measured or unmeasured, to be balanced between the two treatment groups. The two treatment groups are comparable and the observed treatment difference is an unbiased estimate of true treatment difference.

One problem with RCT is that they require very large numbers of individuals that are relatively healthy and that are committed to the research for many years. In addition to, and perhaps because of that, randomized studies are very expensive and it seems that age is inversely related to the willingness to participate in cancer research (Murthy et al., 2004), although the rates of the common cancers in the United States are highest among the elderly. Older African-Americans are very hard to recruit and therefore remain under-represented in cancer prevention and control studies (Paskett et al., 1996).

Keeping this in mind, the focus of this work in on one-arm studies, since they are the most common, their data are easier to access, and one arm studies can also be formed from already collected medical and health data, including sufficiently detailed claims data such as Medicare/Medicaid. One-arm studies are performed on a group of individuals with a specified indication and managed with a specified therapy. These individuals are systematically observed to measure outcomes of interest. A quasi-experimental study can be then generated by comparing the results of one or more single-arm studies of therapy involving the proposed drug with the results of one or more similar studies (usually by different investigators in different settings) of therapy involving its main comparator(s). Data from one-arm studies are usually analyzed and compared to a historical control, where patient-level data of historical control is available and used in treatment comparison. The goal of one-arm studies is to evaluate screening performance in terms of cancer detection, interval cancer cases, tumor characteristics and survival rates. The goal of randomized studies is to prove the reduction in mortality that would be achieved by

early diagnosis of cancers. A major difference between randomized and one-arm studies is the availability of the control-group experience, and for this reason one-arm studies are easier to put in place, and therefore more one-arm data sets are available. Another difference between these studies lies in the pattern of risk of screened subjects and in recall and detection rates (Paci, 2007). Although one-armed studies may be inadequate to assess the impact on mortality, they may provide the means to assess sensitivity and specificity.

Two important parameters are deciding the effect of screening: the sensitivity of the screening test and the distribution of the sojourn time of the preclinical state (Lu et al., 2003). A high sensitivity means a strong power to detect disease. A short sojourn time of the preclinical state means that the disease has little chance to be detected by screening, which in effect means that the proportion of cases detected by screening is low and therefore screening may not be feasible. If the sojourn time is long then the interval between screens may be longer. When the distribution of the sojourn time and the sensitivity of the screening test are known, the lead time bias can be estimated for the assessment of screening. In short, without sufficient sensitivity and lead time, screening is unlikely to provide any benefit.

Lead-time bias depends on the modality of case detection (Zelen, 1976). In building probabilistic models for survival times it is not realistic to assume that all the relevant risk factors or covariates were measured and included. Unmeasured or omitted risk

factors generate a between-case variation often and interchangeably referred to as frailty (in the biomedical literature), extra variation (in the statistical literature), and residual heterogeneity (in the social sciences literature).

One key goal is to model the lead-time distribution. The probability of being screen-detected and the distribution of lead-time depend on the sojourn-time distribution, the frequency of screening, and the sensitivity of the screening test.

I.4 The Prostate Lung Colorectal and Ovarian Cancer Screening (PLCO) Trial

Lung cancer is one of the most common occurring cancers in the United States. According to Jemal (Jemal, Siegel et al., 2008), when compared between 1990 and 2004 lung and bronchus cancer rates have decreased by 22.38% in men and increased by 8.67% in women. Therefore successful screening programs for lung cancer could have a very strong impact on the overall cancer mortality in the U.S. Several one-arm studies evaluating lung cancer screening tests have been performed. In 1999 the Early Lung Cancer Action Project (ELCAP) published the results of its evaluation of early detection of lung cancer in high-risk subjects in New York, showing that low dose computed tomography (CT) is a sensitive tool for the identification of non-calcified pulmonary nodules and early lung cancer (Henschke, McCauley et al., 1999). The International ELCAP group (I-ELCAP) has confirmed, with a greater number of cases, that CT scan is

a sensitive screening test for early lung cancer and has shown high survival rates among patients with stage I screen-detected disease (Henschke, Yankelevitz et al., 2006).

Uncertainty regarding the value of screening for different types of cancer has resulted in conflicting positions in the medical community and confusion in the population at risk. For this reason, the Division of Cancer Prevention (DCP) of the National Cancer Institute (NCI), in collaboration with 10 Screening Centers (SCs) throughout United States, has tried to resolve these uncertainties by conducting a long-term randomized trial (PLCO, 1999). The 10 screening centers are: Birmingham, Ala.; Denver, Colo.; Washington, D.C.; Honolulu, Hawaii; Detroit, Mich.; Minneapolis, Minn.; St. Louis, Mo.; Pittsburgh, Pa.; Salt Lake City, Utah; and Marshfield, Wis. (NCI Fact Sheet 5.34)

The Prostate, Lung, Colorectal and Ovarian cancer screening trial (PLCO) was initiated in 1992 to examine cause-specific mortality reduction from screening for these four cancers in men and women. Between 1993, when the trial opened, and 2001, when enrollment was completed, a total of 154,942 women and men between the ages of 55 and 74 joined PLCO. They were randomized to either, screening and follow-up, or usual care and follow-up. Screening of participants continued until 2006. Additional follow-up will continue until 2014 to determine the benefits or harms of the cancer screening exams being studied. (Prorok et al., 2000)

Those cancer screening regimens were: chest X-ray for lung, flexible sigmoidoscopy for colorectal, prostate-specific antigen serum level and digital rectum examination for prostate, and CA-125 serum level and transvaginal ultrasound for ovary. The remaining half of individuals was advised to seek usual medical care. The details of the study were published as a supplement to Controlled Clinical Trials in December 2000.

In the PLCO trial for colorectal cancer (CRC), researchers are testing flexible sigmoidoscopy (FSG) at the first (baseline), third and fifth annual screen. During a sigmoidoscopy, a thin, lighted viewing instrument is inserted into the rectum to examine the left, or distal, portion of the colorectum. PLCO subjects with a polyp or mass noted on sigmoidoscopy are often referred for further examination with colonoscopy, a procedure that examines the entire colorectum. The final results of the colorectal cancer screening are expected in the next year or so.

Of 77,465 subjects randomized to the screening arm, 64,658 (83%) received the baseline flexible sigmoidoscopy exam. Because of the large size of the study population, the broad geographic representation, and the follow-up criteria, the results of the PLCO trial will offer a benchmark for screening flexible sigmoidoscopy in the United States (Weissfeld et al., 2005).

In Chapter 3 we are going to apply our developed method from Chapter 2 on the PLCO CRC data.

I.5 Preliminary Studies

The pioneers in the theory of screening for chronic disease are Zelen and Feinleib. In (Feinleib, 1967), the author derived the formula $P = mI$, where P, I and m represent the prevalence, incidence and mean duration of a specific illness. This formula holds under the condition known as the stable disease model, which assumes that:

1) The incidence is independent of time.

2) The probability distribution of time with illness is independent of when the disease has been initiated.

3) The preclinical duration probability distribution function (pdf) of time has finite range.

In (Zelen et al., 1969) the authors examined one-shot screening programs, wherein the individual is examined only once. They derived the formula for the mean sojourn time in the preclinical state and the formula for the lead-time distribution, or, synonymously, the forward recurrence-time distribution. Of particular interest is the comparison between the survival of individuals detected early through screening and those detected in the clinical state. To make this comparison valid, mean survival time of the pre-clinical group should be corrected by subtracting the mean lead time. Also, it is reasonable to assume that the clinical course of the disease is positively correlated with the preclinical course. The authors expanded the formula of the stable disease model in (Feinleib, 1967) by rewriting it in the context of time-dependent prevalence and incidence. One consequence of the

model is that if the probability density function of the sojourn time in the preclinical state does not follow an exponential distribution then the formula in (Feinleib, 1967) is biased and the authors determined the bias involved, such that the new formula will be $P(t)/I(t) = mb(t)$, where $b(t)$ is the corresponding approximate time-dependent bias. The information on the age of incidence and the age at which an individual is diagnosed in the pre-clinical state can be useful to assess the lead time. The authors make the assumption that even though prevalence and incidence are taken to be age dependent, the survival from time of diagnosis does not depend on age, but is a function of the stage of the cancer at diagnosis. The paper continues by determining the formula of mean lead time associated with a one-shot screening program, under the following assumptions:

1) Age-specific incidence of preclinical disease is constant over time

2) The survival distribution is independent of age and time

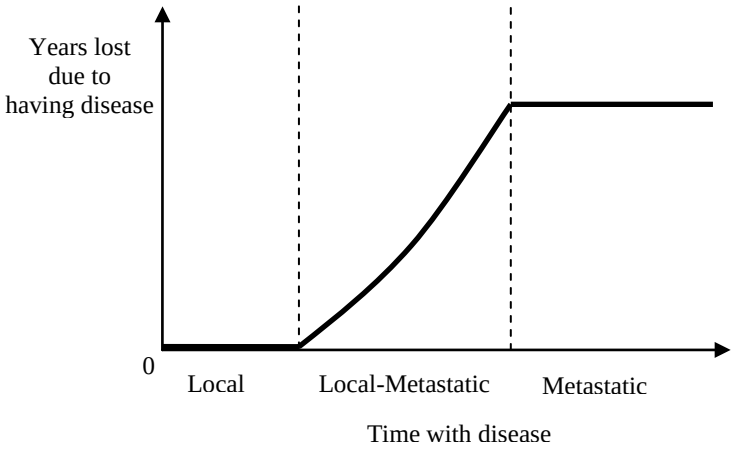
3) The origin is far removed from the time point t, where t, an absolute time, refers to the time when an individual is given the screening procedure.

One other useful result of the model is that when the average age at case-finding (prevalence) is the same as that of incidence, the duration of the preclinical disease follows an exponential distribution.

In a later work by Zelen (Zelen, 1976), the author makes a very good point by saying that an early detection program in which the bulk of cases are in the metastatic (incurable)

phase has little value, even if it succeeds in detecting disease significantly earlier (Figure 2).

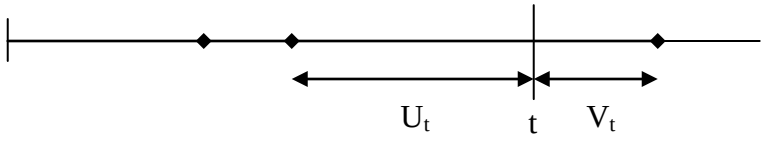
Figure 2: Years of life lost due to having disease versus time with disease for a population of patients



To enhance the theoretical aspects of the problem it is useful to use Cox's theory of stochastic processes (Cox, 1965). He considered as a point process occurrences of the individual events themselves, distinguished only by their positions in time. The Poisson process is the simplest point process. A discrete-time point process $\{X_n\}$, applicable to screening, is stationary if the sets $X_{n_1},...,X_{n_k}$ and $X_{n_1+v},...,X_{n_k+v}$ have the same joint distribution for all $n_1,...,n_k,v$. In other words, the joint distribution depends only on the intervals between the time points $n_1,...,n_k$ and is unaffected by an arbitrary translation v of the time points.

The backward recurrence-time U_t is the length of time measured backwards from t to the last event at or before t. The forward recurrence-time V_t is the length of time measured from t to the next event at or after t (Figure 3). Cox has treated the situation where the point process is sampled by a 'sampling point', but in (Prorok, 1976, a) the point process is sampled by a 'sampling interval' of finite length.

Figure 3: Backward and forward recurrence-time



I.6 Prorok's Model

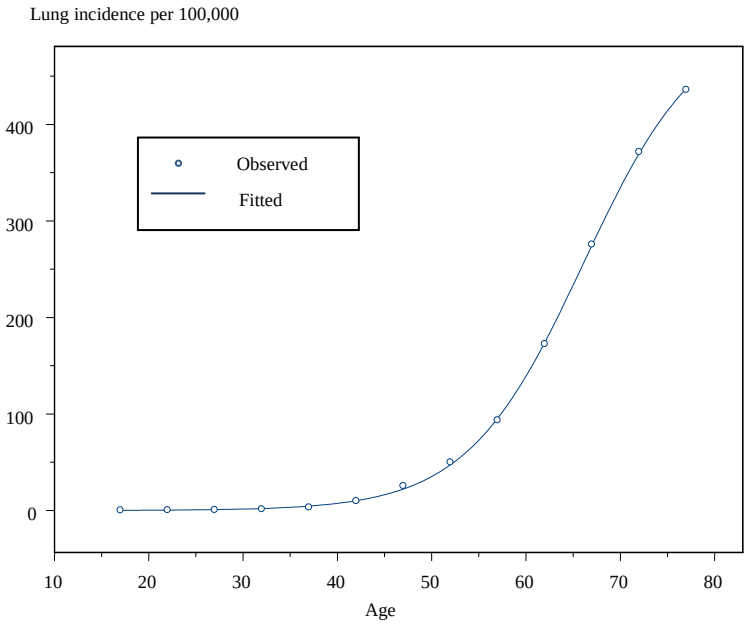
Prior work focused on the single-shot screening. Prorok (Prorok, 1976, a, b) addressed periodic screening, with the period denoted by Δ, using stationary point processes and extending Cox's results (Cox, 1965). The duration of the screening examination and diagnostic process is considered to be small relative to the interval between screens. A sampling point is the time of occurrence of the first screening examination after an individual enters the preclinical state. This point is independent of both the time at which the preclinical state is entered and the duration of stay in the preclinical state. An individual is a member of the i^{th} generation if he enters the preclinical state sometime during the time interval $((i-1)\Delta, i\Delta]$, where $i=\overline{1,k}$ and $k+1$ is the number of screening examinations during the screening process. Denote by S_0 the disease-free state, by S_p the

preclinical disease state and by S_c the clinical disease state. Zero-generation individuals are those who are in S_p at time 0. Let T_i be a non-negative random variable denoting the sojourn time in S_p of the i^{th} generation of individuals. Let V_i be a non-negative random variable denoting the forward recurrence time in S_p of the i^{th} generation of individuals, measured from the time of the first screening point which an i^{th} generation of individuals encounters after entering S_p, which is time $i\Delta$. Let U_i be a non-negative random variable denoting the backward recurrence time in S_p of the i^{th} generation of individuals, measured from time $i\Delta$. The variables T_i, V_i and U_i are assumed to be absolutely continuous, $T_i = U_i + V_i, (\forall) i = \overline{1,k}$ and each series consists of independent and identically distributed (i.i.d.) random variables. Under the restriction that the forward recurrence time is bounded below and the backward recurrence time is bounded above, the author developed the formula of the probability density function (p.d.f.) of the forward recurrence time. Another assumption that the author made is the fact that the entry in S_p occurs uniformly. The result presented is in fact an extension of Cox's result.

Prorok (Prorok, 1976, a; Prorok, 1982), Zelen (Zelen, Feinleib, 1969; Zelen, 1976; Zelen, 2004) and Hutchinson (Hutchinson, Shapiro, 1968) developed and used some of the earliest models of periodic screening that couple a progressive disease model with a preclinical sojourn time distribution, $q(t)$, and incorporate simplifying assumptions of a constant probability, w, of initiating a preclinical disease phase and a constant sensitivity, β, over the course of the preclinical phase. Such recurrence-time models can be used to produce maximum likelihood estimates of the parameters of the lead-time distribution

gained by screening a population. However, for many applications the assumptions that these are constants are unrealistic. For example, in many cancers, the incidence increases with age at an approximately exponential rate. In the SEER (Surveillance Epidemiology and End Results) lung cancer data of 2003-2007 the incidence function that fits the data best is of logistic type and not uniform (Figure 4).

Figure 4: Logistic function that fits the incidence of SEER lung cancer data of 2003-2007



Also, for many screening tests the sensitivity increases as the size or stage of the tumor increases, such as in lung cancer screening by radiography or colorectal cancer screening by fecal occult blood tests.

In addition to this, the assumption of a negative exponential preclinical time distribution is unrealistic, since most cancers don't become clinical at the time of detection by screening, but later on, sometimes on an average of 3 years (mode=2) after being screen detected (Figure 5). Therefore a log-normal distribution or a log-logistic distribution (Figure 6) is a much more reasonable choice of preclinical time distribution.

Figure 5: Log-normal distribution with mean 3 compared to negative exponential distribution with mean 3

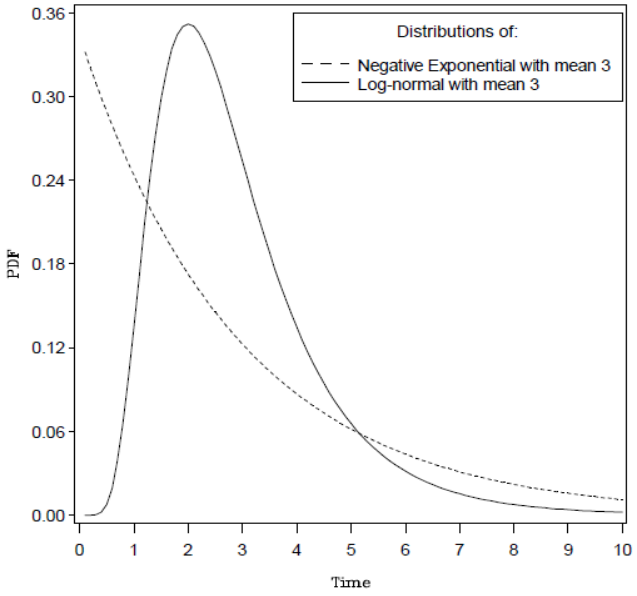
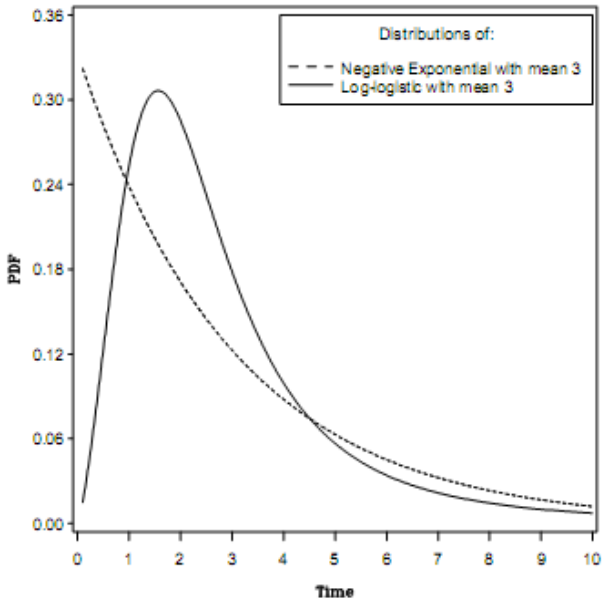


Figure 6: Log-logistic distribution with mean 3 compared to negative exponential distribution with mean 3



Others like Kadafar (Kadafar, Prorok, 1994; Kadafar, Prorok, 1997) have tried to estimate the lead time bias by using a closed form approximation, relying on observed survival curves, both since time of entry into the study and since time of diagnosis.

Loeve (Loeve et al., 1999) used a microsimulation model for evaluating colorectal cancer screening. Likelihood based methods have been proposed by Day (Day, Walter, 1984) and Chen (Chen et al., 2000) in order to estimate the sojourn time distribution. Day made the assumption that the sensitivity is constant throughout the entire screening program, that the sojourn time is age independent and that the incidence of the preclinical state is uniform for an individual over the duration of the study. Chen modeled the disease

process for a chronic disease as a continuous-time Markov process. She made the assumption that the screening sensitivity is 100% and that the preclinical incidence is exponential. Duffy (Duffy et al., 2008) proposed a method to correct for lead-time bias for cancer screening, but he also assumed an exponential distribution for the sojourn time and a uniform distribution of the screen detected probabilities.

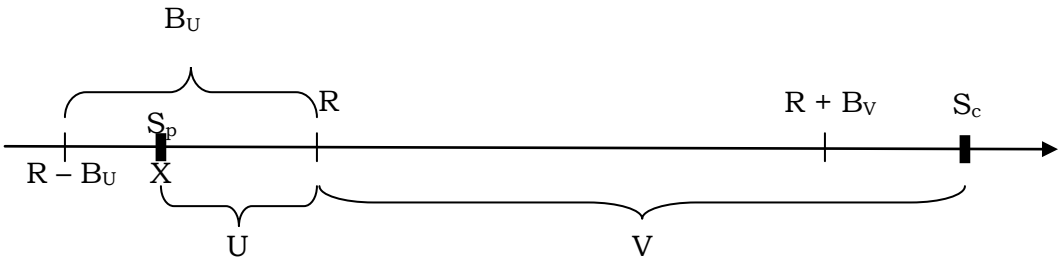
The literature on lead time distribution for cancer screening is quite vast but in 35 years since Prorok's 1976 paper, the formulas developed by Prorok haven't been modified in order to incorporate a non-constant preclinical incidence, non-exponential preclinical duration, or time-dependent sensitivity. Chapter 2 presents a modification of the recurrence-time model of Prorok (Prorok, 1976, a) to incorporate an age-dependent preclinical incidence $w(x)$ and a generation- and screen-dependent sensitivity function β_{ij}. It will extend Prorok's results by deriving the formula for the distribution of the forward recurrence time and lead time based on this modified model. The notation will be based on Prorok's, with some changes to accommodate the modifications.

Chapter II: Recurrence Time Model Derivation and Simulation

II.1 General Result for the Forward Recurrence Time Distribution

R is the focal sampling point of the screening program (Figure 7), assumed to be far removed from the time at which the disease process starts in the population being screened;

Figure 7: Backward and forward recurrence time in relation to the focal sampling point



U is the non-negative random variable denoting the backward recurrence time in S_p measured from the time of the inception of the disease X to the time of screening *R*;

V is the non-negative random variable denoting the forward recurrence time in S_p measured from the time of screening *R* to the time of entering in the clinical phase S_c;

$w(x)dx$ is the probability that an individual enters S_p during the interval $(x, x + dx)$;

$T = U + V$ sojourn time in S_p, $0 \le U < B_U$, $0 \le B_V \le V < +\infty$

$\Longrightarrow T \in [B_V, +\infty)$

$\Longrightarrow R - B_U < x \le R$ and $x + T \ge x + B_V$.

If the patient stays in S_p at least until $R + B_V$ then

$x + T \ge R + B_V \Longrightarrow x \ge R + B_V - T \Longrightarrow R \ge x \ge \max(R - B_U, R + B_V - T)$.

Let $Z(y) = \begin{cases} 1, & if\ patient\ is\ in\ S_p\ at\ time\ y \\ 0, & otherwise \end{cases}$

$$\Longrightarrow P(Z(R + B_V) = 1 | t < T \le t + dt) = \int_{max\,(R - B_U, R + B_V - t)}^{R} w(x)dx \qquad (*)$$

Therefore:

$$P(Z(R + B_V) = 1 | t < T \le t + dt) = \begin{cases} \int_{R + B_V - t}^{R} w(x)dx\ , if\ B_V \le t < B_U + B_V \\ \int_{R - B_U}^{R} w(x)dx\ , if\ B_U + B_V \le t < +\infty \end{cases}$$

Let *g(t)* be the pdf of those T sampled s.t. $B_V \le V$.

$\Longrightarrow g(t)dt = K \cdot P(Z(R + B_V) = 1, t < T \le t + dt)$,

where *K* is the normalizing constant.

$\Longrightarrow g(t)dt = K \cdot P(Z(R + B_V) = 1 | t < T \le t + dt) \cdot P(t < T \le t + dt)$.

Let us now denote by $q(t)dt = P(t < T \leq t + dt)\ and\ Q(t) = \int_t^{+\infty} q(y)dy.$

$$\Longrightarrow K = \left[\int_{R-B_U}^{R} w(x) \cdot Q(R + B_V - x)dx \right]^{-1} \qquad (1)$$

Now, for T fixed, V is distributed uniform over its range and

$T \geq V \geq max(B_V, T - B_U).$

Therefore, for *T* fixed at $t \geq B_V$ we obtain

$$f(v,t) = f(v|t) \cdot g(t)$$

$$\cong \begin{cases} \dfrac{K \cdot q(t)}{t - B_V} \displaystyle\int_{R+B_V-t}^{R} w(x)dx, & if\ B_V \leq v \leq t < B_U + B_V \\ \dfrac{K \cdot q(t)}{B_U} \displaystyle\int_{R-B_U}^{R} w(x)dx, & if\ t \geq B_U + B_V\ and\ v \leq t \leq v + B_U \\ 0, \quad otherwise \end{cases}$$

where *K* is defined in (1).

Next we are going to use the following notation $Q(x,y) = Q(x) - Q(x+y).$

Then, if $B_V \leq v < +\infty$ we get

$$f(v) = \begin{cases} K \cdot \left[\displaystyle\int_v^{B_U+B_V} \frac{q(t)}{t-B_V} \int_{R+B_V-t}^{R} w(x)dxdt + \frac{Q(B_U+B_V, v-B_V)}{B_U} \int_{R-B_U}^{R} w(x)dx \right], \text{ if } B_V \leq v < B_U + B_V \\ \dfrac{K}{B_U} \cdot Q(v, B_U) \cdot \displaystyle\int_{R-B_U}^{R} w(x)dx, \text{ if } B_U + B_V \leq v \\ 0, \text{ otherwise} \end{cases} \qquad (2)$$

II.2 Forward Recurrence Time for Each Generation

Generation 0: those who are in S_p at time 0, considered to be the age at baseline.

Therefore, $R = age$.

Possible times of detection are $age + j\Delta$, $= \overline{0,k}$.

$B_U = +\infty$, $B_V = j\Delta \implies j\Delta \leq V_0, j = \overline{0,k}$

Therefore from (1) we get

$$K_0(age, j) = \left[\int_0^{age} w(x) \cdot Q(age + j\Delta - x)\, dx\right]^{-1} \qquad (3)$$

In conclusion,

$$f_{0,age,j}(v) = K_0(age, j) \cdot \int_v^{\infty} \frac{q(t)}{t - j\Delta} \int_{age+j\Delta-t}^{age} w(x) dx dt \qquad (4)$$

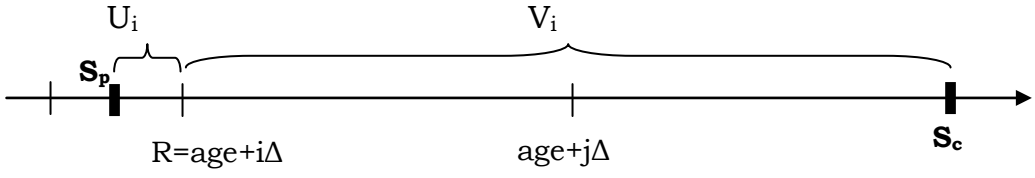
if $j\Delta \leq v < +\infty$, and 0 otherwise.

Generation i>0: those who enter S_p sometime during the time interval

$(age + (i-1)\Delta, age + i\Delta]$ where $i = \overline{1,k}$. We have $B_U = \Delta$ and

$B_V = (j-i)\Delta, j \geq i$.

Figure 8: Backward and forward recurrence time for the i^{th} generation



Here from (1) we get

$$K_i(age,j) = \left[\int_{age+(i-1)\Delta}^{age+i\Delta} w(x)\cdot Q(age+j\Delta-x)dx\right]^{-1} \quad (5)$$

Then

$$f_{i,age,j}(v)=K_i(age,j)\cdot\begin{cases}\left[\int_v^{(j-i+1)\Delta}\frac{q(t)}{t-(j-i)\Delta}\int_{age+j\Delta-t}^{age+i\Delta}w(x)dxdt+\frac{Q((j-i+1)\Delta,v-(j-i)\Delta)}{\Delta}\int_{age+(i-1)\Delta}^{age+i\Delta}w(x)dx\right], \\ \qquad\qquad \text{if } (j-i)\Delta\le v<(j-i+1)\Delta \\ \frac{1}{\Delta}\cdot Q(v,\Delta)\cdot\int_{age+(i-1)\Delta}^{age+i\Delta}w(x)dx, \text{ if } (j-i+1)\Delta\le v \\ 0, \text{ otherwise}\end{cases} \quad (6)$$

II.3 Lead-Time Distribution for Screen Detected Individuals

Let $l_{i,age,j}$ be the lead time for the i^{th} generation of individuals detected by screening at time $age+j\Delta$, where $i = \overline{0,k}$ and $j = \overline{i,k}$.

Generation 0:

$l_{0,age,j} = V_{0,age} - j\Delta \Rightarrow V_{0,age} = l_{0,age,j} + j\Delta, j = \overline{0,k}$.

Let $h_{0,age,j}(l)$ be the p.d.f. of the lead time of those zero-generation individuals who remain in S_p at least until time $age+j\Delta$, are not detected before $age+j\Delta$ but are detected at the $(j+1)^{th}$ screen.

Therefore, from (3) and (4) we obtain:

$$h_{0,age\,,j}(l) \;=\; \begin{cases} K_0(age,j)\cdot\left[\int_{l+j\Delta}^{\infty}\frac{q(t)}{t-j\Delta}\int_{age\,+j\,\Delta-t}^{age} w(x)dxdt\right], if\ l\ \geq\ 0 \\ 0, otherwise \end{cases} \tag{7}$$

Generation i >0:

$l_{i,age\,,j} \;=\; V_{i,age} - (j-i)\Delta \Longrightarrow\; V_i, age \;=\; l_{i,age\,,j} \;+\; (j-i)\Delta$ and

$j\Delta - V_{i,age} \;=\; i\Delta \;-\; l_{i,age\,,j}$

Therefore from (5) and (6) we obtain:

$$h_{i,age\,,j}(l)$$

$$= K_i(age,j)$$

$$\cdot\begin{cases} \left[\int_{l+(j-i)\Delta}^{(j-i+1)\Delta}\frac{q(t)}{t-(j-i)\Delta}\int_{age\,+j\,\Delta-t}^{age\,+i\Delta} w(x)dxdt + \frac{Q((j-i+1)\Delta,l)}{\Delta}\int_{age\,+(i-1)\Delta}^{age\,+i\Delta} w(x)dx\right], \\ \qquad\qquad if\ 0\leq l<\Delta \\ \frac{1}{\Delta}\cdot Q(l+(j-i)\Delta,\Delta)\cdot\int_{age\,+(i-1)\Delta}^{age\,+i\Delta} w(x)dx, if\ \ \Delta\leq l \\ 0, otherwise \end{cases} \tag{8}$$

II.4 Properties of the lead time in a periodic screening program

Weighting factors

Note: withdrawals and lost to follow up cases are ignored for now. Time 0 is considered to be the age at baseline.

We are going to use the following notations:

$E_0 =$ the event that a person is in S_p at time 0;

$P_0 =$ P(person is in S_p at time 0) the prevalence;

$$\Rightarrow P_0 = P(Z(0) = 1) = \int_{-\infty}^{\infty} P(Z(0) = 1 | t < T \leq t + dt) \cdot P(t < T \leq t + dt) dt$$

$$\overset{(*)}{\Rightarrow} P_0(age) = \int_0^{\infty} q(t) \int_{age-t}^{age} w(x) dx dt = \int_0^{age} w(x) \int_{age-x}^{\infty} q(t) dt dx \qquad (9)$$

$W_0 = 1$ by definition, meaning 'identically'.

Now, for $i = \overline{1,k}$ we have:

$E_i =$ the event that a person is a member of the i^{th} generation;

$P(E_i) =$ the incidence rate, $P(E_0, age) = P_0(age)$;

$$\Rightarrow P(E_i, age) = \begin{cases} P_0(age) & , if\ i = 0 \\ \int_{age + (i-1)\Delta}^{age + i\Delta} w(x) dx & , if\ i = \overline{1,k} \end{cases} \qquad (10)$$

R_i = the event that $S_p \leq age + (i-1)\Delta$;

W_i = P(person enters S_p during $(age + (i-1)\Delta, age + i\Delta]$ | not entered previously);

$P(age + i\Delta)$ = P(the i^{th} generation individual is in S_p at time $age + i\Delta$)

$\Longrightarrow$ $if\ i = 0\ then\ P(age + i\Delta) = 1.$

β_{ij} = P(present preclinical disease for the i^{th} generation individual is detected at the (j+1) screen); we take it to be age independent;

$D_{i,age,j}$ = P(i^{th} generation individual remains in S_p at least until time age + $j\Delta$, i.e. $V_i \geq (j-i)\Delta$, and it goes undetected for j-1 screens before time $age + j\Delta$, and is detected at time $age + j\Delta$)

$$= \begin{cases} P(E_i, age) \cdot P(age + i\Delta) \cdot \underbrace{P(V_i \geq (j-i)\Delta)}_{Q_{V_i,age}((j-i)\Delta)} \cdot \beta_{ij} \cdot \prod_{q=i}^{j-1}(1-\beta_{iq}), i = \overline{0,k}, j = \overline{i,k} \\ 0, otherwise \end{cases} \quad (11)$$

Now from (5) we have that

$$K_i(age, i) = \left[\int_{age + (i-1)\Delta}^{age + i\Delta} w(x)Q(age + i\Delta - x)dx\right]^{-1} \quad (12)$$

$$\Longrightarrow Q_{V_i,age}(x) = P(V_i \geq x) = \int_x^{\infty} f_{i,age,i}(v)dv$$

$$\overset{(6)}{\Rightarrow} Q_{V_i,\,age}(x) = K_i(age,i)\cdot \begin{cases} \int_x^{\Delta}\int_v^{\Delta} \frac{q(t)}{t} \int_{age+i\Delta-t}^{age+i\Delta} w(y)dydtdv + \int_x^{\Delta} \frac{Q(\Delta,v)}{\Delta} \int_{age+(i-1)\Delta}^{age+i\Delta} w(y)dydv \\ + \int_{\Delta}^{\infty} \frac{Q(v,\Delta)}{\Delta} dv \cdot \int_{age+(i-1)\Delta}^{age+i\Delta} w(y)dy, \quad \text{if } x\in[0,\Delta) \\ \int_x^{\infty} \frac{Q(v,\Delta)}{\Delta} dv \cdot \int_{age+(i-1)\Delta}^{age+i\Delta} w(y)dy, \quad \text{if } x\in[\Delta,+\infty) \end{cases}$$

$$i = \overline{1,k} \qquad (13)$$

If $i = 0$ then we have from (3)

$$K_0(age,0) = \left[\int_0^{age} w(x)\cdot Q(age - x)\,dx\right]^{-1} \qquad (14)$$

$$\Longrightarrow Q_{V_0,age}(x) = P(V_0 \geq x) = \int_x^{\infty} f_{0,age,0}(v)dv$$

$$\overset{(4)}{\Rightarrow} Q_{V_0,age}(x) = K_0(age,0)\cdot \int_x^{\infty}\int_v^{\infty} \frac{q(t)}{t} \int_{age-t}^{age} w(y)dydtdv, x \geq 0 \qquad (15)$$

Let now

$$Z_i(y) = \begin{cases} 1, & if\ an\ individual\ from\ the\ i^{th}\ generation\ is\ in\ S_p\ at\ time\ y \\ 0, & otherwise \end{cases}$$

Let $R = age + i\Delta, B_U = \Delta, B_V = 0 \Longrightarrow 0 \leq U_i \leq B_U = \Delta$ and

$0 = B_V \leq V_i < +\infty \Longrightarrow R + B_V = age + i\Delta\,, i = \overline{1,k}\,.$

$$P(Z_i(age\ +\ i\Delta) = 1, t < T_i \leq t + dt)$$

$$= \begin{cases} q_i(t) \cdot \int_{age\ +i\Delta-t}^{age\ +i\Delta} w(x)dxdt, & if\ 0\ \leq\ t\ <\ \Delta \\ q_i(t) \cdot \int_{age\ +(i-1)\Delta}^{age\ +i\Delta} w(x)dxdt, & if\ \Delta \leq\ t\ <\ +\infty \\ 0, & otherwise \end{cases}$$

$$\Longrightarrow\ P(age + i\Delta)\ =\ P(Z_i(age\ +\ i\Delta) = 1)$$

$$= \int_{age\ +(i-1)\Delta}^{age\ +i\Delta} w(x)Q_i(age\ +\ i\Delta\ -\ x)dx, i = \overline{1,k}\,. \qquad (16)$$

Now $P(E_0, age) = P_0(age)$ and $P(E_i, age)\ =\ P(E_i, age|R_i, age) \cdot P(R_i, age) +$

$P(E_i, age|\overline{R_i, age}) \cdot P(\overline{R_i, age})$.

The progressive disease model tell us that

$P(E_i, age|R_i, age)\ =\ 0$ and $P(E_i|\overline{R_i})\ =\ W_i\ ,\ i\ =\ \overline{1,k}$. Also we have that

$\overline{R_i}\ =\ \cap_{j=0}^{i-1}\overline{E_j}\ .$

$$\Longrightarrow\ P(\overline{R_i})\ =\ P\left(\bigcap_{j=0}^{i-1}\overline{E_j}\right) =\ P(\overline{E_0}) \cdot \prod_{s=1}^{i-1} P(\overline{E_{i-s}}\ |\ \overline{R_{i-s}})$$

Now, $P(\overline{E_0})\ =\ 1 - P_0$ and define $\delta_{i0}\ =\ \begin{cases} 1, & if\ i = 0 \\ 0, & if\ i \neq 0 \end{cases}$,

$$\pi_i = \begin{cases} 1, & if\ i\ =\ 0\ or\ 1 \\ \prod_{n=1}^{i-1}(1 - W_n), & if\ i\ =\ \overline{2,k} \end{cases}$$

$$\Longrightarrow\ P(E_i, age)\ =\ W_i \cdot (1 - P_0(age)) \cdot\ \pi_i\ , for\ i\ =\ \overline{1,k}$$

If $i = 0 \Longrightarrow P(E_0, age) = P_0(age), W_0 = 1$ and $\pi_0 = 1$.

$$\Longrightarrow P(E_i, age) = W_i \cdot P_0^{\delta_{i0}}(age) \cdot (1 - P_0(age))^{1-\delta_{i0}} \cdot \pi_i \, , i = \overline{0,k} \qquad (17)$$

Let now $\tau_{ij} = \begin{cases} 0, & if\ i = j \\ 1, & if\ i \neq j \end{cases}$.

Therefore, from (13), (15), (16) and (17) we obtain the weighting factor for the i^{th} generation individuals detected at time $age + j\Delta$, i.e. at the $(j+1)^{th}$ screen:

$$D_{i,\,age,\,j} = \begin{cases} W_i \pi_i (1\text{-}P_0(age)) \int_{age + (i-1)\Delta}^{age + i\Delta} w(x) Q_i(age+i\Delta\text{-}x)dx \cdot Q_{V_i,\,age}((j\text{-}i)\Delta)\beta_{ij} \left(\prod_{q=i}^{j-1} (1\text{-}\beta_{iq}) \right)^{\tau_{ij}}, & \text{if } i = \overline{1,k},\ j = \overline{i,k} \\ P_0(age) Q_{V_0,\,age}(j\Delta)\beta_{0j} \left(\prod_{q=0}^{j-1} \left(1\text{-}\beta_{0q}\right) \right)^{\tau_{ij}}, & \text{if } i = 0,\ \ j = \overline{i,k} \\ 0\ , & \text{otherwise} \end{cases} \qquad (18)$$

And now let us denote by $C_{i,age\,,j}$ the weighing factors for those diseased individuals who are preclinical at some time during the course of the screening program, but are not detected by screening. For $i = \overline{0, k-1}$ and $j \in (i, k+1)$, this is the probability that an i^{th} generation individual becomes clinical during the interval

$(age + (j-1)\Delta, age + j\Delta]$.

We denote by

$$Q_{V_i,age}[(j-i-1)\Delta, \Delta] = Q_{V_i,age}\big((j-i-1)\Delta\big) - Q_{V_i,age}\big((j-i)\Delta\big).$$

Then

$$C_{i,age\,,j}$$

$$= \begin{cases} P(E_i, age)[1 - P(age + i\Delta)], & i = \overline{1,k}, \quad j = i \\ P(E_i, age)P(age + i\Delta)Q_{V_i,age}\left((k-i)\Delta\right)\prod_{q=i}^{k}\left(1-\beta_{iq}\right), & i = \overline{0,k}, \ j = k+1 \\ P(E_i, age)P(age + i\Delta)Q_{V_i,age}\left[(j-i-1)\Delta, \Delta\right]\prod_{q=i}^{j-1}\left(1-\beta_{iq}\right), & i = \overline{0,k-1}, \ i < j < k+1 \\ 0, \ otherwise \end{cases} \quad (19)$$

II.4.1 Local and global mean lead times

Let:

$L_{i,age\,,j}(1)$ be the mean lead time of those i^{th} generation individuals detected at the (j+1) $^{\text{th}}$ screen at time $age + j\Delta, i = \overline{0,k}, j = \overline{i,k}$;

$L_D(1, age, j)$ be the mean lead time of all diseased individuals detected at the (j+1)$^{\text{th}}$ screen, $j = \overline{0,k}$;

$L_C(1, age, j)$ be the mean lead time of all diseased individuals who are susceptible to detection at the (j+1)$^{\text{th}}$ screen, $j = \overline{0,k}$;

$L_D(1, age)$ be the mean lead time of all diseased individuals detected by the entire screening program;

$L_C(1, age)$ be the mean lead time resulting from the entire screening program for all individuals who are susceptible to detection sometime during the screening program.

Therefore, from (7) we get, for j = $\overline{0, k}$, that

$$L_{0,age\,,j}(1) \;=\; \int_0^{\infty} s \cdot h_{0,age\,,j}(s)ds \;=\; K_0(age,j) \cdot \int_0^{\infty} s \cdot \left[\int_{s+j\,\Delta}^{\infty} \frac{q(t)}{t-j\Delta} \int_{age\,+j\,\Delta-t}^{age} w(x)dxdt \right] ds \quad (20)$$

From (8) we get, for i= $\overline{1, k}$, j= $\overline{i, k}$, that

$$L_{i,age\,,j}(1) \;=\; \int_0^{\infty} l \cdot h_{i,age\,,j}(l)dl = \; K_i(age,j) \cdot \left[\int_0^{\Delta} \int_{l+(j-i)\Delta}^{(j-i+1)\Delta} \frac{q(t)}{t-(j-i)\Delta} \int_{age\,+j\,\Delta-t}^{age\,+i\Delta} w(x)dxdtdl \right.$$

$$\left. + \int_0^{\Delta} \frac{Q\big((j-i+1)\Delta, l\big)}{\Delta} dl \cdot \int_{age\,+(i-1)\Delta}^{age\,+i\Delta} w(x)dx + \int_{\Delta}^{\infty} \frac{Q(l+(j-i)\Delta, \Delta)}{\Delta} dl \cdot \int_{age\,+(i-1)\Delta}^{age\,+i\Delta} w(x)dx \right] \quad (21)$$

$$\Rightarrow\; L_D(1, age, j) \;=\; \frac{\sum_{i=0}^{j} D_{i,age\,,j} \cdot L_{i,age\,,j}(1)}{\sum_{i=0}^{j} D_{i,age\,,j}}\,, \quad j \;=\; \overline{0, k} \qquad (22)$$

$$\Rightarrow\; L_C(1, age, j) \;=\; \frac{\sum_{i=0}^{j} D_{i,age\,,j} \cdot L_{i,age\,,j}(1)}{\sum_{i=0}^{j}\big(D_{i,age\,,j} \;+\; C_{i,age\,,j} \;+\; C_{i,age\,,j+1}\big)}\,, \quad j \;=\; \overline{0, k} \qquad (23)$$

The global mean lead times:

$$L_D(1, age) = \frac{\sum_{i=0}^{k} \sum_{j=i}^{k} D_{i,age\,,j} \cdot L_{i,age\,,j}(1)}{\sum_{i=0}^{k} \sum_{j=i}^{k} D_{i,age\,,j}}\,, \qquad (24)$$

$$\mathrm{L_C}(1, \mathrm{age}) = \frac{\sum_{\mathrm{i=0}}^{\mathrm{k}} \sum_{\mathrm{j=i}}^{\mathrm{k}} \mathrm{D_{i,\,age,\,j}} \cdot \mathrm{L_{i,\,age,\,j}}(1)}{\sum_{\mathrm{i=0}}^{\mathrm{k}} \sum_{\mathrm{j=i}}^{\mathrm{k}} \big(\mathrm{D_{i,\,age,\,j}} + \mathrm{C_{i,\,age,\,j}} + \mathrm{C_{i,\,age,\,j+1}}\big)}\,, \qquad (25)$$

II.5 Simulation Results

In order to obtain the estimates of parameter values we formed the likelihood using simulated data and maximized its log (MathCAD v.14). The form of the log likelihood is:

$$LL(var, m, a, b, c, \beta_{00}, \tau) = \sum_{age=45}^{75} [\, N1_{age} \cdot ln\left(D_{0,age,0}\right) + N2_{age} \cdot ln\left(C_{0,age,1} + C_{1,age,1}\right)$$

$$+ N3_{age} \cdot ln\left(D_{0,age,1} + D_{1,age,1}\right) + N4_{age} \cdot ln\left(C_{0,age,2} + C_{1,age,2} + C_{2,age,2}\right)$$

$$+ N5_{age} \cdot ln\left(D_{0,age,2} + D_{1,age,2} + D_{2,age,2}\right)$$

$$+ \left(N_{age} - \sum_{i=1}^{5} Ni_{age}\right) \cdot ln\left(1 - D_{0,age,0} - \left(\sum_{i=0}^{2}\sum_{j=i}^{2} C_{i,age,j}\right) - \sum_{i=0}^{2}\sum_{j=i}^{2} D_{i,age,j}\right)\Bigg] \qquad (26)$$

where $N1_{age}$ ... $N5_{age}$ are the corresponding number of subjects in each subgroup for each age cohort and N_{age} is the total number of subjects who are in the age cohort.

The maximization method used by MathCAD for the Maximize function is a quasi-Newton method. Quasi-Newton methods are based on Newton's method to find the point of a function where the gradient is 0. Newton's method assumes that the function can be locally approximated as a quadratic in the region around the maximum, and uses the first and second derivatives (gradient and Hessian) to find the maximum point.

In quasi-Newton methods the Hessian matrix of second derivatives of the function to be minimized or maximized does not need to be computed. The Hessian is updated by analyzing successive gradient vectors instead.

This section will use simulated data in order to test the accuracy of the maximum likelihood estimates (MLEs) obtained from the preceding sections. The goal is to see if one can get reasonable estimates back from the data, when one knows the true parameter values.

The simulated data is a population with 1000 lung cancer cases undergoing a baseline screen and 2 follow-up screens. One needs to figure out how big the population has to be to generate 1000 cases. The simulated population included both men and women age 45 to 75, with age distribution proportional to the US Census data (NP2008_D1), and therefore the population size is related to the probability of cases being generated by a population like the US Census. This data set required a total of 131887 individuals. This number has been obtained by dividing 1000 by the weighted sum of integrated incidence function by each age frequency and was based on the SEER (Surveillance Epidemiology and End Results) lung cancer data of 2003-2007.

We used the functional form previously used to fit age-specific colorectal incidence (Church TR, 1999):

$$w(x) = \frac{c/100000}{1 + e^{-bx-a}} \quad . \quad (27)$$

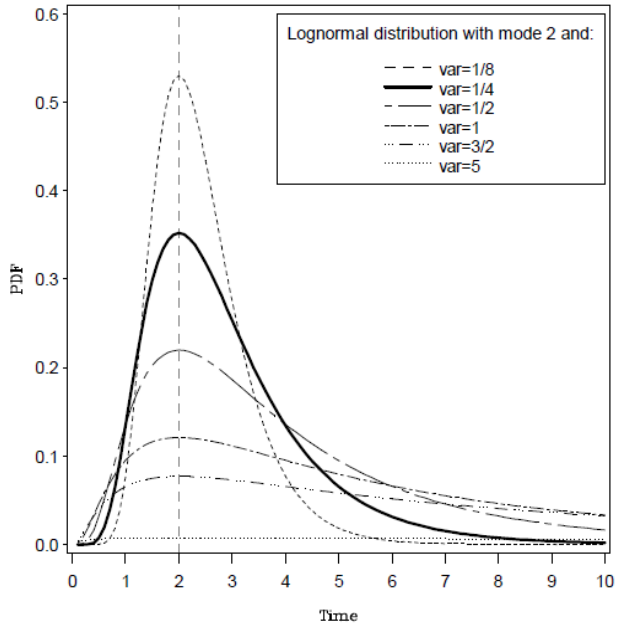
By fitting this function to the SEER data (Figure 4) we obtained the three constants to be

a = -10.7123, b = 0.1621, c = 511.574

which produced a very good fit (goodness of fit test resulted in p = 0.9998). For the purpose of our simulation we chose people ages 45-75.

For the preclinical time distribution we used a lognormal distribution with variance var=0.25 and mode m=2 (Figure 5) since we wanted to represent a cancer where the most common value for the preclinical duration is 2 years and majority of clinical symptoms occur at an interval greater than 2/10 years but less than 6 years from the time of cancer starting point (Figure 9).

Figure 9: Log-normal distribution with mode 2 and various variances



To simulate screening we let the generation index *i* take values from 0 to 2 and screen *j* take values from *i* to 2. For sensitivity of screen *j* among generation *i* we defined

$$\beta_{ij} = \frac{\beta_{00}}{2^i} + (j - i)\tau \quad (28)$$

where $\beta_{00} = 0.4$ and $\tau = 0.1$.

Figures (10 through 12) present the graphical results obtained for the lead time distribution for generations 0, 1 and 2, for individuals who are 45 years old.

Generation 0:

Figure 10a

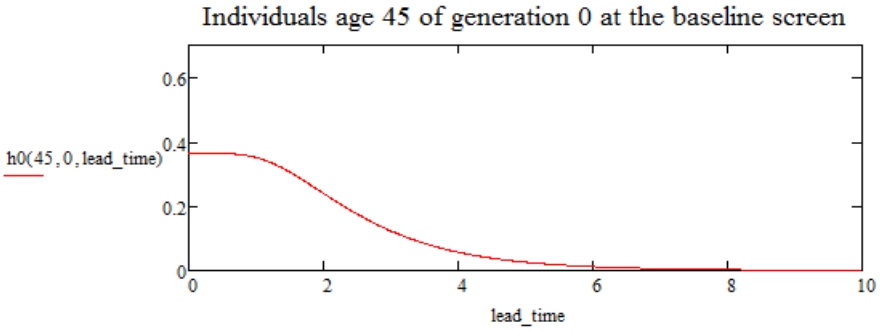


Figure 10b

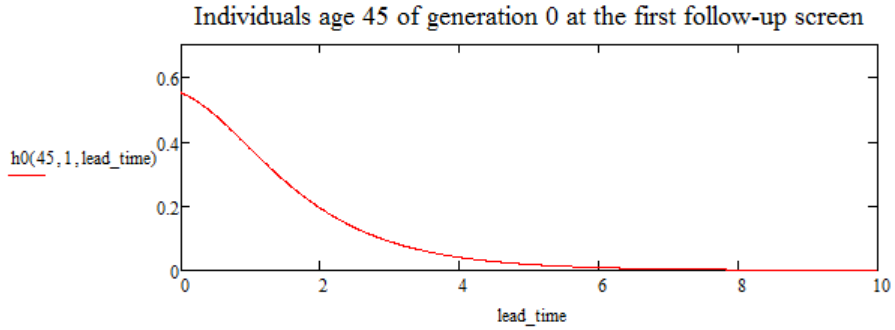
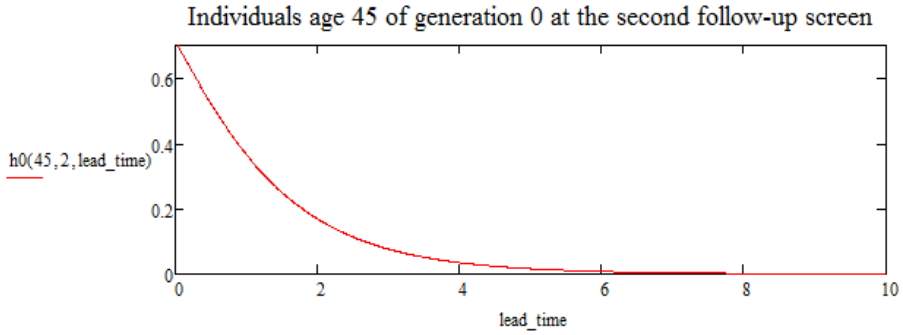


Figure 10c



Generation 1:

Figure 11a

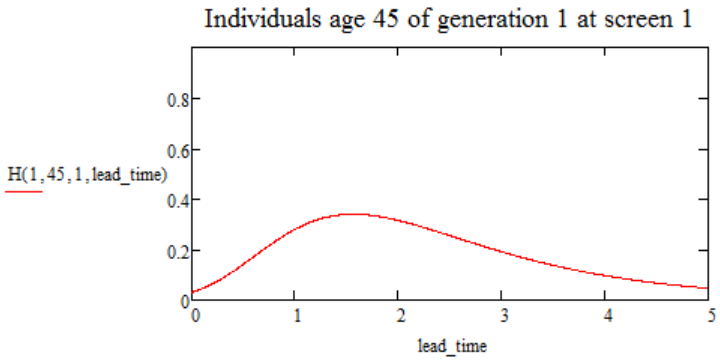
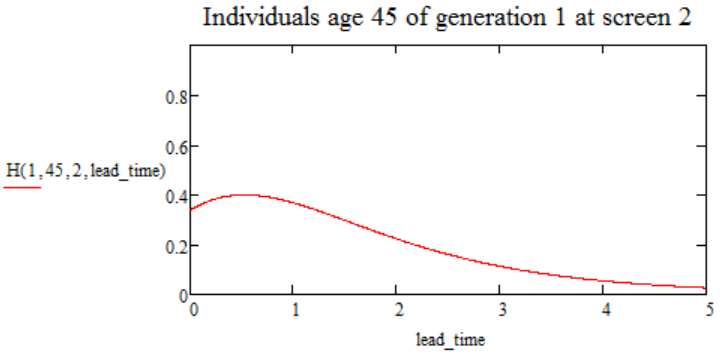
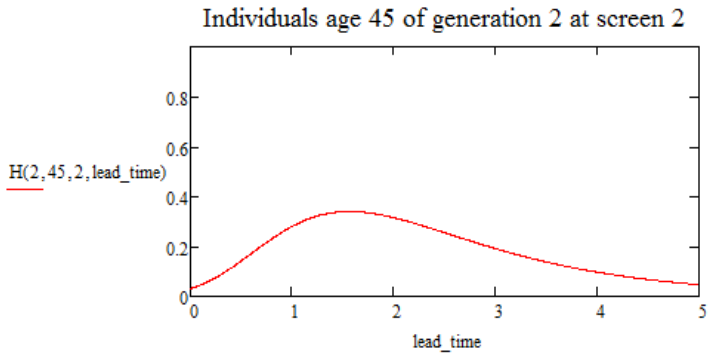


Figure 11b



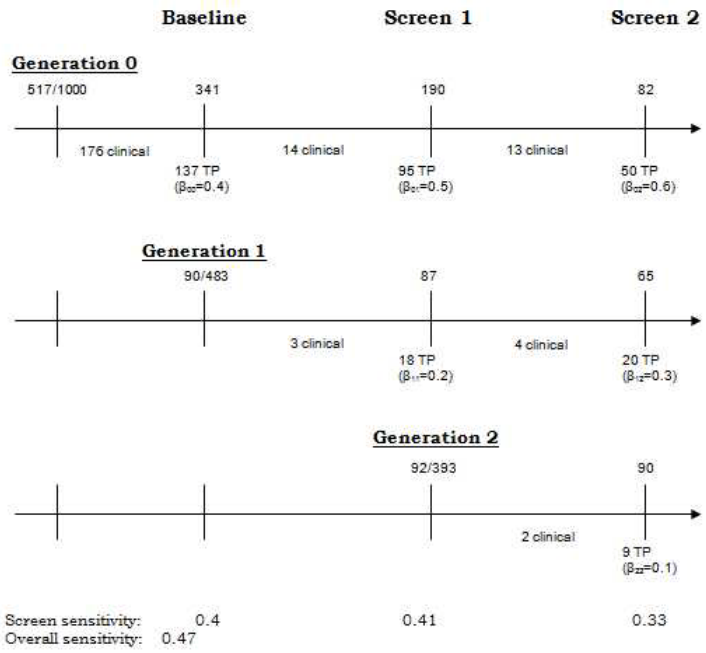
Generation 2:

Figure 12



The summarized simulation results are given in Figure 13.

Figure 13: Summary of the lung cancer simulation results



MathCAD's numerical integration algorithms make successive estimates of the value of the integral and return a value when the two most recent estimates differ by less than the value of the built-in tolerance parameter, TOL. Unfortunately due to computer limitations we cannot use a TOL that is smaller than 0.01 in order to maximize the log likelihood, and therefore the results that we have obtained for the MLE's are restricted to TOL=0.01. On the other hand we can compute the information matrix using TOL=0.001, just that it takes a very long time to execute.

The results obtained after maximizing the log likelihood function are:

var = 0.27161, m = 5.15056, a = - 10.99997, b = 0.15542, c = 564.20607,

β_{00} = 0.17862 and τ = 0.29086, giving us a maximum log-likelihood of – 2806.39, compared to the value of – 2936.85 at the original parameter values.

The covariance matrix is obtained by inverting the Information matrix (negative of the expectation of the Hessian matrix). The standard error for each parameter can then be obtained by square rooting the elements of the main diagonal. We can then use this s.e. in order to determine a 95% CI for the parameter values and then check if our MLEs are belonging to these intervals.

These findings are very important in the context of determining the optimal screening regimens for detecting the different types of cancer. Our model has the flexibility of being able to be applied to a wide variety of progressive disease models, due to its general format. Once we determine the possible shape of the incidence function, sensitivity function and preclinical time distribution, we substitute them in our model and maximize the log likelihood. By doing so, we determine the estimates of the parameters that define the functions and are able to determine the lead time distribution for the specified screening regimen. The methods used until now to determine the lead time distribution have been quite limited by the inflexibility of choosing the more appropriate functions for the corresponding settings. Our model on the other hand opens the door to

obtain much more reliable results that are essential in determining the efficacy of the tested screening programs.

Chapter III: Application of the Recurrence Time Model on the Screening Arm of the PLCO Colorectal Cancer Trial

III.1 Characteristics of the Model Applied to the PLCO CRC Data

Colorectal cancer is the third leading cause of cancer-related deaths in the United States when men and women are considered separately, and the second leading cause when both sexes are combined. It is estimated that about 141,210 people will be diagnosed with colorectal cancer and about 49,380 will die because of it during 2011 (American Cancer Society 2011). The majority of these cancers and deaths could be prevented by applying existing knowledge about cancer prevention and by efficiently using the established screening tests.

In Chapter 1 Section 4 we introduced the PLCO screening trial and we described the PLCO trial for colorectal cancer (CRC). This chapter will focus on applying the model developed in Chapter 2 on a data set that is comprised of 16.312 patients ages 60 to 75 at the baseline screen, from the PLCO CRC screening trial, who received the initial screen (at baseline). Table 2 presents the patient age distribution at baseline (AgeN) and the number of patients who had a screen detected cancer either at baseline (N1), at the 3 years follow-up (N3), or during the interval between baseline and the 3 year follow-up (N2).

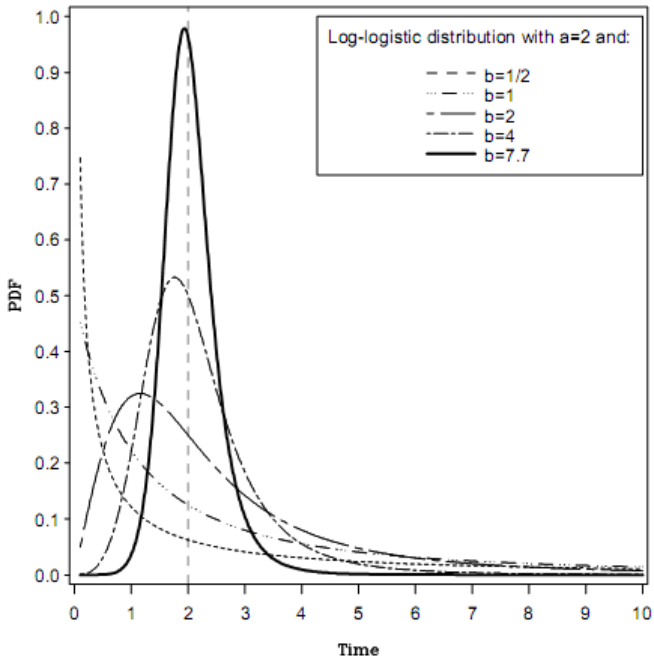
Table 2: PLCO CRC age distribution (AgeN), and number of cases diagnosed at baseline (N1), during the interval between baseline and the follow-up screen (N2), and at first follow-up screen (N3)

age_t0	AgeN	N1	N2	N3
60	1220	4	4	2
61	1316	3	3	0
62	1447	5	1	1
63	1404	4	2	2
64	1465	6	1	1
65	1323	4	4	0
66	1335	6	2	1
67	1168	2	1	0
68	1078	5	4	0
69	1003	0	4	0
70	869	1	1	2
71	792	2	3	2
72	704	5	6	0
73	634	5	1	0
74	511	4	4	1
75	43	1	0	0
Total	16312	57	41	12

In order to approximate the age-specific colorectal preclinical cancer incidence one can fit an exponential distribution of the form $\frac{c}{1000} \cdot e^{dx}$.

For the preclinical time distribution *q(t)*, we use a log-logistic distribution with scale parameter a > 0 and shape parameter b > 0. Figure 14 present different log-logistic distributional forms, with the bolded one corresponding to a distribution with mean=2.057, mode=1.933, median=a=2 and variance=0.251.

Figure 14 Log-logistic distribution with median 2 and various shape parameters



By using the log-logistic distribution instead of the log-normal distribution we keep very similar distributional forms with the advantage of increasing the speed of computation.

For sensitivity of screen *j* among generation *i*, where i= $\overline{0,1}$ and j= $\overline{i,1}$, define:

$$\beta_{ij}(\tau) = 1.5^{-i} \cdot (1 - \tau^{j+1}). \quad (29)$$

This sensitivity function, while still generation- and screen- dependent, and thus not constant throughout the entire screening process, differs from the sensitivity function in the simulation study previously presented. It is limited to a single parameter to ensure identifiability across the two screens being modeled.

Here Δ=3, since the time interval between the first screen at baseline and the first follow-up screen is 3 years.

Therefore the formula for the log-likelihood in this case is:

$$LL(a,b,c,d,\tau) = \sum_{age=60}^{75} [\, N1_{age} \cdot ln\left(D_{0,age,0}\right) + N2_{age} \cdot ln\left(C_{0,age,1} + C_{1,age,1}\right) + N3_{age} \cdot ln\left(D_{0,age,1} + D_{1,age,1}\right) + \left(N_{age} - \sum_{i=1}^{3} Ni_{age}\right) \cdot ln\left(1 - D_{0,age,0} - C_{0,age,1} - C_{1,age,1} - D_{0,age,1} - D_{1,age,1}\right)] \quad (30)$$

where the functions involved in this expression have been described in Chapter 2.

III.2 Results

After maximizing the log-likelihood presented in formula (30), the following parameter estimates have been obtained:

Table 3: PLCO CRC: parameter estimates and 95% confidence intervals

	Estimated value	95% CI
a	2.3098	(0, 37.947)
b	2.3485	(0, 28.947)
c	6.3803	(0, 128.094)
d	0.0558	(0, 0.216)
τ	0.4978	(0.384, 0.612)

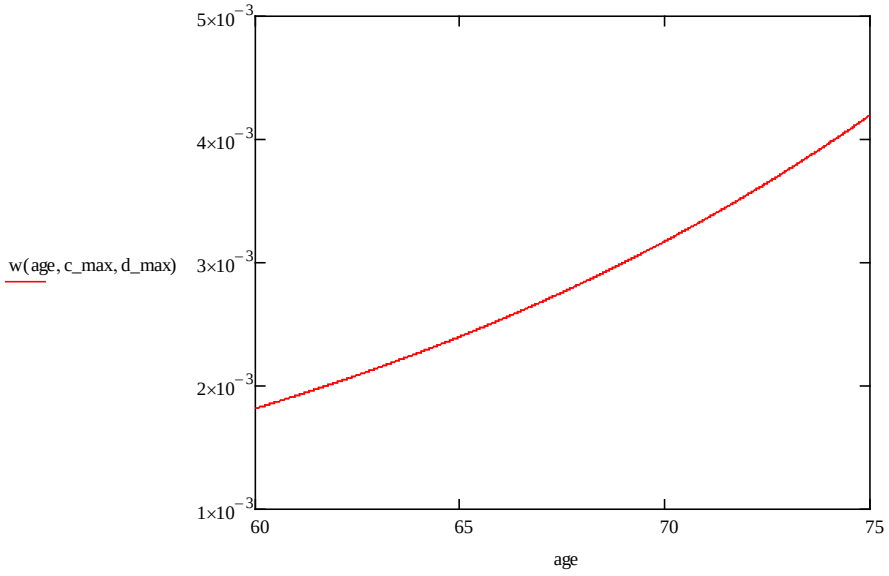
with the corresponding correlation matrix of:

$$\text{CorrM} = \begin{pmatrix} 1 & 0.999479571 & -0.975953171 & 0.917349136 & 0.39989327 \\ 0.999479571 & 1 & -0.973462994 & 0.913559406 & 0.39087325 \\ -0.975953171 & -0.973462994 & 1 & -0.981949234 & -0.389676355 \\ 0.917349136 & 0.913559406 & -0.981949234 & 1 & 0.364660346 \\ 0.39989327 & 0.39087325 & -0.389676355 & 0.364660346 & 1 \end{pmatrix}$$

Therefore we can derive estimates and confidence intervals for the parameters of the model from the actual data. Some of the confidence intervals are wide and the high collinearity between the parameters *a* through *d* may explain the reason behind it.

The graph of the exponential preclinical incidence function based on estimated values for *c* and *d* is presented in Figure 15.

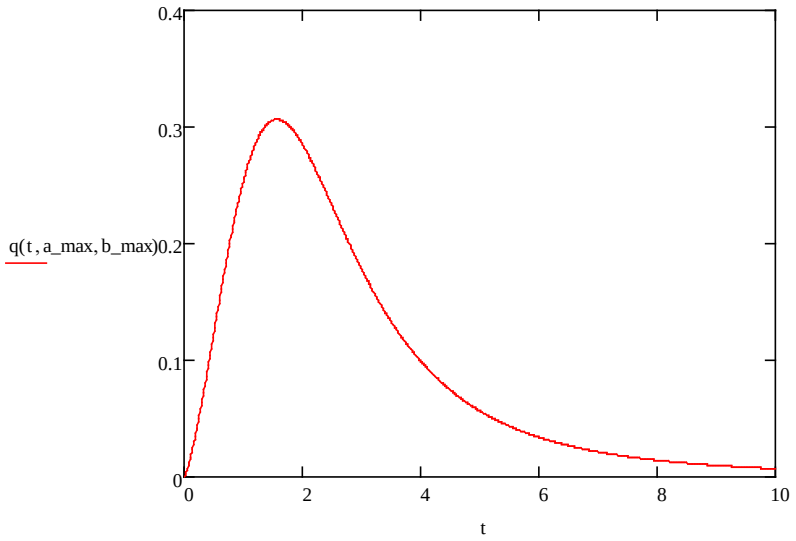
Figure 15: PLCO CRC estimated exponential preclinical incidence function



From here we can see that the preclinical colorectal cancer incidence rate for individuals age 80 is 5.34 times higher than for individuals age 50. In Table 6.11 of the SEER Cancer Statistics Review 1975-2008 on colorectal cancer the incidence rate for individuals ages 75-79 is 6.2 times higher than the incidence rate for individuals ages 50-54. This tells us that our result is consistent with the observed clinical incidence rates and thus much more realistic than the assumption of constant incidence for all age groups.

The graph of the preclinical time distribution based on the estimates of *a* and *b* is presented in Figure 16.

Figure 16: PLCO CRC estimated log-logistic preclinical sojourn-time distribution



From here we can see that the mean time an individual spends in the preclinical stage is of 3.1757 years, with a 95% CI = (0,12.3). The percentage of cases that have a preclinical duration of less than 1 year about 12%; of less than 2 years, 42%; of less than 3 years, 65%. From this last value we expect to screen more than (100-65) % = 35% of the cancers from 3 years to the next follow-up screen. The percentage of cases with preclinical duration less than 5 years is 86%. The PLCO colon cancer study is looking at 3 and 5 years follow-up screens, which means, based on our results, that the study has a very high probability of screen detection.

The value of $\hat{\tau}$ indicates that the estimated screening sensitivity for generation 0, baseline screen, is 0.5021, first follow-up screen is 0.7521, and for generation 1 first screen is 0.5014.

For illustration purposes we present the lead time distributions for 60 year olds (Figure 17, 18 and 19).

Figure 17

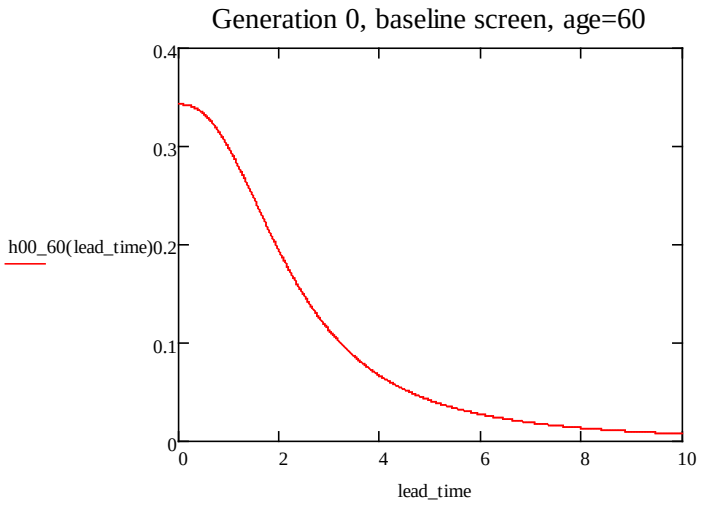


Figure 18

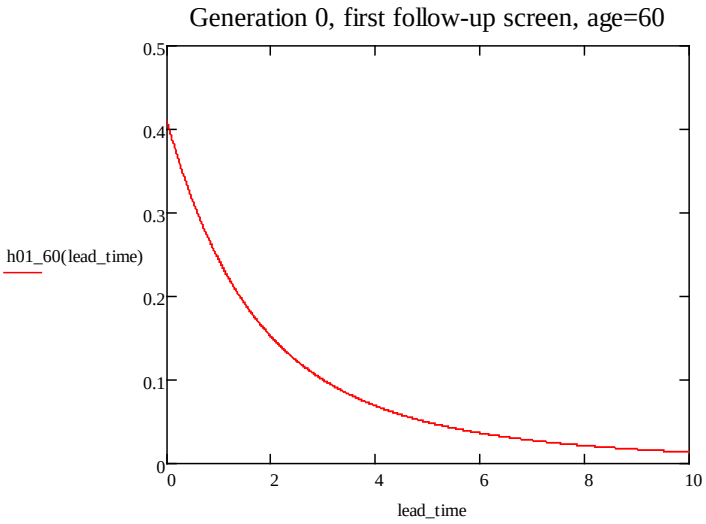
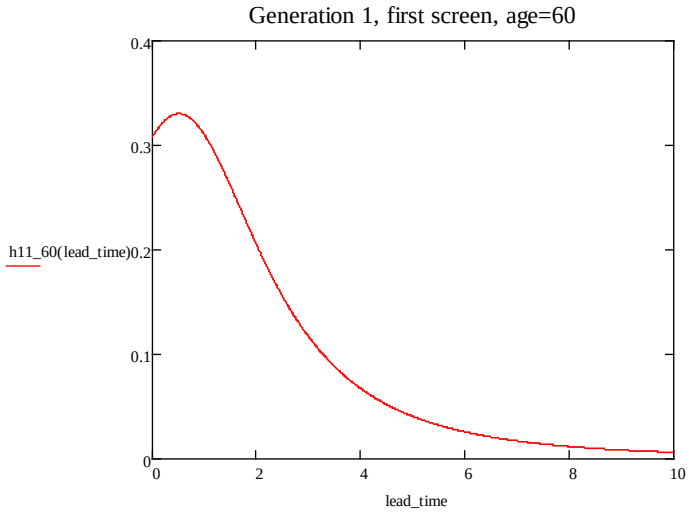


Figure 19



The lead time distributions for individuals 60-75 are very similar, as shown in tables 3 and 4.

Table 4: Mean lead time for individuals detected by screening, by age, generation and screen

Age	Mean Lead Time Generation 0, screen 0	Mean Lead Time Generation 0, screen 1	Mean Lead Time Generation 1, screen 1	Weighted mean lead time at screen 1
60	2.86916	4.14549	2.64058	4.10803
61	2.86945	4.14635	2.64058	4.10678
62	2.86971	4.14715	2.64058	4.10535
63	2.86996	4.14791	2.64058	4.10375
64	2.87019	4.14861	2.64058	4.10198
65	2.87041	4.14926	2.64058	4.10001
66	2.87061	4.14987	2.64058	4.09788
67	2.87079	4.15044	2.64058	4.09555
68	2.87097	4.15096	2.64058	4.09302
69	2.87113	4.15145	2.64058	4.09030
70	2.87128	4.15191	2.64058	4.08737
71	2.87142	4.15233	2.64058	4.08424
72	2.87155	4.15272	2.64058	4.08089
73	2.87167	4.15308	2.64058	4.07732
74	2.87178	4.15342	2.64058	4.07352
75	2.87188	4.15373	2.64058	4.06949

The mean lead time for those individuals detected by screening, is broken down in Table 3, by generation, screen and age, and also we present the weighted overall mean lead time for both generations at screen 1. The overall mean lead time across all ages of all diseased individuals detected by the entire screening program is 2.982 years and the mean lead time resulting from the entire screening program for all individuals who are susceptible to detection sometime during the screening program, across all ages, is 0.705 years.

Table 5: Mean lead time for all diseased individuals susceptible for detection or detected by the screening program

Age	Mean lead time for all diseased individuals susceptible to detection at baseline	Mean lead time for all diseased individuals susceptible to detection at the first follow-up screen	Mean lead time of all diseased individuals detected by the entire screening program	Mean lead time of all individuals susceptible for detection sometime during the screening program
60	1.66938	0.25348	2.98109	0.70317
61	1.66955	0.25377	2.98125	0.70339
62	1.66971	0.25407	2.98138	0.70361
63	1.66986	0.25438	2.98149	0.70383
64	1.66999	0.25470	2.98159	0.70406
65	1.67012	0.25503	2.98166	0.70430
66	1.67024	0.25538	2.98171	0.70454
67	1.67035	0.25574	2.98175	0.70480
68	1.67045	0.25612	2.98177	0.70507
69	1.67055	0.25652	2.98177	0.70534
70	1.67064	0.25693	2.98176	0.70563
71	1.67072	0.25737	2.98172	0.70593
72	1.67080	0.25782	2.98167	0.70624
73	1.67087	0.25830	2.98161	0.70657
74	1.67094	0.25880	2.98153	0.70691
75	1.67100	0.25933	2.98143	0.70727

The formulas for all these lead times and for the weighted averages have been derived in the previous chapter.

Chapter IV: Summary and Conclusion

Evidence from several studies shows that screening for CRC, detecting and then removing precancerous adenomatous polyps can reduce its incidence and mortality (Pignone et al., 2002). But important factors in CRC detection are methods of screening, screening frequency and screening age. As we have just seen in this work, screening age plays a major role in CRC preclinical incidence. In previous articles and computations, preclinical incidence has been considered to be uniform over different ages, assumption that is most definitely not valid. As the preclinical incidence plays a major role in computing the average lead time by age and overall, we can see how it can affect the final conclusion about the efficiency of screening in the context of CRC and other progressive disease models.

Creating the correct screening regimen, implementing it, extracting the results and then applying the most efficient model to analyze the data, those are all steps that are not only important to ensure the success of the screening regimen, but the implications here are major, from saving lives to also being cost-effective. One would want to know what the cost-effectiveness of cancer screening by different methods is. Or what is the incremental cost-effectiveness of continuing screening after the age of 80 or 75 compared to stopping at the age of 70. Or what is the incremental cost-effectiveness of starting screening before the age of 50, like at 45 or 40. As we can see these are all questions where the age of screening at baseline plays a major role not only from the statistical and therefore health

point of view, but also from the financial one. The present work is opening the door to progressive disease screening models where preclinical incidence is a function of age at baseline and where one has the flexibility to choose the best preclinical time distribution in the context of the corresponding progressive disease under screening. Sensitivity can also be adjusted as a function, to include both, the generation and screen information. So far in previous publications this also has not been the case, the unrealistic assumption being that sensitivity is a constant throughout the screening process.

Fitting complex models such as the one presented here is challenging, and depends upon the computer speed and capability of computation. Tuning approximation algorithms to the exact computation capabilities was essential. Eventually we hope to convert the model to a lower level programming language such as C++, so that fine control of approximation and root finding algorithms can be used to create efficient computation of likelihood maxima and first and second derivatives.

What the formulas that have been presented do not account for is overdiagnosis. Overdiagnosis needs to be modeled by including mortality, both from the cancer and all causes, in the model. This is a work in progress which means that the already complex formulas are going to be enhanced even further in the near future, to also account for overdiagnosis.

Bibliography

Albert A., Gertman P.M., Louis T.A. (1978). Screening for the early detection of cancer I. The temporal natural history of a progressive disease state. Math Biosci 40, 1-59.

Albert A., Gertman P.M., Louis T.A., Liu S. (1978). Screening for the early detection of cancer II. The impact of screening on natural history of the disease. Math Biosci 40, 61-109.

American Cancer Society (2010). Cancer facts and Figures. Atlanta, GA: American Cancer Society (http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-026238.pdf)

American Cancer Society (2011). Colorectal Cancer Facts & Figures 2011-2013. Atlanta, GA: American Cancer Society (http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-028323.pdf)

Begg C.B., McNeil B.J. (1988). Assessment of radiologic tests: control of bias and other design considerations. Radiology*;* 167:565-569.

Brookmeyer R., Day N.E. (1987). Two-stage models for the analysis of cancer screening data. Biometrics ; 43(3): 657-669.

Chen T.H.H., Kuo H.S., Yen M.F., Lai M.S., Tabar L., Duffy S.W. (2000) Estimation of Sojourn Time in Chronic Disease Screening without Data on Interval Cases Biometrics; Vol. 56, No. 1, pp. 167-172

Church T.R. (1999). A Novel Form of Ascertainment Bias in Case-Control Studies of Cancer Screening. J Clin Epidemiol; 52(9): 837-847.

Cole P., Morrison A.S. (1980). Basic issues in population screening for cancer. J Natl Cancer Inst, 64: 1263 -1272.

Commission on Chronic Illness (CCI). (1957). Chronic illness in the United States. Vol I. Prevention of chronic illness. Cambridge, MA: Harvard University Press, 1:45.

Cronin K.A., Weed D.L., Connor R.J., Prorok P.C. (1998). Case-control studies of cancer screening: Theory and practice. J Natl Cancer Inst 90: 498–504.

Cuzick J. (1999), Screening for cancer: future potential. Eur. J. Cancer 35: 685–692.

Cox D.R., Miller H.D. (1965). The Theory of Stochastic Processes, Chapman: London.

Day N.E. (1989). Quantitative approaches to the evaluation of screening programs. World J Surg 13: 3–8.

Day N.E., Walter S.D. (1984). Simplified models of screening for chronic disease: Estimation procedures from mass screening programmes. Biometrics 40: 1-13.

Day N.E., Williams D.R.R., Khaw K.T. (1989). Breast cancer screening programmes: the development of a monitoring and evaluation system. Br J Cancer ; 59:954-958.

Duffy S.W., Cuzick J., Tabar L., Vitak B., Chen T.H.H., Yen M.F., Smith R.A. (2002). Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. Appl Statist; 51(2): 235-243.

Duffy S.W., Nagtegaal I.D., Wallis M. et al. (2008). Correcting for lead time and length bias in estimating the effect of screen-detection on cancer survival. Am J Epidemiol; 168: 98–104.

Eddy D.M. (1980). Screening for cancer: theory, analysis, and design / Englewood Cliffs, N.J.: Prentice-Hall.

Feinleib M. (1967). The stable disease model (Abstract), Biometrics 23, No.2: 1299.

Henschke C.I., McCauley D.I., Yankelevitz D.F., et al. (1999). Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet. Jul 10; 354(9173): 99-105.

Henschke C.I., Yankelevitz D.F., Libby D.M., Pasmantier M.W., Smith J.P., Miettinen O.S.; International Early Lung Cancer Action Program Investigators. (2006). Survival of patients with stage I lung cancer detected on CT screening. N Engl J Med.; 355 (17): 1763-71.

Hutchinson G.B., Shapiro S. (1968). Lead time gained by diagnostic screening for breast cancer. J National Cancer Institute; 41: 665-681.

Jemal A., Siegel R., Ward E., Hao Y., Xu J., Murray T., Thun M.J. (2008). Cancer Statistics, 2008. CA Cancer J Clin 58: 71-96.

Kadafar K., Prorok P.C. (1994). A data-analytic approach for estimating lead time and screening benefit based on survival curves in randomized cancer screening trials. Statist. Med.; 13, 569-586.

Kafadar K., Prorok P.C. (1997). Estimating the difference in location parameters of two survival curves with applications to screening. J Stat Planning and Inference; 57(2), 165-179.

Lee C.H. (2002). Screening mammography: proven benefit, continued controversy. Radiol Clin North Am; 49: 395-407.

Lee J.M. (1993). Screening and informed consent. N Engl J Med; 328: 438-440.

Loeve F., Boer R., van Oortmarssen G.J., van Ballegooijen M., Habbema J.D. (1999). The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res; 32:13–33.

Louis T.A., Arthur A., Heghinian S. (1978) Screening for the early detection of cancer III. Estimation of disease natural history. Math Biosci 40, 111-144.

Lu Y., Fang J.Q. (2003). Advanced Medical Statistics, Published by World Scientific

Marcus P.M., Bergstralh E.J., Fagerstrom R.M., Williams D.E., Fontana R., Taylor W.F., Prorok P.C. (2000). Lung Cancer Mortality in the Mayo Lung Project: Impact of Extended Follow-up. J. Natl. Cancer Inst.; 92: 1308-1316.

Marcus P.M., Bergstralh E.J., Zweig M.H., Harris A., Offord K.P., Fontana R.S. (2006). Extended Lung Cancer Incidence Follow-up in the Mayo Lung Project and Overdiagnosis**.** J Natl Cancer Inst, June; 98: 748 - 756.

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. (2010) CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomized trials. BMJ 23;340:c869. doi: 10.1136/bmj.c869

Morabia A., Zhang F. F. (2004). History of medical screening: from concepts to action, Postgrad Med J, 80: 463-469.

Murthy V.H., Krumbolz H.M., Gross C.P. (2004). Participation in cancer clinical trials, race-, sex-, and age-based disparities. JAMA; 291(22):2720–6.

National Cancer Institute Fact Sheet 5.34 (2005), Colorectal Cancer Screening (http://www.cancer.gov/cancertopics/factsheet/detection/Fs5_34.pdf)

National Cancer Institute, Colorectal Cancer Research from the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: NCI Fact Sheet (2005) (http://www.cancer.gov/cancertopics/factsheet/detection/plco-colorectal)

Obuchowski N.A., Graham R.J., Baker M.E., Powell K.A. (2001). Ten criteria for effective screening: their application to multislice CT screening for pulmonary and colorectal cancers. AJR Am J Roentgenol, 176: 1357-1362.

Oken M.M., Marcus P.M., Hu P., Beck T.M., Hocking W., Kvale P.A., Cordes J., Riley T.R., Winslow S.D., Peace S., Levin D.L., Prorok P.C., Gohagan J.K. for the PLCO Project Team (2005). Baseline Chest Radiograph for Lung Cancer Detection in the Randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. J. Natl. Cancer Inst. 97 (24): 1832-1839.

Olsen A.H. et al. (2006). Overdiagnosis, Sojourn Time, and Sensitivity in the Copenhagen Mammography Screening Program. Breast Journal 12, no. 4 (July): 338-342.

Paci E. (2002). Lung cancer screening: the methodological debate. Lung Cancer; 38: S17-S21.

Paci E., Miccinesi G., Puliti D. et al. (2006). Estimate of overdiagnosis of breast cancer due to mammography after adjustment for lead time. A service screening study in Italy. Breast Cancer Res; 8:R68.

Paci E. (2007). Observational, one-arm studies and randomized population-based trials for evaluation of the efficacy of lung cancer screening, J Thorac Oncol.; 2: Suppl 1, 45-46.

Paskett E., DeGraffinfreid C., Tatum C., Margitic S. (1996). The recruitment of African-Americans to cancer prevention and control studies. Prev Med; 25: 547–53.

Pignone MP, Rich S, Teutsch S, Berg A, Lohr K. (2002) Screening for colorectal cancer in adults at average risk: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med.; 137: 132-141.

PLCO – Manual of Operations and Procedures. November 01, 1999.

Prorok P.C. (1976). The theory of periodic screening. I. Lead time and proportion detected. Adv Appl Probl; 8:127-143.

Prorok P.C. (1976). The theory of periodic screening. II. Doubly bounded recurrence times and mean lead time and detection probability estimation. Adv Appl Probl; 8:460-476.

Prorok P.C. (1982). Bounded Recurrence Times and Lead Time in the Design of a Repetitive Screening Program. J Appl Probl; Vol. 19, No. 1 (Mar., 1982): 10-1.

Prorok P.C. et al. (2000). Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Controlled Clinical Trials; 21(Issue6, Spplement1): 273S–309S.

SEER Cancer Statistics Review 1975-2008, table 6.11
http://www.seer.cancer.gov/csr/1975_2008/results_merged/sect_06_colon_rectum.pdf

Shapiro S., Goldberg J.D., Hutchison G.B. (1974). Lead time in breast cancer detection and implications for periodicity of screening. Am. J. Epidemiol. 100(5): 357-366.

Shen Y., Zelen M. (1999). Parametric estimation procedures for screening programmes: stable and nonstable disease models for multimodality case finding. Biometrika; 86(3): 503-515.

Shen Y., Zelen M. (2001). Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. J Clinical Oncology; 19(15): 3490-3499.

Shen Y., Zelen M. (2005). Robust modeling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. Biostat; 6(4): 604-614.

Sobue T. (2000). A case-control study for evaluating lung cancer screening in Japan. Cancer 89(S11): 2392-2396.

Spix C., Michaelis J., Berthold F., Erttmann R., Sander J., Schilling F.H. (2003). Lead-time and overdiagnosis estimation in neuroblastoma screening. Statist. Med.; 22: 2877-2892.

U.S. National Center for Health Statistics: http://www.cdc.gov/nchs/

Walter S.D., Day N.E. (1983). Estimation of the duration of a preclinical disease state using screening data. Am J Epidemiol 118: 856-886.

Walter S.D., Stitt L.W. (1987). Evaluating the survival of cancer cases detected by screening. Statistics in Medicine; 6: 885–900.

Weissfeld JL, Schoen RE, et al. (2005) "Flexible Sigmoidoscopy in the PLCO Cancer Screening Trial: Results from the Baseline Screening Examination of a Randomized Trial." Journal of the National Cancer Institute. Vol. 97, No. 13. July 6.

Zelen M. (1976). Theory of early detection of breast cancer in the general population. In: Heuson JC, Mattheiem WH, Rozenzweig M, eds. Breast cancer trends in research and treatment. New York: Raven, 287-300.

Zelen M., Feinleib M. (1969). On the theory of screening for chronic diseases. Biometrika; 56: 601-614.

Zelen M. (1993). Optimal scheduling of examinations for the early detection of disease. Biometrika; 80(2): 279-293.

Zelen M. (2004). Forward and Backward Recurrence Times and Length Biased Sampling: Age Specific Models. Lifetime Data Analysis; 10(4): 325-334.