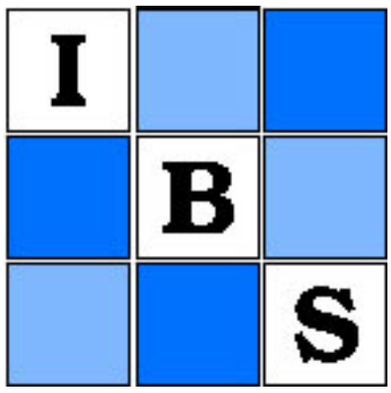

BIOMETRICS 43, 127–144
March 1987

The Natural History of Lung Cancer in a Periodically Screened Population

Betty J. Flehinger

IBM Thomas J. Watson Research Center,
P.O. Box 218, Yorktown Heights, New York 10598, U.S.A.

and

Marek Kimmel*

Memorial Sloan-Kettering Cancer Center,
1275 York Avenue, New York, New York 10021, U.S.A.

SUMMARY

A mathematical model of the progression kinetics of lung cancer in a periodically screened population is proposed and data collected by the Memorial Sloan-Kettering Cancer Center in New York are used for parameter estimation. It is assumed that the development of adenocarcinoma of lung is a stochastic process with two stages, early and advanced, characterized by mean times, detection probabilities, and cure probabilities. Confidence regions of these parameters are estimated using a number of novel techniques. It is found, surprisingly, that the mean duration of the early stage is at least 4 years, the detectability less than .2, and the curability less than .5. These estimates imply that annual radiographic screening from age 45 to 80 might decrease mortality from adenocarcinoma of lung by something less than 20%.

1. Introduction

At the present time, there is controversy about the advisability of routine periodic radiographic examinations of individuals at high risk for lung cancer. Although the American Cancer Society currently advises against any screening for lung cancer (ACS, 1980), many health professionals share the view that "the weight of evidence at this time supports the prudent medical practitioner who recommends regular screening of the asymptomatic person at high risk of lung cancer" (Melamed and Flehinger, 1984). It is estimated that more than 125,000 men and women in the United States will die of the disease this year; 90% of them will survive less than 2 years after diagnosis. It is well established that survival probability is very closely related to stage at treatment. Patients with asymptomatic, localized (AJCC Stage I) disease have a high probability of surviving after surgical resection. The vast majority of those who seek medical attention because of symptoms have metastasis to regional lymph nodes or distant sites before detection. They have low probability of surviving regardless of the type of treatment.

The only screening techniques known to detect Stage I lung cancer are chest radiography and sputum cytology. For the past 10 years, a clinical trial utilizing these techniques has been carried out by the Memorial Sloan-Kettering Cancer Center as part of the National

---

* This author is a Visiting Investigator from Institute of Automation, Silesian Technical University, Gliwice, Poland.

Key words: Detection; Lung cancer; Mortality; Screening.

Cancer Institute Cooperative Early Lung Cancer Group and the results have been reported (Melamed et al., 1977, 1981, 1984; Flehinger et al., 1983, 1984; Heelan et al., 1984). This trial was designed to assess the value of 4-monthly sputum cytology examinations as adjuncts to annual chest Xrays in reducing lung cancer mortality in male smokers over the age of 45. Ten thousand volunteers were randomly assigned to one of two screening schedules: the "Xray only" group of 5072 men were offered postero-anterior and lateral chest Xrays annually; the "dual screen" group of 4968 men were asked to submit sputum specimens every 4 months in addition to the annual radiographic examinations. Active screening was continued until all participants had been enrolled for at least 5 years. Any positive or suspicious findings from either screening test were carefully investigated and surgery was offered whenever resection of lung cancer seemed feasible.

The results of this research may be summarized as follows:

(i) Sputum cytology is effective in the early detection of slow-growing squamous-cell lung cancer, which accounted for less than one-third of the cases. In the absence of cytologic screening, the annual chest Xray finds these cancers somewhat later but while still localized and resectable. Patients diagnosed in Stage I through either detection mode have a high probability of surviving their lung cancer. No statistically significant effect on mortality can be ascribed to the addition of cytologic screening to a program of annual chest Xrays.

(ii) Half of all the lung cancers found during the screening period have been adenocarcinoma and early detection of this type was restricted to chest radiography. These tumors tend to be peripheral and can often be identified while they are very small; 72% were found by routine screening Xrays. Fifteen percent of the cases were oat cell cancer, almost all found through symptoms in advanced stage with poor survival.

(iii) Since all participants were offered annual chest Xrays, no direct evaluation of the effect of annual radiographic screening on lung cancer mortality is possible. However, it is noteworthy that 40% of all lung cancers were detected in Stage I and 70–80% of the patients whose cancers were treated in Stage I do not die of that disease. Overall, the 5-year survival of patients who developed lung cancer while enrolled in the screening program has been 35%, contrasted with less than 10% for the United States as a whole.

There remains a major question that must be answered. How strong is the effect of lead time, length bias, and overdiagnosis bias on these survival statistics? Phrased another way, if the Stage I cancers had not been detected and resected, what course would they have followed? Would most have progressed to symptomatic, disseminated disease followed quickly by death? Or might many have grown so slowly that the patients would have remained unaware of their presence until they died of other causes?

Answers to these questions may be obtained from a characterization of the natural history of lung cancer as a stochastic process. In this paper a model is described which embodies most of the major essential features of the progression kinetics of lung cancer but which is as simple as possible. It is assumed that a susceptible individual goes through a sequence of transitions: from a cancer-free state to an asymptomatic early cancer state to an asymptomatic advanced cancer state to a symptomatic state followed immediately by death. Each state is characterized by a distribution function of time spent in the state, a detection probability, and a cure probability. In addition, there is a distribution function of time to death from other causes. An individual who is not screened will never know about his cancer if the time to symptoms is longer than the time to death from other causes. The most important parameters in evaluating the efficacy of screening are the probability that an individual is susceptible to lung cancer, and the mean duration, detectability, and curability of the early stage.

The lung cancer screening program can be modeled by a periodic grid superimposed on the preceding stochastic process. The model provides that, at the grid points, all cases found in the asymptomatic states are treated and may or may not be cured. Data collected during the 10 years of the screening program are used to estimate the parameters of the stochastic process.

For simplicity, we restrict ourselves to analysis of the progression kinetics of adenocarcinoma, since that cell type was detected in Stage I only by chest Xray and it constituted the major part of detected lung cancers.

The relevant data collected in this annual screening program consist of the following: (i) the age at first screen of each of 10,040 participants; (ii) the number of years of screening of each participant; (iii) the date of each death from any cause of any participant; (iv) a characterization of each case of adenocarcinoma by (a) detected by screening or in the interval between screens, (b) age at detection, (c) screening year at detection, (d) stage at detection, and (e) survival time after detection; (v) for cases of adenocarcinoma detected by Xray after the initial screen, retrospective evidence of the presence or absence of cancer in the previous year's roentgenogram. During the screening period, 291 lung cancers were confirmed, of which 140 were adenocarcinoma, 67 found in early stage (AJCC Stage I) and 73 in late stage (AJCC Stages II and III). We must recognize that there are key quantities in the model that by their very nature can never be observed, such as age of onset of asymptomatic early-stage disease, and time spent in that early stage. Also, it is impossible to determine whether or not an individual belongs to the susceptible subgroup if he does not develop diagnosed cancer.

In this paper, based on the results of the study, it is assumed that all early-stage cancers are asymptomatic, while the late-stage cancers are either asymptomatic or symptomatic. All asymptomatic stages correspond to what is usually called "preclinical disease" (cf. Zelen and Feinleib, 1969). The "clinical" disease would correspond to the symptomatic phase of adenocarcinoma, which is usually very short and almost inevitably fatal.

This paper is organized as follows. In Section 2 we present the basic assumptions of the mathematical model. In Section 3 the parameters of the model are discussed. In Section 4 methods of finding confidence regions for the early-stage parameters are given. The results are presented and discussed in Section 5. These results suggest that the early stage is of long duration and detection probability is low. The effect of screening on mortality from adenocarcinoma is discussed in Section 6. Section 7 is a summary of results and discussion.

Many papers have been written concerning modeling of cancer progression and screening procedures. The pioneer work of Zelen and Feinleib (1969) concerns the estimation of the mean lead time (i.e., the time from screening detection to the point when symptoms would have appeared), as well as the relation between prevalence and incidence for a single-screen program. This approach has been generalized in several directions with application of the modeling techniques to data collected in screening programs for breast and cervix cancer (Prorok, 1976a, 1976b; Blumenson, 1976, 1977; Albert, Gertman, and Louis, 1978; Albert, Gertman, Louis, and Liu, 1978; Louis, Albert, and Heghinian, 1978; Dubin, 1981; Chen and Prorok, 1983; Chen, Prorok, and Graff, 1983; Walter and Day, 1983; Day and Walter, 1984). The classical clinical trial to evaluate a screening program which produced the breast cancer data that most modelers have tried to fit was carried out by Health Insurance Plan in New York (Shapiro, Goldberg, and Hutchison, 1974). To our knowledge, the only mathematical modeling study of respiratory cancer is related to the dynamics of lung cancer in the absence of screening (Manton and Stallard, 1982).

The model discussed in this paper has several features that distinguish it from any of the above. An attempt has been made to construct a comprehensive model of the natural history of lung cancer in the presence of a screening program. The concepts of age-dependent incidence and mortality are introduced and used in the estimation process. Cure

probability is defined, estimated, and utilized to project the impact of screening on mortality from lung cancer.

2. Mathematical Model of Cancer and Screening

2.1 *Assumptions*

(i) In the high-risk population volunteering for screening, a subgroup of the participants are susceptible to adenocarcinoma of lung. The probability that an individual belongs to this subgroup is ρ.

(ii) In the absence of screening and treatment, lung adenocarcinoma, after its onset, progresses through three stages: early, advanced asymptomatic, and symptomatic, followed immediately by cancer death (see Figure 1).

(iii) For an individual in the susceptible subgroup, the age of onset of the early stage, τ_0, is a random variable with a trapezoidal distribution $F_0(\cdot)$ and density

$$f_0(\tau_0) = \frac{2\alpha}{80 - R} + 2(1 - 2\alpha)\frac{\tau_0 - R}{(80 - R)^2}, \quad R \leqslant \tau_0 \leqslant 80.$$

Parameter R may assume values from 0 to 40 and α from 0 to 1.

(iv) The durations τ_1 and τ_2 of the early and advanced stages are independent exponential random variables with means μ_1 and μ_2, respectively. We denote their densities by $f_1(\cdot)$, $f_2(\cdot)$ and cumulative distributions by $F_1(\cdot)$, $F_2(\cdot)$. τ_1 and τ_2 are also independent of the age of the onset of disease, τ_0. The duration of the symptomatic stage is set equal to 0.

(v) A screening program consists of periodic examinations intended to detect the cancer. If an individual enters the program at age t, he is examined at ages $t_1, t_2, \ldots$, where $t_j = t + (j - 1)\Delta t$.

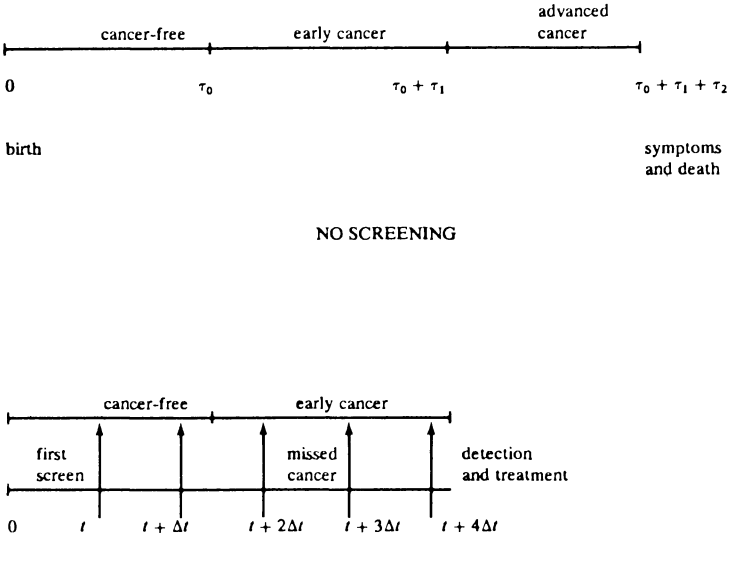

Figure 1. Sequence of events in the model of cancer and screening. In the case depicted, screenings begin before cancer is present. The cancer is missed by two screens, but detected while still in early stage.

(vi) Given the presence of early (advanced asymptomatic) cancer, a single examination detects it with probability p_1 (p_2). Detection on successive examinations is independent.

(vii) When a cancer is detected, screening is aborted and the patient is treated. The probability of "cure" for early stage is equal to c_1 and for advanced stage is set equal to 0. Cure is defined pragmatically in terms of the patient's survival after detection: If the patient is cured, his survival distribution is the same as if he had never had cancer; if he is not cured, his survival distribution is the same as if his cancer had not been detected through screening.

(viii) Members of the population are characterized by a density $q(\cdot)$ of age at initial screen and a family of distributions $H_t(\cdot)$ of time to death from causes other than adenocarcinoma of the lung for participants with initial screen at age t. The time of death from other causes is independent of onset of cancer, death from cancer, and screening procedures. However, these cancer-related events occur only on condition that they are not preceded by death from other causes and death from other causes occurs only if not preceded by cancer death.

2.2 *Remarks*

(i) Although the assumptions of this model may apply to a variety of populations, the estimated parameters will relate specifically to high-risk volunteers in New York City. Therefore, any conclusions will be affected by self-selection bias as well as conditions specific to the geographical area.

(ii) It is implicitly assumed that early-stage adenocarcinoma is always asymptomatic. In our experience, deviations from this generalization occur so rarely that their effect is negligible.

(iii) In this model, for simplicity, the time from symptoms to death is completely neglected, since that interval is rarely longer than a few months.

(iv) The trapezoidal distribution of age of onset was chosen as a convenient vehicle to incorporate the following known features of the natural history of lung cancer and of the Memorial Sloan-Kettering study: (a) The observed incidence rate increases with age. (b) The youngest age was 45 and some 45-year-olds had lung cancer detected at enrollment. (c) The oldest participants were under 80. In the expression of assumption (iii), R represents a lower bound for the age of onset and may assume values from 0 to 40. The upper bound for age of onset is taken as 80, since no data beyond that point are available. Parameter α controls the shape of the density function, in that $\alpha < \frac{1}{2}$ ($\alpha > \frac{1}{2}$) corresponds to an increasing (decreasing) density function. For all values of α, the hazard rate $f_0(t)/[1 - F_0(t)]$ increases as t increases from its minimum value, R, to its maximum, 80.

(v) Throughout this paper, any dependence of incidence, progression, or mortality on compliance with screening is neglected. The probabilities of detection p_1 and p_2 must be interpreted as probabilities that the individual complies with the screen *and* that his cancer is detected.

2.3 *Detection Probabilities*

The stochastic description of the screening program is built on the following probabilities defined for an individual enrolled at age t:

$E_k(t)$ is the probability that early adenocarcinoma is detected for the first time in the kth screen.

$R_k(t)$ is the probability that early adenocarcinoma is present and missed in the $(k-1)$st screen, and is detected for the first time in the kth screen.

We note that $E_k(t)$ may be expressed as the sum of probabilities of disjoint events. Event j consists of the following sequence: early cancer begins between screen $j-1$ and screen j; screens $j, j+1, \ldots, k-1$ fail to detect it; detection occurs at screen k. Thus, $E_k(t) = \rho e_k(t)$, where

$$e_k(t) = \sum_{j=1}^{k} P_{jk}(t)(1-p_1)^{k-j}p_1 \tag{2.1}$$

and $P_{jk}(t)$ is the probability that for a susceptible individual, the early stage begins between t_{j-1} and t_j and does not end before t_k. Based on our assumptions about age of onset and duration of early stage, we have

$$P_{jk}(t) = \frac{2\mu_1 e^{-t_k/\mu_1}}{(80-R)^2}\{[\alpha\cdot 80 - (1-\alpha)R](e^{t_j/\mu_1} - e^{t_{j-1}/\mu_1}) + (1-2\alpha)[(t_j-\mu_1)e^{t_j/\mu_1} - (t_{j-1}-\mu_1)e^{t_{j-1}/\mu_1}]\}, \tag{2.2}$$

where $t_0 = R$.

By analogous reasoning, we find

$$R_k(t) = \rho r_k(t) = \rho \sum_{j=1}^{k-1} P_{jk}(t)(1-p_1)^{k-j}p_1. \tag{2.3}$$

Now we can eliminate the conditioning on age at initial screen by integrating $E_k(t)$ and $R_k(t)$ over the density function $q(t)$; the corresponding unconditional functions are denoted E_k and R_k. Correspondingly, we have the unconditional probabilities for the susceptible subgroup denoted e_k and r_k.

2.4 *Survival Density of Detected Cancers*

We now express the density functions $g_1(\cdot)$ and $g_2(\cdot)$ with associated distributions $G_1(\cdot)$ and $G_2(\cdot)$ of survival time after detection of early or advanced asymptomatic cancers. Since it is assumed that the early cancer is cured with probability c_1 and survival time of those cases not cured has the distribution of the sum of exponential times in early and advanced stages, we have

$$g_1(x) = (1-c_1)\left(\frac{e^{-x/\mu_1} - e^{-x/\mu_2}}{\mu_1 - \mu_2}\right). \tag{2.4}$$

Advanced cancer is considered incurable, and survival after detection is exponential:

$$g_2(x) = \frac{e^{-x/\mu_2}}{\mu_2}. \tag{2.5}$$

2.5 *Population Mortality*

We seek an expression for $\psi_t(\tau)$, the probability that a member of the population enrolled for initial screen at age t and screened annually thereafter dies of adenocarcinoma of lung before age $t+\tau$:

$$\psi_t(\tau) = \int_0^{\tau} [1 - H_t(u)]\, d\varphi_t(u), \tag{2.6}$$

where $H_t(u)$ is the distribution function of time from enrollment to death from other causes and $\varphi_t(u)$ is the probability of adenocarcinoma death before $t+u$. We observe that $\varphi_t(u)$ may be expressed as the probability that an individual passes through the cancer sequence

of onset, progression, and death minus the probability that his cancer is detected in early stage and cured. Thus,

$$\varphi_t(u) = \rho(F_0 * F_1 * F_2)\Big|_t^{t+u} - c_1 \sum_{k\geq 1} E_k(t)(F_1 * F_2)\Big|_0^{t+u-t_k}. \tag{2.7}$$

We may integrate $\psi_t(\tau)$ over the density $q(t)$ of age at initial screen to obtain the distribution of time in study to adenocarcinoma death for the study population.

3. Parameters of the Model

The model defined by the preceding assumptions contains a large number of parameters. The methods of handling them in the analysis that follows may be divided into two categories: (a) variable assigned values, and (b) estimated values. Category (a) parameters are allowed to vary over broad ranges and are investigated for robustness. The essential features of this paper lie in estimating confidence regions of the parameters in category (b) and demonstrating their effect on the projected efficacy of screening programs.

Parameters in category (a) include R, the earliest possible age of onset of cancer, and α, the slope of the trapezoidal density function. These have been assigned values ranging from 0 to 40 and 0 to 1, respectively. Based on these assignments, parameters in category (b) were estimated and it was found that the estimates were sufficiently robust with respect to the variations.

In category (b) are found ρ, the probability that an individual belongs to the subgroup susceptible to adenocarcinoma of lung; μ_1, the mean time spent in the early stage, and μ_2, the mean time in the advanced asymptomatic stage; p_1, the probability of detection by screening in the early stage; c_1, the probability of cure for cancers detected in the early stage; $q(\cdot)$, the density function of age at initial screen; and $H_t(\cdot)$, the distribution functions of death from causes other than adenocarcinoma of lung. The density $q(\cdot)$ and distributions $H_t(\cdot)$ are estimated directly and nonparametrically from data collected in the screening program. Estimation of confidence regions for the other five parameters which are key to evaluating the efficacy of screening is discussed in the following sections.

4. Estimation Techniques

In this section, we consider methods of estimation of the important parameters of our model from statistics collected in the screening program. First, in Section 4.1 we describe the construction of confidence regions for μ_1, p_1, and ρ based on data about detection of early adenocarcinoma through screening. In Section 4.2 we describe a method for utilizing survival data to estimate μ_2 and c_1. In Section 4.3 we discuss data on mortality from adenocarcinoma of lung for groups defined by age at first screen. It is demonstrated that matching this data with theoretical mortality rates restricts the confidence regions for μ_1 and p_1. Numerical results are presented in Section 5.

4.1 *Confidence Regions for μ_1 and p_1*

Certain features of the data on detection of early adenocarcinoma have been used to build joint confidence regions for the mean time in early stage and the detection probability.

First, we considered the screening years in which the early cancers were detected. We computed the ratio of the number detected in the K_1 earliest screens (X_1) to the total number of early adenocarcinomas detected in K screens (Z_1). It is intuitively clear that a high value of this ratio indicates that either the mean time in early stage (μ_1) or the early-stage detectability (p_1) or both are large.

We note that the expectations of Z_1 and X_1 may be expressed in terms of the probabilities E_k defined in Section 2:

$$\mathrm{E}(Z_1) = N \sum_{k=1}^{K} E_k; \quad \mathrm{E}(X_1) = N \sum_{k=1}^{K_1} E_k, \tag{4.1}$$

where N is the population size. K_1 is selected to be approximately $K/2$, but may be modified to improve the accuracy of the estimation.

Second, we utilized data obtained in a special study of the radiographs of patients whose non–small-cell cancers were detected by screening after their initial screen (Heelan et al., 1984). Previous negative radiographs, when available, were reevaluated in light of the findings at detection. We computed the ratio of number of early adenocarcinomas (X_2) evident on previous Xray films to total number of early adenocarcinomas reexamined (Z_2). A high value of this ratio indicates a long mean time (μ_1) in the early stage and a low value of the early-stage detectability (p_1).

The expectations of Z_2 and X_2 are expressed in terms of the probabilities E_k and R_k defined in Section 2:

$$\mathrm{E}(Z_2) = Ns \sum_{k=2}^{K} E_k, \quad \mathrm{E}(X_2) = Ns \sum_{k=2}^{K} R_k, \tag{4.2}$$

where s is the proportion of early cases detected after the initial screen for which prior radiographs were available. (It is assumed that a random sample of prior radiographs was obtained.)

Note that E_k and R_k are functions of ρ, μ_1, p_1, α, and R. For the moment, we fix the values of α and R. Next, ρ is treated as a nuisance parameter and eliminated in the following way: Given values of μ_1 and p_1, the likelihood of Z_1 detected early cases in a population of size N in K screens expressed as a function of ρ is

$$L(\rho) = \left(\rho \sum_{k=1}^{K} e_k\right)^{Z_1} \left(1 - \rho \sum_{k=1}^{K} e_k\right)^{N-Z_1}. \tag{4.3}$$

Therefore, under the constraint $\rho \leqslant 1$, the maximum likelihood estimator is

$$\tilde{\rho} = \tilde{\rho}(\mu_1, p_1) = \min\left(\frac{Z_1/N}{\sum_{k=1}^{K} e_k}, 1\right). \tag{4.4}$$

Now we can proceed to substitute $\tilde{\rho}$ for ρ in the expressions for E_k and R_k, so that they depend only on μ_1 and p_1.

It is proved in the Appendix that the joint distribution of the random vector $(X_1/Z_1, X_2/Z_2)$ is asymptotically bivariate normal. The expectation and covariance matrix displayed there are functions of μ_1 and p_1 only. Furthermore, numerical studies have indicated that the sample sizes of our data provide good approximations to normality and to the calculated means and covariance matrix. Therefore, using the standard transformation to obtain a chi-square distributed statistic (with 2 degrees of freedom), we proceed to construct confidence regions for parameters μ_1 and p_1. We then vary α and R over a grid of values and obtain a confidence region for (μ_1, p_1) for each point on that grid.

4.2 *Estimation of μ_2 and c_1*

In Section 2, we derived expressions [eqs (2.4) and (2.5)] for the density of survival times after detection of adenocarcinoma in early and advanced asymptomatic stages in terms of the mean times, μ_1 and μ_2, and cure probability, c_1. The survival times were observed

during a limited time period in the screening study, during which some patients died of other causes. Survival after early-stage detection generated censored samples from a defective distribution, with the defect equal to the cure probability. Survival of advanced cases is governed by a proper exponential distribution. The likelihood functions for the survival times of a population of n_i individuals with early ($i = 1$) or advanced asymptomatic ($i = 2$) cancer measured from time of detection may be expressed as in Meeker (1984) and in unpublished work of Bandes and Nadas (IBM report TR22.1345, 1971):

$$L_i(c_i, \mu_i) = \binom{n_i}{n_{ia}} \prod_{m=1}^{n_{ia}} g_i(x_{im}; c_i, \mu_i)[1 - \Theta(x_{im})] \prod_{m=n_{ia}+1}^{n_i} \theta(y_{im})[1 - G_i(y_{im}; c_i, \mu_i)] \quad (4.5)$$

where n_{ia} is the number of individuals who died of their cancer and Θ and θ are the cumulative distribution function and density of the censoring time (due to death from other causes or limited follow-up). The x_{im} are times to cancer death and the y_{im} are censoring times. We note that c_2 is assumed equal to 0.

We first estimate μ_2 by maximizing the expression for L_2 by standard numerical procedures. The approximate asymptotic variance of the estimator is found in a standard way. We substitute the estimate of μ_2 in the expressions for g_1 and G_1 [eq. (2.4)] so that the likelihood L_1 is a function of μ_1 and c_1 only. Then we obtain confidence interval estimates of c_1 as a function of μ_1.

4.3 *Limits on μ_1 and p_1*

We further limit the confidence regions of (μ_1, p_1) by considering age-specific lung cancer (adenocarcinoma) mortality. First, we develop expressions for the expected values and variances of deaths in specified age groups. Then we construct a statistic with an approximate chi-square distribution and compute this statistic for a grid of values of (μ_1, p_1) and for various values of the parameters α and R. We find that only a restricted number of these parameters yield acceptably small values of the statistic.

The population of participants is divided into groups specified by age at enrollment, with the range of ages divided into equal segments with k_i being the number of participants and t_i the mean age in the ith group. We define a_i^* as the lung cancer (adenocarcinoma) death rate of the ith group, i.e., as the number of cancer deaths divided by the years in study summed over all participants in the group. To compute the moments of a_i^*, we make the simplifying assumption that the time in study, T, was identical for all participants. Then, based on Section 2.5, we have, approximately

$$a_i = \mathrm{E}a_i^* = \frac{\psi_{t_i}(T)}{T}, \quad (4.6)$$

$$b_i = \sigma_{a_i^*} = \left\{\frac{\psi_{t_i}(T)[1 - \psi_{t_i}(T)]}{k_i T^2}\right\}^{1/2}. \quad (4.7)$$

Clearly, for all k_i large, the statistic

$$A^2 = \sum \left(\frac{a_i^* - a_i}{b_i}\right)^2 \quad (4.8)$$

is approximately chi-square. Thus, we can compute A^2 for various values of the important parameters and eliminate those combinations which yield values greater than a preselected threshold.

5. Estimation Results

In this section we present the application of the estimation techniques developed in Section 4, to the study data. The results will be stated in subsections, corresponding to analogous subsections of Section 4.

5.1 *Confidence Regions for* μ_1 *and* p_1

In Section 4.1 methods were developed for confidence region estimation of the critical parameters (μ_1, p_1, ρ) for fixed values of the minimum age of onset (R), and coefficient of the age-of-onset distribution (α). In the present section, we discuss these confidence regions for values of R ranging from 0 to 40 years and α from 0 to 1.

The data on which this analysis is based are presented in Table 1. It should be noted that the population eligible for screening decreased from $N_1 = 10{,}040$ at the initial screen to $N_8 = 4914$ at the eighth screen. The decreases from first to fifth screen ($N_5 = 9462$) were due to deaths in the study population. The subsequent decreases arose from the fact that it required three calendar years to complete the enrollment while screening was phased out within one calendar year. Every participant was offered at least 5 years of screening and half were eligible for eight screens. In the analysis below, variances of X_1/Z_1 and X_2/Z_2 were calculated as if 9028 men, a weighted average, were all offered eight screenings. We note that not every eligible participant actually complied with each screen. Therefore, p_1, the early-stage detectability, must be interpreted as the probability that the patient with early-stage cancer complies with screening procedures *and* has his cancer detected.

During the screening period, a total of $Z_1 = 67$ early adenocarcinomas were detected, $X_1 = 36$ of them in the first four screens. Thus, with $K = 8$, $K_1 = 4$, we found $X_1/Z_1 = .54$. Of the 58 early cases detected after the initial screen, $Z_2 = 36$ had previous radiographs reexamined and $X_2 = 28$ had evidence of cancer visible in retrospect, so that $X_2/Z_2 = .78$. Furthermore, 44 advanced asymptomatic cases were found in screening and 29 symptomatic cases surfaced in the intervals between screens.

We allowed the minimum age of onset R and the slope of the trapezoidal onset density α to vary over a grid from 10 to 40 years and from 0 to 1, respectively. For each (R, α) pair considered, we plotted .95 confidence region estimates of μ_1 and p_1. In no case did the region contain values of μ_1 less than 2.3 years or of p_1 greater than .35. Moreover, the plots suggest that greater likelihood is associated with higher values of μ_1 and lower detectability p_1. This is not surprising in view of the high value of X_2/Z_2.

Table 1
Summary of adenocarcinoma detection

	Screening number (k)								
	1	2	3	4	5	6	7	8	Total
Population (N_k)	10,040	9879	9729	9603	9462	8924	7015	4914	
Number of early cases detected	9	10	7	10	8	9	11	3	67
Number of early cases detected and reexamined	—	8	4	9	7	4	3	1	36
Number of early cases visible in retrospect	—	7	3	6	6	3	2	1	28
Number of advanced asymptomatic cases detected	15	5	3	5	5	2	6	3	44
Number of symptomatic cases	3	2	5	7	3	4	5	0	29

We were then able to estimate the value of ρ, the fraction of study population who are susceptible to adenocarcinoma of lung. This fraction can be safely placed below .2.

5.2 *Estimation of μ_2 and c_1*

Following the discussion of Section 4.2, we present the results of estimation of μ_2, the mean time in the advanced asymptomatic stage, and c_1, the cure probability in early stage from data on survival of confirmed cancer cases.

The estimate of μ_2 is based on the survival of advanced asymptomatic cases. There were 44 such patients, of whom 35 died of lung cancer, with an estimated 5-year survival of .20. With the assumption of zero cure probability, the maximum likelihood estimate of μ_2 is 2.9 years with a standard deviation of .5 year.

The survival of cases of early adenocarcinoma was used to estimate the cure probability c_1 as a function of mean time μ_1. The sample consisted of 67 cases of early adenocarcinoma detected. Their survival has been tracked through the year 1984, although screening ended in 1982. The number of patients who died of adenocarcinoma during the entire period was equal to 17, with an estimated 5-year survival of .75. The maximum observation period for a single case was equal to 10 years. Figure 2 depicts the estimate of c_1 as a function of μ_1, demonstrating that $\hat{c}_1(\mu_1)$ is a decreasing function of μ_1, close to linear, with $\hat{c}_1(0) = .68$ and $\hat{c}_1(\mu_1) = 0$ for $\mu_1 \geq 9.2$.

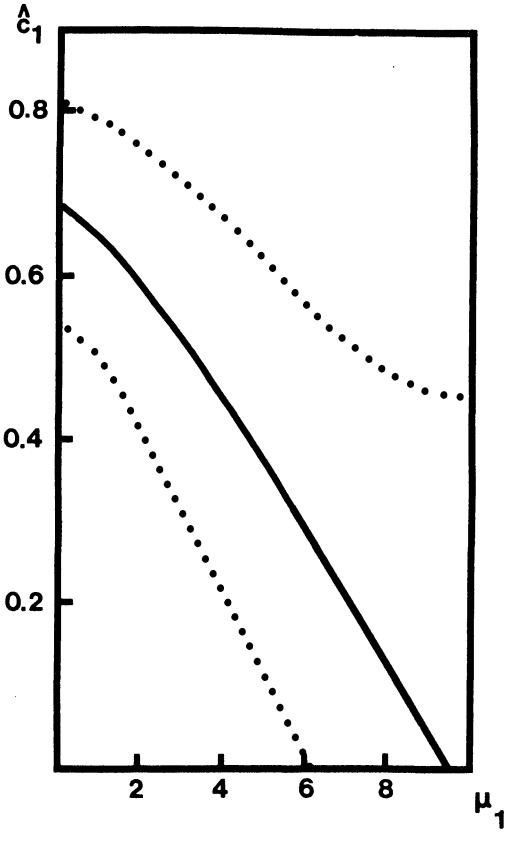

Figure 2. Confidence interval (.95) estimate of c_1 as a function of μ_1.

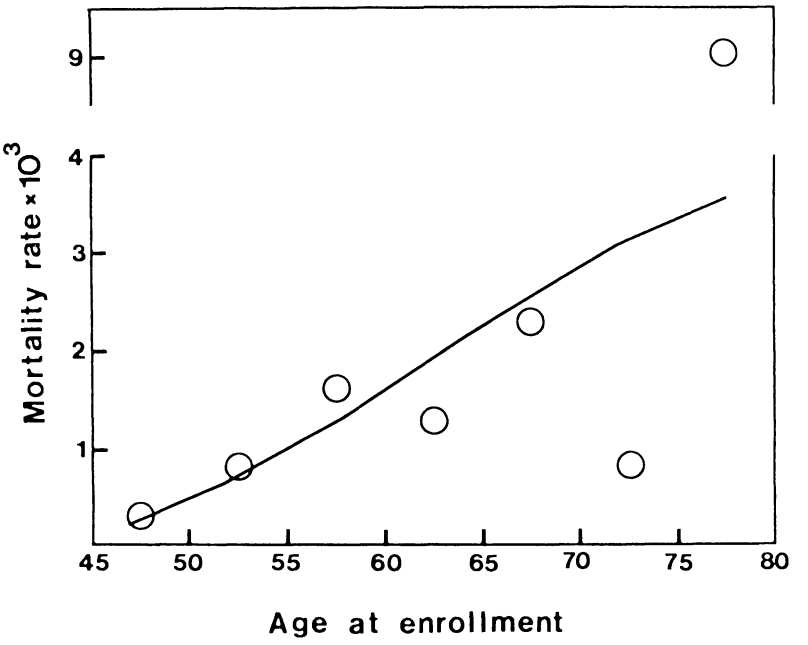

Figure 3. Mortality rates a_i^* plotted against age at enrollment for seven age groups. Circles represent data; the curve corresponds to the best fit $\alpha = 0$, $R = 40$, $\mu_1 = 5$, $p_1 = .12$, $\tilde{\rho} = .11$, $\hat{c}_1 = .37$.

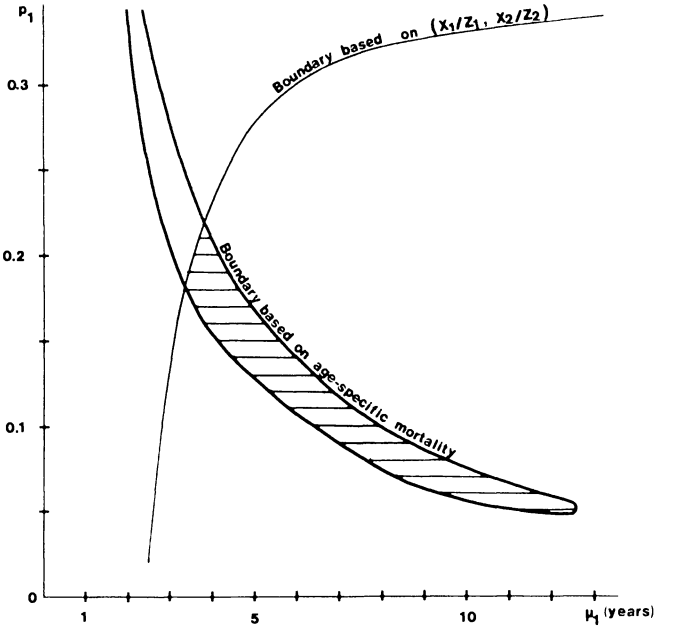

Figure 4. Intersection of confidence regions (.95) for (μ_1, p_1) based on $(X_1/Z_1, X_2/Z_2)$ and on age-specific mortality, indicated by hatching.

5.3 *Limits on μ_1 and p_1*

Using the method described in Section 4.3, we utilized age-specific population mortality data to limit further the confidence regions for (μ_1, p_1). The study population whose age at enrollment ranged from 45 to 80 was divided into seven groups with a 5-year age range in each group. The values of the adenocarcinoma death rates are plotted in Figure 3. It should be noted that only 190 men were in the 70–75 age group and 36 in the 75–80 group.

Based on the statistic defined in equation (4.8) and on the estimates of μ_2 and c_1 found in Section 5.2, we plotted .95 confidence regions for (μ_1, p_1) for values of R and α varying over the grid used in Section 5.1. It was found that the best fit to the data was associated with $R = 40$, $\alpha = 0$. In Figure 4 we exhibit the intersection of confidence regions for (μ_1, p_1) described in Section 5.1 with those found in this section, all based on $R = 40$, $\alpha = 0$.

6. Effect of Screening on Mortality

It should be recalled that the major reason for this modeling effort is an attempt to assess the effect of screening on lung cancer mortality, in view of the fact that the entire study population was offered annual Xray examination. Now that the structure of the model is completely defined and the region of feasible parameters has been determined, we turn our attention to predicting mortality. Consider a population in which periodic screening is initiated at a specified age and continued indefinitely. First we model the predicted mortality in our study population during the tenure of the study and compare the calculated results with data. Finally, we consider the impact of screening at various intervals on a population screened from age 45 to death.

The total number of deaths from adenocarcinoma occurring during the screening period, in cases detected by screening or through symptoms in the intervals between screens, was 61. As a check on the plausibility of our model, we calculated the expected number of adenocarcinoma deaths during the study period for a grid of (μ_1, p_1) points inside the hatched area of Figure 4. These results are displayed in Table 2. It may be observed that values of μ_1 ranging from 4 to 11 years with p_1 ranging from .18 down to .06 give expected numbers of cancer deaths very close to the observed value of 61.

As a major application of our model, we asked the following question: If members of a high-risk population enter a radiographic screening program at age 45 and are screened by chest Xrays at regular intervals up to age 80, what is the impact on the mortality from adenocarcinoma? In Table 3 we present the expected mortality in populations screened annually compared to those not screened at all, for parameter values that seem most plausible based on all the preceding analysis.

It should be noted that the most favorable estimate indicates an 18% reduction in

Table 2
Expected numbers of deaths in study period based on model

Parameters of model				Estimated parameters		Adenocarcinoma deaths
α	R	μ_1	p_1	$\tilde{\rho}$	$\hat{c}_1$	
0	40	4	.18	.09	.46	59
		5	.14	.09	.37	60
		7	.10	.10	.21	59
		9	.08	.11	.05	57
		11	.06	.12	0	59

Table 3
Expected adenocarcinoma mortality (per 10,000) *in population screened from age* 45 *to age* 80 *based on model*

Model parameters				Estimates				
α	R	μ_1	p_1	$\tilde{\rho}$	$\hat{c}_1$	No screen	Screening	% Difference
0	40	4	.18	.09	.46	426	351	18
		5	.14	.09	.37	425	370	13
		7	.10	.10	.21	411	383	7
		9	.08	.11	.05	394	388	2
		11	.06	.12	0	416	416	0

adenocarcinoma mortality over a 35-year period with annual screening. These figures relate to values of 4 years for μ_1, the mean time in early stage; .18 for p_1, the early-stage detectability; and .46 for c_1, the early-stage cure probability. As μ_1 increases and p_1 and c_1 decrease, these effects are diminished.

7. Conclusions

We have postulated a mathematical model of the natural history of lung cancer in the presence of a periodic screening program. It is assumed that a member of a susceptible subgroup of the population develops a cancer that progresses from early stage to advanced asymptomatic stage to symptoms and death. The age of onset and the times spent in successive stages are independent random variables. Each screening examination detects the cancer independently with a probability that depends only on the stage. Upon detection, the cancer is treated and the probability of cure depends on stage. Confidence regions of the important parameters of this model have been estimated (for lung adenocarcinoma) based on data collected in the National Lung Program carried out by Memorial Sloan-Kettering Cancer Center.

It must be recognized that the major objective of screening for lung cancer is to detect the disease in early stage with the hope of curing it. It is generally known and this study bears out the fact that advanced lung cancer, even though asymptomatic and detected, is almost invariably fatal. Therefore, the most important parameters in determining the effectiveness of a screening program are those related to early-stage cancer, the mean time μ_1, the detection probability p_1, and the cure probability c_1. In Sections 4 and 5 of this paper, confidence interval estimates of these parameters are derived from stage-specific incidence, survival, and mortality data.

Based on these multiple approaches to the estimation problem, we believe that the following feasibility regions hold for lung adenocarcinoma. The mean time in early stage μ_1 is at least 4 years and may be as great as 12 years. The susceptibility ρ is approximately .1. The probability of detection by each screening examination is less than .2. The estimate of early-stage cure probability is negatively correlated with the estimate of μ_1, close to .5 for μ_1 close to 4 years and decreasing to 0 for μ_1 greater than 9. Correspondingly, the estimated impact of annual screening for 35 years from age 45 on mortality from adenocarcinoma of lung in a population of male smokers ranges from 0 up to 18%. The analysis carried out in this paper results only in an upper limit on the possible benefit. This is related to the fact that the long survival of patients with early-stage cancer may be attributed indistinguishably to cure or to lead-time and length-biased sampling. It is clear that currently available screening techniques will not solve the problem of lung cancer mortality in smokers. However, it may be hoped that improved detection techniques, in increasing p_1, might make a very substantial difference.

In this paper, we have postulated a model for periodic screening for a chronic disease and developed novel methods of estimating the parameters using data collected entirely in one long-term clinical trial in which no relevant control information was available. While the resultant confidence regions are quite broad, a number of new insights have resulted.

In conclusion, we consider those assumptions of the model which may not be realistic and which we intend to investigate further. These include independence of times in successive stages, τ_0, τ_1, τ_2; exponential distribution of τ_1 and τ_2; independence of detection in successive screening examinations; and constancy of detectability p_1 throughout the duration of the early stage.

It is well known that fast-growing disease is distinguished from slow-growing disease, so that τ_1 and τ_2 should be positively correlated. While this would have very little effect on estimates of μ_1, it might impact the estimates of mortality benefit. We have conducted some preliminary studies of the robustness of results with respect to assumptions about the distributional form of τ_1 and τ_2 and have found little effect. The subject of changing detection probability in successive examinations as the cancer grows will be considered in a future paper.

Acknowledgements

The authors gratefully acknowledge the leadership and stimulation of Myron R. Melamed, M.D., throughout the development of this work, and the helpful comments of Emmanuel Yashchin and the referees. We thank Barbara Newman for her diligent assistance in preparing the manuscript.

This work was supported in part by NCI Contract #N01-CN-45007, ACS Grant #RD-207, and NIH Grant #CA-42830.

Résumé

La cinétique de progression du cancer du poumon dans une population régulièrement surveillée est étudiée par un modèle mathématique dont les paramètres sont estimés sur les données recueillies au Centre Anticancéreux du Memorial Sloan-Kettering à New York. Le développement d'un adénocarcinome est, par hypothèse de départ, un processus stochastique à deux états: stade précoce et stade avancé, caractérisés par leur durée moyenne, leur probabilité de détection, et leur probabilité de guérison. Les intervalles de confiance des paramètres ont été étudiés à l'aide de techniques originales. Nous avons eu la surprise de trouver une durée moyenne du stade précoce d'au moins 4 ans, avec une probabilité de détection de moins de 20% et une curabilité inférieure à 50%. Selon ces estimations, un dépistage radiographique annuel de l'age de 45 ans jusqu'à 80 ans ne pourrait entraîner qu'une diminution inférieure à 20% de la mortalité par adénocarcinome pulmonaire.

References

Albert, A., Gertman, P. M., and Louis, T. A. (1978). Screening for the early detection of cancer—I. The temporal natural history of a progressive disease state. *Mathematical Biosciences* **40,** 1–59.

Albert, A., Gertman, P. M., Louis, T. A., and Liu, S.-I. (1978). Screening for the early detection of cancer—II. The impact of the screening on the natural history of the disease. *Mathematical Biosciences* **40,** 61–109.

American Cancer Society (1980). Report on the cancer-related health checkup. *Ca* **30,** 194–240.

Bickel, P. J. and Doksum, K. A. (1977). *Mathematical Statistics.* San Francisco: Holden-Day.

Blumenson, L. E. (1976). When is screening effective in reducing the death rate? *Mathematical Biosciences* **30,** 273–303.

Blumenson, L. E. (1977). Detection of disease with periodic screening: Transient analysis and application to mammography examination. *Mathematical Biosciences* **33,** 73–106.

Chen, J. S. and Prorok, P. C. (1983). An age-dependent stochastic model of periodic screening: Basic properties. *Mathematical Biosciences* **65,** 69–91.

Chen, J. S., Prorok, P. C., and Graff, K. M. (1983). An age-dependent stochastic model of periodic screening: Length bias at a prevalence screen. *Mathematical Biosciences* **65,** 93–123.

Day, N. E. and Walter, S. D. (1984). Simplified models of screening for chronic disease: Estimation procedures from mass screening programmes. *Biometrics* **40,** 1–14.
Dubin, N. (1981). Predicting the benefit of screening for disease. *Journal of Applied Probability* **18,** 348–360.
Flehinger, B. J., Melamed, M. R., Zaman, M. B., Heelan, R. T., Kimmel, M., Perchick, W. B., and Martini, N. (1983). Screening for early detection of lung cancer in New York. A Report of a UICC International Workshop held at Venice, Italy. P. C. Prorok and A. B. Miller (eds). *UICC Technical Report Series* **78,** 123–135.
Flehinger, B. J., Melamed, M. R., Zaman, M. B., Heelan, R. T., Perchick, W. B., and Martini, N. (1984). Early lung cancer detection: Results of the initial (prevalence) radiologic and cytologic screening in the Memorial Sloan-Kettering study. *American Review Respiratory Disease* **130,** 555–560.
Heelan, R. T., Flehinger, B. J., Melamed, M. R., Zaman, M. B., Perchick, W. B., Caravelli, J. F., and Martini, N. (1984). Non-small-cell lung cancer: Results of the New York Screening Program. *Radiology* **151,** 289–293.
Louis, T. A., Albert, A., and Heghinian, S. (1978). Screening for the early detection of cancer—III. Estimation of disease natural history. *Mathematical Biosciences* **40,** 111–144.
Manton, K. G. and Stallard, E. (1982). A population-based model of respiratory cancer incidence, progression, diagnosis, treatment and mortality. *Computers and Biomedical Research* **15,** 342–360.
Meeker, W. Q., Jr. (1984). A comparison of accelerated life test plans for Weibull and lognormal distributions and Type I censoring. *Technometrics* **26,** 157–171.
Melamed, M. R. and Flehinger, B. J. (1984). Screening for lung cancer. *Chest* **86,** 2–3.
Melamed, M., Flehinger, B., Miller, D., Osborne, R., Zaman, M., McGinnis, C., and Martini, N. (1977). Preliminary report of the lung cancer detection program in New York. *Cancer* **39,** 369–382.
Melamed, M. R., Flehinger, B. J., Zaman, M. B., Heelan, R. T., Hallerman, E. T., and Martini, N. (1981). Detection of true pathologic Stage I lung cancer in a screening program and the effect on survival. *Cancer* **47,** March 1 Supplement, 1182–1187.
Melamed, M. R., Flehinger, B. J., Zaman, M. B., Heelan, R. T., Perchick, W. A., and Martini, N. (1984). Screening for early lung cancer. Results of the Memorial Sloan-Kettering Study in New York. *Chest* **86,** 44–53.
Prorok, P. C. (1976a). The theory of periodic screening I: Lead time and proportion detected. *Advances in Applied Probability* **8,** 127–143.
Prorok, P. C. (1976b). The theory of periodic screening II: Doubly bounded recurrence times and mean lead time and detection probability estimation. *Advances in Applied Probability* **8,** 460–476.
Shapiro, S., Goldberg, J. D., and Hutchison, G. B. (1974). Lead time in breast cancer detection and implications for periodicity of screening. *American Journal of Epidemiology* **100,** 357–366.
Walter, S. D. and Day, N. E. (1983). Estimation of the duration of a pre-clinical disease state using screening data. *American Journal of Epidemiology* **118,** 865–886.
Zelen, M. and Feinleib, M. (1969). On the theory of screening for chronic diseases. *Biometrika* **56,** 601–614.

Received March 1985; *revised February and October* 1986.

Appendix

It will be proved that the distribution of the random vector $(X_1/Z_1, X_2/Z_2)$ defined in Section 4.1 is asymptotically bivariate normal.

We require some preliminary arguments and definitions. We have

$$Z_i = \sum_{n=1}^{N} z_{in}; \qquad X_i = \sum_{n=1}^{N} x_{in}, \quad i = 1, 2,$$

where z_{in} and x_{in} are independent, identically distributed indicator functions of the following events for the nth participant, $n = 1, \ldots, N$:

$\mathscr{Y}_1 = \{$early cancer detected in screen 1, 2, . . . , or $K\}$,

$\mathscr{X}_1 = \{$early cancer detected in screen 1, 2, . . . , or $K_1\}$,

$\mathscr{G}_2$ = {early cancer detected in screen 2, 3, . . . , or K and prior Xray available},

$\mathscr{X}_2$ = {early cancer detected in screen 2, 3, . . . , or K and found in prior Xray}.

Then

$$\beta_1 = \Pr\{\mathscr{G}_1\} = \sum_{k=1}^{K} E_k,$$

$$\gamma_1 = \Pr\{\mathscr{X}_1\} = \sum_{k=1}^{K_1} E_k,$$

$$\beta_2 = \Pr\{\mathscr{G}_2\} = s \sum_{k=2}^{K} E_k,$$

$$\gamma_2 = \Pr\{\mathscr{X}_2\} = s \sum_{k=2}^{K} R_k,$$

$$\beta_{12} = \Pr\{\mathscr{G}_1 \cap \mathscr{G}_2\} = s \sum_{k=2}^{K} E_k,$$

$$\gamma_{12} = \Pr\{\mathscr{X}_1 \cap \mathscr{X}_2\} = s \sum_{k=2}^{K_1} R_k,$$

$$\delta_{12} = \Pr\{\mathscr{X}_1 \cap \mathscr{G}_2\} = s \sum_{k=2}^{K_1} E_k,$$

$$\delta_{21} = \Pr\{\mathscr{X}_2 \cap \mathscr{G}_1\} = s \sum_{k=2}^{K} R_k.$$

Define the random vector

$$\mathbf{Y}_N = (X_1, X_2, Z_1, Z_2)^{\mathrm{T}}.$$

Its expectation is

$$\mathrm{E}(\mathbf{Y}_N) = N(\gamma_1, \gamma_2, \beta_1, \beta_2)^{\mathrm{T}}$$

and its covariance matrix is

$$\mathbf{NA} = \mathbf{N} \begin{pmatrix} \gamma_1(1-\gamma_1) & \gamma_{12} - \gamma_1\gamma_2 & \gamma_1(1-\beta_1) & \delta_{12} - \gamma_1\beta_2 \\ * & \gamma_2(1-\gamma_2) & \delta_{21} - \gamma_2\beta_1 & \gamma_2(1-\beta_2) \\ * & * & \beta_1(1-\beta_1) & \beta_{12} - \beta_1\beta_2 \\ * & * & * & \beta_2(1-\beta_2) \end{pmatrix}, \tag{A.1}$$

the asterisks denoting symmetry.

Theorem. The sequence of random vectors $\mathbf{V}_N$,

$$\mathbf{V}_N = \begin{pmatrix} \dfrac{X_1}{Z_1} - \dfrac{\gamma_1}{\beta_1} \\ \dfrac{X_2}{Z_2} - \dfrac{\gamma_2}{\beta_2} \end{pmatrix} \sqrt{N},$$

tends in distribution, as $N \to \infty$, to the bivariate normal with zero expectation and covariance matrix

$$\mathbf{Q} = \begin{pmatrix} \dfrac{\gamma_1(\beta_1 - \gamma_1)}{\beta_1^3} & \dfrac{\gamma_{12}}{\beta_1\beta_2} - \dfrac{\gamma_1\delta_{21}}{\beta_1^2\beta_2} - \dfrac{\gamma_2\delta_{12}}{\beta_1\beta_2^2} + \dfrac{\gamma_1\gamma_2\beta_{12}}{\beta_1^2\beta_2^2} \\ * & \dfrac{\gamma_2(\beta_2 - \gamma_2)}{\beta_2^3} \end{pmatrix}.$$

Proof. By the central limit theorem,

$$\frac{\mathbf{Y}_N - \mathrm{E}(\mathbf{Y}_N)}{\sqrt{N}} \xrightarrow{d} \mathrm{N}(\mathbf{0}, \mathbf{A}).$$

Define random matrices $\mathbf{B}_N$,

$$\mathbf{B}_N = \begin{pmatrix} \frac{N}{Z_1} & 0 & -\frac{\gamma_1 N}{\beta_1 Z_1} & 0 \\ 0 & \frac{N}{Z_2} & 0 & -\frac{\gamma_2 N}{\beta_2 Z_2} \end{pmatrix}.$$

By the law of large numbers, $\mathbf{B}_N$ tend with probability 1 to $\mathbf{B}$,

$$\mathbf{B} = \begin{pmatrix} \frac{1}{\beta_1} & 0 & -\frac{\gamma_1}{\beta_1^2} & 0 \\ 0 & \frac{1}{\beta_2} & 0 & -\frac{\gamma_2}{\beta_2^2} \end{pmatrix}.$$

By a matrix generalization of the Slutski theorem (Bickel and Doksum, 1977, theorem A.14.9), we have

$$\mathbf{V}_N = \mathbf{B}_N \frac{\mathbf{Y}_N - \mathrm{E}(\mathbf{Y}_N)}{\sqrt{N}}$$

tends in distribution to the bivariate normal with covariance matrix $\mathbf{BAB}^{\mathrm{T}} = \mathbf{Q}$. The theorem is proved.

Remark. Let us note that the scalar random variable $\mathbf{V}_N^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{V}_N$ has, for N large, a distribution close to chi-square with 2 degrees of freedom. This fact is used to construct confidence regions in Section 4.1.