Published in final edited form as:

Sci China Math. 2012 August 1; 55(8): 1565–182. doi:10.1007/s11425-012-4475-y.

Dynamic Optimal Strategy for Monitoring Disease Recurrence

Hong Li¹ and Constantine Gatsonis²

Hong Li: hong_li2@rush.edu; Constantine Gatsonis: gatsonis@stat.brown.edu

¹Department of Preventive Medicine, Rush University Medical Center, Chicago, IL 60612, U.S.A

²Department of Biostatistics, Brown University, Providence, RI 02912, U.S.A

Abstract

Surveillance to detect cancer recurrence is an important part of care for cancer survivors. In this paper we discuss the design of optimal strategies for early detection of disease recurrence based on each patient's distinct biomarker trajectory and periodically updated risk estimated in the setting of a prospective cohort study. We adopt a latent class joint model which considers a longitudinal biomarker process and an event process jointly, to address heterogeneity of patients and disease, to discover distinct biomarker trajectory patterns, to classify patients into different risk groups, and to predict the risk of disease recurrence. The model is used to develop a monitoring strategy that dynamically modifies the monitoring intervals according to patients' current risk derived from periodically updated biomarker measurements and other indicators of disease spread. The optimal biomarker assessment time is derived using a utility function. We develop an algorithm to apply the proposed strategy to monitoring of new patients after initial treatment. We illustrate the models and the derivation of the optimal strategy using simulated data from monitoring prostate cancer recurrence over a 5-year period.

Keywords

Keywords Biomarker trajectory; Cancer recurrence surveillance; Latent class model; Optimal strategy; Time-dependent hazard

1 Introduction

Monitoring to detect disease recurrence is an important and costly part of health care for patients with chronic conditions. In particular, more than 10 million cancer survivors are candidates for such surveillance in the U.S. and are monitored following a variety of strategies, which include imaging tests and other biomarkers ([1], [2], [3], [4], [5], [6]). The frequency and modality of monitoring is typically decided on the basis of a combination of clinical experience and pragmatic considerations about practicality and cost. However, there is limited systematic investigation on how to develop, optimize, and evaluate surveillance strategies, leading to a commonly expressed concern that current practices may result in suboptimal use of resources and put a heavy burden on the health care system ([4], [7], [8],

[9], [10]). Thus the development of optimal strategies for patient monitoring is an important current need.

In this paper, we address the question of how to decide when patients should come back for follow up after initial treatment using a biomarker as a tool to monitor disease recurrence according to the characteristics of patient and disease. Indeed, the phenomenal recent growth of all types of biomarkers, including molecular, genomic, genetic, and imaging based markers, has brought forth a large number of potential new modalities for surveillance for cancer recurrence. For example, Prostate Specific Antigen (PSA) values are used to monitor for recurrence of prostate cancer and changes in Standard Uptake Value (SUV) measurements obtained by Positron Emission Tomography (PET) are known to precede clinically detectable and symptomatic recurrence in several cancers. The availability of longitudinal observations of biomarker values has led to increasing recognition that the *longitudinal trajectory* of the biomarker may provide important additional information beyond the biomarker's current value ([11], [12], [13], [14]). As noted in recent studies, the trajectory describes biomarker behavior and risk factors over time and is influenced by characteristics of patients and disease, such as age, gender, tumor size, and tumor stage.

In our work, we use the biomarker trajectory to classify patients and construct optimal monitoring strategies for cancer recurrence. The statistical approach incorporates multiple components, including utility and risk analysis, modeling of the natural history of the disease, the biomarker trajectory, and the hazard rate for recurrence. To define optimality we adapt an expected utility criterion originally developed for cancer screening ([59]). We assume a three-state Markov chain model with time-dependent transition probabilities for the natural history of the disease, which is an extension of the model with constant transition probabilities used in cancer screening ([15]). To link the biomarker trajectory with the risk of recurrence we adopt a latent class joint model (LCJM) which makes it possible to assess and incorporate patient heterogeneity and to discover distinct classes of patient trajectories. These classes are used to assign patients into different risk groups and to predict the risk of disease recurrence ([16], [17]). LCJM is also implemented in recent work by Proust-Lima et al. (2009) ([14]), where it is used for prediction of recurrence in prostate cancer. In our paper, the focus is on the development of optimal monitoring strategies, which can be adapted dynamically.

Monitoring strategies in current clinical practice use fixed time intervals for all patients without incorporating information on the individual patient's disease characteristics. By comparison, a novel aspect of our approach is that it explicitly incorporates the heterogeneity of patients and develops a strategy which is based on prediction of disease recurrence using patients' biomarker trajectory and periodically updated risk. The optimal interval between monitoring examinations depends on the speed with which important clinical changes develop. Hence, the developed optimal strategy is class-based, i.e. patients in different latent classes may follow different monitoring schedules and a monitoring schedule for patients in the same class may vary as time goes on. Most importantly we develop an algorithm in which we could easily adapt the class-based strategy for new patients whose biomarker trajectory data do not yet exist.

In Section 2, we describe the methods including the natural history model, the LCJM which will be used to analyze the longitudinal surveillance, the derivation of optimal monitoring strategies, a proposed algorithm to apply the developed monitoring strategy to new patients whose biomarker trajectory data do not yet exist, and the criteria for evaluating monitoring strategies. In Section 3, we describe the estimation method used for model fitting. In Section 4, we apply the proposed method to simulated data from monitoring for recurrence of prostate cancer using PSA. The proposed strategy is shown to have better performance than a strategy used in routine clinical practice. We then conclude with a discussion in Section 5.

2 The Method

2.1 The Model

We consider the setting of a prospective cohort study of monitoring disease recurrence in which the initially enrolled patients have undergone therapy and achieved remission. For every patient at each monitoring time, the following data are collected: biomarker measurements Y , covariates V associated with the class membership of biomarker trajectories, fixed effects X , random effects $\tilde{X}$, group effects U used to model biomarkers, and covariates Z associated with the hazard for disease recurrence.

The natural history model assumes three disease states, remission (none or undetectable disease) denoted by S_0 , recurrence (recurrence without symptom) denoted by S_1 , and clinical recurrence (symptomatic recurrence) denoted by S_2 . Disease status may change from remission to recurrence, remission to clinical recurrence, and recurrence to clinical recurrence. The transition process is modeled as a Markov chain in continuous time.

We also assume that, a) the result of the monitoring exam is positive or negative, b) the sensitivity is a constant, and c) clinical recurrence is observed without error.

The pattern of the biomarker trajectory is viewed as defining a latent class in the LCJM of biomarker values and risk of recurrence. In other words, each trajectory pattern represents a subpopulation of patients which has its own biomarker response and level of risk to trigger disease recurrence. We make the usual assumption of conditional independence (CI) between the time to recurrence of disease and the longitudinal biomarker response given the latent class ([17]). We also discuss testing for this assumption and explore the robustness of the results in Section 5.

The LCJM consists of three components: class membership, longitudinal biomarker trajectories, and hazard for time to disease recurrence ([17]). We assume the number of latent classes, L , is unknown and indexed by $l = 1, 2, \dots, L$, the number of patients, n_r , indexed by $i = 1, 2, \dots, n_r$ at time t_r . We define $c_i = (c_{i1}, \dots, c_{iL})^\top$, where $c_{il} = 1$ if patient i is a member of class l and 0 otherwise. The probability of patient i belonging to latent class l , denoted by π_{il} , is modeled by multinomial logistic regression with covariate vector $V_i = (V_{i1}, V_{i2}, \dots, V_{id})^\top$ which represents characteristics of the patient known at baseline,

$$\pi_{il}=P(c_{il}=1)=\frac{\exp(\omega_l V_i^\top)}{\sum_{j=1}^L \exp(\omega_j V_i^\top)} \quad (2.1)$$

where ω_l is the coefficient vector for class l with $\omega_1=0$ to make model identifiable.

Each latent class has its own linear model for the longitudinal biomarker measurements Y . For patient i in latent class l at time t_r we assume,

$$Y_i=\gamma_{0l}+U_i(\Gamma c_i)+X_i\gamma_2+\tilde{X}_i\gamma_{3i}+\varepsilon_i \quad (2.2)$$

where $Y_i=(Y_{it_1}, \dots, Y_{it_r})^\top$ is the t_r -dimensional biomarker vector; $U_i=(U_{it_1}^\top, \dots, U_{it_r}^\top)^\top$ represents the $t_r \times q_1$ class-specific; covariate; the class-specific parameters are in the $q_1 \times L$ matrix Γ , where $\Gamma=(\gamma_{11}, \dots, \gamma_{1L})$ with γ_{1l} a q_1 -dimensional column vector containing the parameters specific to class l ; $X_i=(X_{it_1}^\top, \dots, X_{it_r}^\top)^\top$ represents the $t_r \times q_2$ fixed effect covariates; the q_2 -dimensional vector γ_2 is the fixed effect coefficient;

$\tilde{X}_i=(\tilde{X}_{it_1}^\top, \dots, \tilde{X}_{it_r}^\top)^\top$ represents the $t_r \times q_3$ random effect covariates; the q_3 -dimensional vector γ_{3i} is the random effect coefficient; the t_r -dimensional vector γ_{0l} is the class-specific intercept; and ε_i is multnormally distributed with mean 0 and covariance matrix $\sigma^2 I_{tr}$ and is uncorrelated with γ_{3i} .

Each latent class also has its own model for the event process. For patient i in latent class l at t_r , the hazard function is

$$\lambda_{ilh}(t_r, Z_i)=\delta_{l0h} \exp\left(\delta_{lh}^\top Z_i\right) \quad (2.3)$$

where $h=0, 1$ represents S_0 and S_1 disease states and $l=1, \dots, L$; $\lambda_{ilh}(t_r, Z_i)$ is a time-dependent hazard function and represents the hazard of patient i in latent class l with disease state h , which depends on each patient only through covariate Z_i ; δ_{l00} and δ_{l01} are the class-specific baseline hazards of S_0 and S_1 ; Z_i represents a set of covariates which can be time-dependent, such as the PSA doubling time or time-independent such as patient age or cancer stage before therapy; and δ_{l0} and δ_{l1} are the class-specific coefficient vectors for covariates Z_i for S_0 and S_1 , respectively.

The likelihood function for monitoring detection up to the current measurement time t_{k-1} is

$$L(\delta, \beta_{10}; Z, e, \Delta)=\prod_{r=1}^{k-1} \prod_{l=1}^L \prod_{i=1}^{n_{rl}} (p_{irl}^{e_{irl}} (1-p_{irl})^{1-e_{irl}})^{\Delta_{irl}} \quad (2.4)$$

where δ denotes combined parameters including $\{\delta_{l0h}\}$ and $\{\delta_{lh}^\top\}$, e denotes the vector of the indicator of monitoring detection status, $\{e_{irl}\}$, defined below, and Δ denotes the vector of latent class membership indicator, $\{\Delta_{irl}\}$, defined below. The likelihood is derived under the following assumptions. Whether or not patient i in class l has disease recurrence detected

at time t_r follows a Bernoulli distribution with probability denoted by p_{irl} . For patients in latent class l at time t_r , n_{rl} is the total number of patients attending the monitoring examination, $e_{irl}=1$ if disease has been detected by monitoring in patient i , and $e_{irl}=0$ otherwise, and $i_{rl}=1$ if patient i belongs to latent class l at time t_r , and $i_{rl}=0$ otherwise. The probabilities of monitoring detection at time t_1 and t_r are $p_{i1l} = \theta_{i101}(t_1)(1 - \beta_{10})$ and $p_{irl} = (1 - \beta_{10})(\theta_{i101}(t_r)(1 - P_{il}(T_{r-1}^1)) + \theta_{i111}(t_r)\beta_{10}P_{il}(T_{r-1}^1))$ (see Supplementary Materials 1 and 2 for detailed derivations), where $r = 2, 3, \dots$, $\beta_{10}=1$ -sensitivity, T_{r-1}^1 =the true disease status is S_1 at t_{r-1} , and $\theta_{ijb}(t_r)$ is the transition probability from state j to b for patient i in latent class l from t_{r-1} to t_r . Note that $\theta_{i111}(t_r) = \exp^{-\int_{t_{r-1}}^{t_r} \lambda_{i11}(t)dt}$, $\theta_{i112}(t_r) = 1 - \exp^{-\int_{t_{r-1}}^{t_r} \lambda_{i11}(t)dt}$, $\theta_{i100}(t_r) = \exp^{-\int_{t_{r-1}}^{t_r} \lambda_{i10}(t)dt}$, $\theta_{i101}(t_r) = \int_{t_{r-1}}^{t_r} \exp^{-\int_{t_{r-1}}^v \lambda_{i10}(t)dt} * \exp^{-\int_v^{t_r} \lambda_{i11}(t)dt} * \lambda_{i10}(v)dv$ and $\theta_{i102}(t_r) = 1 - \theta_{i100}(t_r) - \theta_{i101}(t_r)$ (see Supplementary Materials 4 for detailed derivations). These transition probabilities are functions of the hazard functions $\lambda_{i11}(t)$ and $\lambda_{i10}(t)$ defined by (2.3). Since the likelihood function (2.4) is implicitly a function of the hazard functions $\lambda_{i11}(t)$ and $\lambda_{i10}(t)$, it follows that, by maximizing the likelihood function (2.4), we can obtain estimates for the parameters of the hazard model (2.3).

2.2 Optimal Strategy

2.2.1 Expected utility maximization—The goal of monitoring for cancer recurrence is to detect disease in the *recurrence without symptom* stage (S_1) and thus prevent patients from reaching symptomatic recurrence. The expectation is that by successfully detecting disease in an earlier and more treatable stage would make it possible to begin treatment as early as possible which may lead to slow down or halt disease progression. Hence, early detection by monitoring may prolong the time to progression of clinical recurrence and may also have a significant impact on disease mortality. To select a monitoring strategy we adapt the optimality criterion of Parmigiani et al. (2002) ([59]) which was originally intended for use in selecting screening strategies that lead to reduced the probability of first occurrence of symptomatic disease. The utility function, $u(t, W)$, which depends on the monitoring time and clinical outcomes, is defined as a negative value a if clinical recurrence and 0 otherwise, where W are the random variables associated with the natural history of the disease and the outcome of the exam. The optimal objective function, which is the expected value of the utility function with respect to all unknowns, is

$$\tilde{U}(t) = aP(\text{clinical recurrence at } t) \quad (2.5)$$

The form of the expected utility function makes it straightforward to implement across different disease settings. More elaborate utility measures can also be used, such as the expected delay in detecting tumor recurrence ([18], [19]), and the total expected loss ([20]). However, the first two of these utility functions require detailed information on the biology of the disease, while the third requires detailed knowledge about the financial and other sequelae of monitoring. The computation of the optimal monitoring time requires

maximization of the expected utility function (2.5), or equivalently, minimization of the probability of clinical recurrence at time t (since a is a negative number).

The selection of an optimal monitoring time can be made for individual patients or for groups of patients. For an individual patient the monitoring time can be computed by fitting models (2.1)-(2.4) using the data up to the current measurement time, substituting parameter estimates and the individual covariate measurements in function (2.5), and maximizing (2.5) with respect to time t . We believe that a more practical approach is to select monitoring times that would apply to patients in the latent classes, which as noted earlier, are likely to consist of patients similar with respect to observed covariates, biomarker trajectory patterns, and risk of recurrence. Patients in different trajectory classes have different trajectory patterns and risk, therefore they are more likely to follow different monitoring schedules. Therefore, it seems reasonable to apply the optimal criterion to every latent class and calculate the class-based optimal time instead of the monitoring time for each individual. As we discuss below, a class-based strategy can be readily adapted to use in the case of new patients, for whom trajectory data do not yet exist.

The expression for the class-based probability of clinical recurrence at the next unknown examination time t_k is

$$P_l(\text{clinical recurrence at } t_k) = \theta_{l02}(t_k)P_l(T_{k-1}^0) + \theta_{l12}(t_k)P_l(T_{k-1}^1)\beta_{10} \quad (2.6)$$

where $\theta_{jlb}(t_k)$ is the class-based transition probability for disease transferring from state j to b for patients in latent class l from t_{k-1} to t_k (see the detailed derivation in Supplementary Materials 4). The detailed derivation of (2.6) is given in Supplementary Materials 3. In order to calculate $\theta_{jlb}(t_k)$ in (2.6), we use the average values of the covariates for patients in each latent class to calculate the class-based transition probabilities since patients in the same latent class are similar with respect to observed covariates. By minimizing function (2.6), we obtain the class-based optimal monitoring time.

To summarize the computation of the group optimal strategy at time t_k , we use available data up to t_{k-1} to classify patients into different latent classes and get parameter estimates by fitting models (2.1)-(2.4) and substituting the parameter estimates into function (2.6). The next optimal monitoring time t_k is then calculated as the time that minimizes (2.6). This process is repeated at each of the subsequent time points until the end of the study or to a time point when every patient has had cancer recurrence. If no recurrence is recorded in the database for members of a particular latent class, it is not possible to compute the optimal strategy for this class. A reasonable choice for the monitoring time interval for patients in that class would be to set it as the maximum among all monitoring time intervals for patients in all other classes with recurrent events. It is also possible to set up an upper bound, denoted by $\tilde{t}_U$, and lower bound, denoted by $\tilde{t}_L$, for the monitoring time interval $\tilde{t}$, which represents the difference between the next monitoring time and the current monitoring time. For example, we can set up $\tilde{t}_U$ at t_k as the optimal time interval calculated at t_{k-1} , since the risk of cancer recurrence at later times is equal to or larger than the risk at earlier times.

Chapter 18 in Taha (1997) ([21]) provides the algorithm to solve the optimization with inequality constraints.

2.2.2 Monitoring Schedule for New Patients—In previous sections, we discussed the derivation of a class-based monitoring strategy that can be developed if a longitudinal set of observations on the marker and the true recurrence status are available. The calculation of the monitoring schedule can be started only from the time of the first asymptomatic recurrent event. However, the optimal strategy needs to address the entire span of time a patient is a candidate for monitoring. In particular, in order to monitor new patients, we first need to decide the initial monitoring interval before recurrences are observed. If t_r is the time of the first asymptomatic recurrence for patients in class l and the optimal strategy computed on the basis of previous data calls for monitoring $\tilde{t}_r$ months later, then it is reasonable to monitor patients in this class at least every $\tilde{t}_r$ months before t_r because the risk of cancer recurrence before t_r is lower than or equal to the risk after t_r .

To implement the strategy it is also necessary to calculate the probability of class membership for new patients in the absence of their own longitudinal data. If t_r is the time of the first asymptomatic recurrent event confirmed by the reference standard, then for the time intervals up to t_r the probability of class membership for every new patient can be computed from model (2.1) using the coefficient estimate ω_l of each latent class at t_r and baseline covariates from this new patient. After t_r , the probability of class membership for every new patient can be computed from (2.1) using the coefficient estimate ω_l of each latent class at each monitoring time t_{r+1} , t_{r+2} , ... and baseline covariates from this new patient.

A strategy for new patients can proceed by first assigning every new patient to the class with the highest probability of class memberships and then monitoring each patient according to the schedule computed for patients in that class. Class membership and corresponding monitoring strategy can be updated and revised on the basis of information obtained at subsequent monitoring times.

2.3 Evaluation of monitoring strategies

A number of metrics for evaluating the performance of monitoring strategies can be adapted from proposals made in the screening literature ([22], [23]). On the basis of these metrics we propose the following five evaluation criteria to evaluate the efficiency of the proposed monitoring strategy.

- i. *The frequency of monitoring tests* defined as the total number of monitoring tests needed for each patient.
- ii. *The number of months earlier of monitoring detection* defined as the total number of months that the proposed strategy is able to detect disease recurrence earlier than a comparator strategy for each patient. The latter may be a strategy used in current clinical practice.
- iii. *Monitoring detection rate and error rate*: The cumulative monitoring detected disease recurrence rate f_1 is defined as the ratio of the total number of monitoring detected disease recurrence and the total number of disease recurrence, and the

cumulative error rate f_2 is defined as the ratio of the total number of symptomatic recurrence before scheduled examination and the total number of disease recurrence. A small value of f_2 or a large value of f_1 imply that the strategy could detect more asymptomatic recurrent events.

- iv. *Total cost* defined as the cost for an initial monitoring examination plus the cost of a confirmatory test if needed. The costs associated with time and travel to the test are not included in this calculation. The total cost up to time t_{k-1} , denoted by *cost*,

is $cost = \sum_{r=1}^{k-1} n_r cost_1 + \eta_r cost_2$ where $cost_1$ is the cost of the initial monitoring examination per patient, $cost_2$ is the cost of the confirmatory test per patient, n_r is the total number of patients attending the monitoring examination at time t_r , and η_r is the total number of patients whose true disease status has been verified at time t_r .

- v. *The percentage of the monitoring detected length (PMDL)* defined as

$$PMDL = \frac{\text{monitoring detected time} - \text{time at change point}}{\text{time at clinical recurrence} - \text{time at change point}} * 100$$

The smaller the PMDL is, the earlier stage that disease recurrence can be detected. The monitoring detected time is observed. We define the monitoring detected length as the difference between the monitoring detected time and the time of recurrence. Since we are unable to observe the time of recurrence, we use the time at biomarker change point, which is a time point cancer initiates and biomarker value starts to increase quickly, to approximate the time of recurrence. As an example, consider the monitoring of prostate cancer recurrence with PSA, which forms the setting for our simulated data in Section 4 below. The clinical disease-free interval, which is the time difference between the time at clinical recurrence and the time at recurrence, and PSADT, which is the time it takes for PSA to double, are linearly correlated and the time to symptomatic prostate cancer progression is estimated to be 3 times his PSADT ([24], [25]). PSADT is calculated by the two-point method, i.e. $PSADT = \log(2) * (t_k - t_1) / (\log(y_k) - \log(y_1))$ where y_1 is the biomarker value at the measurement time t_1 , and y_k is the biomarker value at the measurement time t_k , t_1 is any time after nadir and $PSA > 0.4$ ng/ml in an increasing PSA trend, t_k is any time before sending patients to the treatment, t_1 and t_k are any successive time at least 3 months apart and $y_k - y_{k-1}$ is at least larger than 15% of y_{k-1} . Based on the above definitions, it is easy to calculate the PSADT and clinical disease-free interval. We note that the definition of PMDL requires information that may not be available for all cancer types or other diseases.

3 Estimation

The full set of parameters in the joint model of (2.1)-(2.4) can be estimated directly through the joint likelihood function, using the CI assumption. However, the computational procedure can be complicated since we have a large dimensional parameter space and the joint likelihood function involves four components: longitudinal responses, time to event process, Markov chain model for disease transition, and latent class model. In addition, the fitting of the full model also needs to accommodate model choices (such as choosing the

number of latent classes which requires several such models to be fit), the missingness in latent class membership, and the repeated computation to update the latent class membership and monitoring strategy at many time points as we describe below. As a more practical alternative we pursued an ad-hoc approach to estimation which uses the special structure of the model. In particular, a) the latent classes correspond to biomarker trajectories, which capture all heterogeneities of the patients and disease, and b) given the latent class, the longitudinal model and survival model are assumed to be independent. We followed a two-stage approach to estimation, in which response variables and other covariates are used to categorize individuals by biomarker latent trajectory patterns (latent classes) in the first stage and the likelihood model is fitted in the second stage using the class membership assignments. In this two-stage approach, the first step addresses the question how to distinguish groups of patients with distinct patterns of longitudinal biomarker measurements and patient and disease characteristics, and the second step addresses the question how monitoring schedules would differ across classes of patients. Two-stage estimation has been used in the analysis of correlated survival data (for example, [26], [27], [28]), and has also been used in the latent class joint modeling literature (for example, [29], [30], [31], [32], [33]). As pointed out by Putter et al. (2008) ([31]) for a given number of classes, the two-stage procedure guarantees consistent estimates in both stages. We further discuss the efficiency of parameter estimation and compare the two-stage method and the full likelihood method via simulation in Section 5.

In the first stage, we fit the latent class model (2.1) and (2.2) to estimate parameters in these two models, discover the number of latent classes, and assign class membership for each patient. The fitting can be done using the R flexmix package ([34]), which implements a general framework for finite mixture models and latent class regression models. We use the Bayesian Information Criterion (BIC) to select the number of latent classes. Individual patients are assigned to the latent class with the highest posterior probability among all posterior probabilities for all latent classes. In the second stage, we estimate the parameters in the hazard function using the likelihood function (2.4) given the latent class assignments obtained from the first stage. We need to repeat this two-stage estimation procedure since in the setting of a prospective cohort study, parameter estimates and the probability of class membership need to be updated after collecting additional measurements at each future monitoring time until the end of the study or such time as all patients have disease recurrence.

A proper accounting of uncertainty in the final estimates needs to incorporate the uncertainty of latent class membership from the first stage (chapter 6 in [35], [36]). We use multiple imputation (MI) to account for this uncertainty, as has been done by other authors in the literature of latent class analysis (for example, [37], [36], [38]). In particular, we adapted the approach of Harel et al. (2007) ([37]), in which latent classes are viewed as variables that are missing completely at random (MCAR). We assume a joint model for the complete data $G = (G_{obs}, G_{mis})$ and the missingness indicator R , where G_{obs} are the observed data, G_{mis} are the missing latent class memberships, and R is a set of missingness indicators. We can ignore the missingness indicator R and impute the latent class membership based on the posterior probabilities of class memberships given the observed data and parameter estimates since

the latent class memberships are MCAR. We impute m independent sets of the latent class memberships, $G_{mis}^{(1)}, \dots, G_{mis}^{(m)}$ and then fit the latent class joint model on the m sets of complete data separately. Finally, we use Rubin's rules ([39]) to combine the m sets of point estimates and standard errors to obtain a single set of estimates for the model parameters and their standard errors. The latter are estimates which take the uncertainty about the unknown latent class membership into account. As pointed out by Little et al. (2002) ([39]) MI is a general and flexible method to handle the identifiability issues, allows users to derive excellent inferences for a broad range of estimands with complete-data method, and the resulting complete-data analyses can be easily combined to create an inference that validly reflects sampling variability because of the missing values. Note that it may be possible to avoid MI at some monitoring times under special circumstances, such as if the biomarker trajectories can be well separated or the overlap of components happens to be in a group of patients who did not have yet any disease recurrent events. We discuss how to assess the latent class model performance in Section 4.1.

In summary, the two-stage estimation approach offers two main advantages. First it simplifies the computation. The estimation of parameters from the joint likelihood function can be done via the EM-algorithm at each monitoring time but the procedure would be slow due to the following reasons: 1) large dimensional parameter space, 2) accommodating model choices, 3) numerical analysis needed due to no closed form solutions in the M step for majority of the parameters, and 4) running MI m times in order to take the estimation uncertainty into account. In addition, we need to repeatedly update the class membership and compute the monitoring schedule at many monitoring time points. The two-stage approach requires the EM-algorithm only in the first step to identify the number of latent classes and avoids fitting the joint model repeatedly, and also it requires the MI only in the second stage which will greatly simplify the estimation and save a lot of computation time. Second it conforms closely to the logic of the model and facilitates the interpretation of the results. Our interest is to distinguish biomarker trajectory patterns first and then to link each biomarker trajectory class directly to the risk of recurrence for patients in that latent class when we develop the monitoring strategies. We assess estimation bias for this two-stage estimation method using simulation in Sections 4 and 5.

4 Simulation

In order to evaluate and compare alternative monitoring strategies in terms on their impact on patient outcomes such as mortality, randomized trials could be conducted, in which one study arm would be monitored with a commonly used strategy in current clinical practice and the other arm would be monitored with the proposed strategy ([40]). However, in lieu of randomized studies and also in preparation for such studies it would be helpful to demonstrate the validity and utility of the approach proposed in this paper and to evaluate the efficiency of the proposed strategy. In order to assess the properties of the proposed strategies we conducted a simulation study.

4.1 Simulated dataset and model fitting

We simulated data from monitoring 150 prostate cancer patients for recurrence using PSA, over a 5-year period. In the simulated data set, 80 patients had monitoring detected recurrence and there was no symptomatic recurrence at the end of 5 years. We selected a period of 5 years because the risk of cancer recurrence is smaller after 5 years ([41]) and a sample size of 150, well within the range of used in published studies evaluating recurrence monitoring strategies for breast, ovarian, prostate, and bladder cancer ([42], [43], [44], [45]). Since the PSA level typically decreases below detectable levels for a certain time period after treatment, the starting point for the simulation was chosen to be the time point at which the PSA level was at nadir, which would be different for different patients, but detectable after treatment. The simulated data consisted of 150 patients with recorded variables: Gleason score (measure of tumor aggressiveness), tumor stage, pre-treatment PSA, age, monitoring time which would be different for different patients, log of biomarker value at each time, and true cancer recurrent status indicator. The detailed simulation model is given in the Appendix. Using the simulated data, we derived an optimal strategy using our methods and evaluated using the criteria defined in Section 2.3. The performance of the optimal strategy was then compared to that of a commonly used strategy (henceforth, the “routine strategy”), which requires monitoring patients' PSA values every 3 months for the first 2 years and every 6 months thereafter ([41]).

The joint model included the following variables: 1) the covariates used to predict class membership in (2.1) were tumor stage, Gleason score, and pre-treatment PSA value, 2) the biomarker value Y in (2.2) was the vector of the longitudinal measurements of PSA (in the log scale). The fixed-effect and random-effect covariates in the linear model were the linear and quadratic terms of time in months since enrollment. We did not include covariates in the model for the hazard of the risk for recurrence. The main risk factors to trigger prostate cancer recurrence are Gleason score and PSADT ([44]). However, we already stratified the biomarker trajectory based on Gleason score and PSADT is considered as a surrogate biomarker for recurrence ([46]) which should be independent of diagnosis time given the disease risk class ([17]). The parameter estimates using 5-year data are shown in the “Full Data” column in Table 1. Comparing “Full Data” column with “Parameter True Value” column, it can be seen that the parameter estimates from the proposed models (“Full Data” column) are very close to the true values of those parameters used in the simulation model (“Parameter True Value” column). The relative bias for intercept, coefficient for t , coefficient for t^2 , and baseline hazard is 0%, 2%, and 2.3% for latent class 1, 0%, 0%, 0.6%, and 2.6% for latent class 2, and 0.4%, 1.6%, 0%, and 3.9% for latent class 3, respectively.

4.2 Simulation results

(i) Computation and evaluation of the optimal strategy based on 150 patients

—In the simulated dataset, the first asymptomatic cancer recurrence event confirmed by the reference standard was at 24 months. We illustrate the computation procedure by using the simulated data up to 24 months to compute the dynamic optimal strategy for later times.

We first examine the posterior distribution of the probability of class membership after each iteration of fitting the latent class model. By plotting the histogram of the maximal posterior

probability of class membership, a peak at probability 1 indicates that components can be well separated from each other, while no peak at 1 indicates overlap with other components. Figure 1 presents such histogram using data up to 24 months. The graph shows that biomarker trajectories were well separated into 3 groups by the latent class model. Hence, it was not necessary to perform MI at 24 months. However, when the components are not well separated (as is the case in the analysis at 41 months, for example), multiple imputation of latent class indicators is needed. For the analysis at 41 months, we imputed the latent class indicator 20 times and fitted models (2.3) and (2.4) within each iteration. The parameter estimates from each imputation were combined to arrive at the final estimates. Parameter estimates for data up to 41 and 45 months are presented in Table 1. The top graph in Figure 2 presents the schedule calculated using the simulated data collected up to 24 months following the computation procedure outlined in Section 2.2.1. Based on the first 24-month data, the latent class model suggested 3 latent trajectory classes, which were labeled as the low risk group (class 1), intermediate risk group (class 2), and high risk group (class 3). For patients in the high risk group, the next monitoring time was 29 months after the end of treatment, and patients in other risk groups without cancer recurrence followed the same monitoring schedule as the patients in the high risk group. At 35 months some patients in class 2 experienced cancer recurrence, and we were able to fit models and calculate the optimal time for patients in this group. Since there was no cancer recurrence for patients in class 1, the monitoring schedule for patients in this class was based on the maximal time interval among all monitoring intervals for patients in all other classes with recurrent events.

Evaluation of the optimal strategy: We applied the following two evaluation criteria to the simulated dataset: 1) *Number of months earlier of monitoring detection*: all patients in class 2 and 3 were detected as having cancer recurrence at the end of 5 years. The second column in Table 2 shows that the proposed strategy could detect cancer recurrence from 1 month up to 4 months earlier than could the routine strategy for a majority of the patients. On average, cancer recurrence had been detected more than 2 months earlier by the proposed method for this simulated data and at least 50% of cancer recurrence could be detected at least 3 months earlier. 2) *Percentage of the monitoring detected length*: we also compared the PMDL between two strategies. The PMDL range was divided into the following categories, <10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, and >60%. The number of patients in each of PMDL categories was 9, 30, 18, 11, 9, 3, and 0 for the proposed strategy and 5, 19, 22, 14, 10, 10, and 0 for the routine strategy, respectively. It is clear that the proposed strategy can detect cancer recurrence at an earlier stage compared with the routine strategy. As noted in Section 2.2.2, the dynamic updating of the strategy begins only after asymptomatic recurrent events are observed. In the simulated dataset, this meant that optimal strategy began at 24 months. In order to form a complete strategy we assumed that the routine strategy will be used up to the 24 month time point on the simulated data. Hence it is not appropriate to compare the other two criteria, *frequency of monitoring tests* and *total costs*, between the routine strategy and the proposed strategy. Note that f_1 and f_2 are equal between two strategies since there were no patients with symptomatic recurrences in this simulated dataset.

Classification error rate of latent class model: The true class membership for each patient is known in this simulated dataset, and the estimated class membership was obtained from LCJM. Using 5-fold cross validation we estimated the average classification error rate (chapters 7, 9 in [47]). For this simulated dataset, the estimated average classification error rate is 0.02, 0.032, 0.034, 0.082, and 0.032 at 24, 29, 32, 35, and 41 months, respectively.

(ii) Optimal strategy and results for new patients

Optimal strategy for new patients: According to the proposed approach in Section 2.2.2, we first calculated the monitoring interval before 24 months, the time of the first asymptomatic recurrence. For patients in the high risk group, they were monitored 5 months later after 24 months, hence it is reasonable to monitor these patients at least every 5 months before 24 months. The same approach applied to patients in other risk groups as well. The bottom graph in Figure 2 presents the recommended class-based monitoring strategy for a 5-year period which can be used for new prostate cancer patients who are in remission.

Results for 50 new patients: We simulated 50 additional patients using the same assumptions and models that were used for the initial simulation, and used them as a simple example to describe how to apply the proposed strategy (see bottom graph in Figure 2) to monitor a group of new patients after their initial treatment according to the approach we discussed in Section 2.2.2. In this dataset, 40 patients had monitoring detected cancer recurrence at the end of 5 years. As we mentioned before, the first time we observed asymptomatic recurrent events in the simulated dataset with 150 patients is at 24 months. Hence, up to 24 months, these 50 patients were classified into 3 latent groups in which the class memberships were estimated using the coefficient estimates in the latent class model at time t_r , which is 24 months in this example, and patients were monitored according to the schedule of the class to which they were originally assigned. The next monitoring time, denoted by t_{r+1} , was 29 months after the end of treatment for patients in class 3 and 30 months for patients in other classes. At time t_{r+1} , the class memberships of those patients were updated using the coefficient estimates in the latent class model at time t_{r+1} , and patients were monitored according to the schedule of their updated class memberships. We repeated this procedure until the end of 5 years. We applied the routine monitoring strategy on these 50 new patients as well and compared the evaluation criteria defined in Section 2.3 between two strategies. The third column in Table 2 shows that the proposed strategy could detect cancer recurrence from 1 month up to 4 months earlier than could the routine strategy for a majority of the patients. On average, cancer recurrence had been detected at least 2 months earlier by the proposed method and more than 50% of cancer recurrence could be detected at least 3 months earlier. Table 3 shows that on average, the routine strategy needed almost 4 more tests than did the proposed method. The total cost for the proposed strategy, given by $277cost_1 + 53cost_2$, was less than the total cost for the routine strategy, given by $422cost_1 + 59cost_2$. The number of patients in each of PMDL range categories, which is <10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, and >60% was 23, 10, 2, 3, 2, 0, and 0 for the proposed strategy, and 12, 17, 4, 2, 4, 1, and 0 for the routine strategy, respectively. According to the proposed strategy, patients in class 1 and 2 were monitored only three times, and patients in class 3 were monitored four times before 24 months, which is the first time we observed the asymptomatic recurrences. However, if using the routine monitoring

strategy, all patients were monitored seven times before 24 months. The proposed strategy will save a lot of time and expense when no events happen.

5 Discussion

In this paper, we developed a method for determining an optimal monitoring strategy using latent class models with joint analysis of longitudinal biomarker and event process data. The strategy dynamically modifies the monitoring intervals based on the periodically updated risk from patients in different biomarker trajectory groups. The results from a simulated 5-year monitoring prostate cancer recurrence study suggest that the proposed strategy has advantages over routine strategies that monitor all patients at fixed intervals. In particular, the proposed strategy was shown to lead to cost savings, lower frequency of monitoring tests, and earlier detection of cancer recurrence. The adaptive strategy can be readily applied to new patients who have the same type of disease and are in remission after initial treatment. The approach can be easily generalized to other types of cancer or chronic disease where biomarkers for disease progression are available.

The choice of monitoring examination time by the proposed method is based on the underlying time course of disease progression and is directly related to the risk of disease recurrence. By identifying patients with increased risk of recurrence the approach can detect asymptomatic recurrence earlier than the routine strategy. Assuming that appropriate therapies are available, the proposed strategy has the potential to prolong the time of progression-free survival from symptom which will provide patients more time on advanced treatment. Therefore, the proposed strategy has a potential impact on prolonging patients' overall survival and reducing mortality rate. The selection of optimal strategy rests importantly on the choice of utility function. Although we used a broadly applicable utility function in this paper, more elaborate functions can be developed and used in specific disease settings.

The overall structure of the approach for deriving optimal strategies is straightforward. However its implementation entails considerable computational burden. The complexity stems from the need to repeatedly fit the LCJM and calculate the optimal time through a utility function. In view of the advantages of the proposed approach, the required computational effort seems to be readily justifiable.

The two-stage estimation procedure for fitting the LCJM is practical for the settings we study and has been shown to be consistent but not necessarily as efficient as a procedure that utilizes the joint likelihood. However, when the class membership is certain, the parameter estimation does not lead to loss of efficiency ([31]). Although we only used biomarker trajectory data in the first stage to estimate the latent class membership, the misclassification rate should be low since the trajectory describes biomarker behavior and significantly correlated with survival ([48], [49], [50], [51], [52], [53]) and the biomarker trajectory pattern has been viewed as a latent class, and moreover we used MI to account for the uncertainty of latent class membership which helps minimize loss of efficiency. Specifically in our example the estimated misclassification rate, presented in Section 4.2, is low. The

relative bias of parameter estimation results presented in Section 4.1 verified that the two-stage estimation approach could provide the reliable results.

We conducted simulation studies to assess the performance of the two-stage estimation approach. First, we simulated 200 datasets using the simulation model described in the Appendix and performed the two-stage procedure for each data set. The relative bias for baseline hazard for class 2 and class 3 was 6.5% and 5.7%, respectively, and the coverage rate of the parameter for class 2 and class 3 was 94.8% and 95.4%, respectively. Second, we compared the two-stage estimation approach to maximum likelihood estimation by examining the differences in parameter estimates from the two approaches using simulated data for a simple situation. Specifically we simulated 50 datasets with 150 observations in each dataset. There was only one covariate (Gleason score) in the latent class model, the biomarker value was assumed to be a linear function of time, and baseline hazards were assumed constant. The parameter estimates of the longitudinal biomarker model for every latent class were the same between the two approaches up to the fourth decimal place. The difference between the estimates of the baseline hazard (defined as the parameter estimate from the full likelihood approach minus the parameter estimate from the two-stage estimate approach), ranged from -0.0005 to 0.004 and -0.006 to 0.001 for latent class 2 and 3, respectively. Among those 50 differences between the estimates of the baseline hazard, 47 parameter estimates are different at the fifth decimal place for class 2 and 46 parameter estimates are different at the fifth decimal place for class 3, respectively.

Conditional independence is an important assumption in the LJCM, and its appropriateness needs to be assessed ([17], [54]). In order to investigate the robustness of results to the CI assumption we simulated 200 additional datasets after introducing dependence between the hazard function and the biomarker models, via a shared random effects model. The coefficient for the shared random effect was set equal to 1. We applied the two-stage estimation on each of the 200 datasets. The relative bias of the baseline hazard for class 3 at 24 months, 29 months, and 32 months was 2.4%, 15.9%, and 6.1%, respectively, and the relative bias of the baseline hazard for class 2 at 35 months and 41 months were 11.9% and 10.1%, respectively. However, the monitoring time intervals for those 200 simulated datasets are still the same as the monitoring time intervals obtained using the simulated data with valid CI assumption. The simulation results show that the invalid CI assumption has a moderate impact on parameter estimates, but does not affect appreciably the final monitoring strategy.

Potential non-identifiability of latent class model parameters is a well-known problem ([55]). However Teicher et al. (1963) ([56]) proved the theory that the class of all finite mixtures of normal distributions is identifiable. Moreover label switching only affects interpretation of results and does not pose a problem for parameter estimation itself ([34]). In the analysis of our example, we tried to avoid overfitting by specifying different component parameters differently, and we also checked output to make sure there were no convergence problems with multiple starting points.

In this paper, we propose to use class-based monitoring strategies. Alternatively, individual strategies can be developed as we described in Section 2.2.1. Individual strategies would

generally differ from class-based strategies if the hazard function involved individual covariates. As noted earlier, the advantage of a class-based monitoring strategy is that it can be readily applied to other patients who have similar disease but their biomarker trajectory data do not yet exist by only identifying their class memberships using model (2.1).

In Section 4, we compared the adaptive monitoring strategy with the routine monitoring strategy at fixed, predetermined intervals. Because the estimation with the latent class model implies a heavy computational burden we compared the strategy obtained with the latent class model to the strategy obtained using a simpler model in which there is no latent class model involved, i.e. all patients follow the same biomarker model and hazard model. The optimal monitoring strategy from this simpler model required monitoring patients at 6, 12, 18, 24, 30, 36, 42, 48, 54, and 60 months. This strategy has the same monitoring schedule as the routine strategy after 24 months and thus would have the same performance after 24 months as the routine strategy in terms of PMDL and the number of months earlier of monitoring detection. The proposed strategy can detect cancer recurrence on average at least 2 months earlier and has shorter PMDL comparing with the results from the simpler model although the monitoring strategy developed by the simpler model has similar cost, which is $262cost_1 + 59cost_2$. Through this comparison, we verified one more time that the proposed monitoring strategy using LJCM has better performance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Work partially supported by NCI grant U01 CA079778 (PI: C Gatsonis).

Appendix

Simulated observations for biomarker values and true recurrence status were generated using the following assumptions:

1. The recurrence time $T \sim F(T) = 1 - \exp(-\lambda T)$. The indicator of true recurrence status D was set to 1 at any monitoring time t after T and to 0 before T , where $\lambda=0.027$ and 0.048 for latent class 2 and 3, respectively,
2. For $T < t_{max}$, the changepoint for the biomarker value $\tau \sim \text{Uniform}(T, T + 3)$,
3. At any monitoring time t the values of the log of biomarker (Y) were simulated as follows: $\log(Y) \sim \text{Normal}(w_0 + w_1 t + w_2 t^2, \sigma^2)$, where $\sigma^2 = 1e - 04$ and all w s are constants plus a random variation which follows normal distribution
for latent class 1: $w_0 \sim 0.26 + \text{Normal}(0, 0.003^2)$, $w_1 \sim 5e-04 + \text{Normal}(0, (3e - 04)^2)$, and $w_2 \sim 1.3e-05 + \text{Normal}(0, (3e - 06)^2)$,
for latent class 2: $w_0 \sim 0.4 + \text{Normal}(0, 0.003^2)$, $w_1 \sim 0.05 + \text{Normal}(0, (3e - 03)^2)$, and $w_2 \sim 2.86e-03 + \text{Normal}(0, (1e - 04)^2)$,

for latent class 3: $w_0 \sim 0.5 + \text{Normal}(0, 0.01^2)$, $w_1 \sim 0.025 + \text{Normal}(0, (3e - 03)^2)$, and $w_2 \sim 6e-3 + \text{Normal}(0, (1e - 04)^2)$,

4. If the indicator of true recurrence status D was equal to 1 and the test result was positive, then the indicator of monitoring detected status was set to 1 and 0 otherwise,
5. The definition of test results is ([57])
if age was less than 60 then
a) if PSA was larger than or equal to 2.5ng/ml then test result was positive, b) if PSA was less than 2.5ng/ml then test result was negative,
if age was larger than or equal to 60 then
a) if PSA was larger than or equal to 4ng/ml then test result was positive, b) if PSA was less than 4ng/ml then test result was negative,
6. Age at enrollment $\sim \text{Normal}(66, 8^2)$,
7. The following covariates were simulated according to literature review:
for latent class 1: Gleason score, pre-treatment PSA, and tumor stage is randomly selected within [2,6], [2.5, 10]ng/ml, and [T1, T2b], respectively,
for latent class 2: Gleason score, pre-treatment PSA, and tumor stage is randomly selected within [2,7], [10, 20]ng/ml, and [T1, T2c], respectively,
for latent class 3: Gleason score, pre-treatment PSA, and tumor stage is randomly selected within [6,9], 20ng/ml, and [T2c, T4], respectively.

References

1. Pollack LA, Greer GE, Rowland JH, Miller A, Doneski D, Coughlin SS, Stovall E, Ulman D. Cancer Survivorship: a New Challenge in Comprehensive Cancer Control. *Cancer Causes and Control*. 2005; 16(suppl. 1):51–59. [PubMed: 16208574]
2. Edelman MJ, Meyers FJ, Siegel D. The Utility of Follow-up Testing after Curative Cancer Therapy. *J Gen Int Med*. 1997; 12:318–331.
3. Kattlove H, Winn RJ. Ongoing Care of Patients after Primary treatment for Their Cancer. *CA Cancer J Clin*. 2003; 53:172–196. [PubMed: 12809410]
4. Pfister DG, Berson AB III, Somerfield MR. Surveillance Strategies after Curative Treatment of Colorectal Cancer. *The New England Journal of Medicine*. 2004;2375–2382. [PubMed: 15175439]
5. Rodriguez-Moranta F, Salo J, Arcusa A, Boadas J, Pinol V, Bessa X, Batiste-Alentorn E, Lacy AM, Delgado S, Maurel J, Piqu JM, Casells A. Postoperative Surveillance in Patients with Colorectal Cancer Who Have Undergone Curative Resection: A Prospective, Multicenter, Randomized, Controlled Trial. *J Clin Oncol*. 2006; 24:386–393. [PubMed: 16365182]
6. Sugiyama T, Hirose T, Hosaka T, Kusumoto S, Nakashima M, Yamaoka T, Okuda K, Ohmori T, Adachi M. Effectiveness of Intensive Follow-up after Response in Patients with Small Cell Lung Cancer. *Lung Cancer*. 2008; 59:255–261. [PubMed: 17900754]
7. Berman JM, Cheung RJ, Weinberg DS. Surveillance after Colorectal Cancer Resection. *The Lancet*. 2000; 355:395–399.
8. Kievit J. Colorectal Cancer Follow-up: a Reassessment of Empirical Evidence on Effectiveness. *European Journal of Surgical Oncology*. 2000; 26:322–328. [PubMed: 10873350]

9. Glasziou PP, Irwig L, Mant D. Monitoring in Chronic Disease: a Rational Approach. *BMJ*. 2005; 330:644–648. [PubMed: 15774996]
10. Parker CB, DeLong ER. ROC Methodology within a Monitoring Framework. *Statistics in Medicine*. 2003; 22:3473–3488. [PubMed: 14601014]
11. Pauler DK, Finkelstein DM. Predicting Time to Prostate Cancer Recurrence Based on Joint Models for Non-linear Longitudinal Biomarkers and Event Time Outcome. *Statistics in Medicine*. 2002; 21:3897–3911. [PubMed: 12483774]
12. Taylor JMG, Yu M, Sandler HM. Individualized Predictions of Disease Progression Following Radiation Therapy for Prostate Cancer. *J Clin Oncol*. 2005; 23:816–825. [PubMed: 15681526]
13. Thompson IM, Ankerst DP, Chi C, Lucia MS, Goodman PJ, Crowley JJ, Parnes HL, Coltman CA. Operating Characteristics of Prostate-Specific Antigen in Men with an Initial PSA Level of 3.0 ng/ml or Lower. *JAMA*. 2005; 294:66–70. [PubMed: 15998892]
14. Proust-Lima C, Taylor JMG. Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach. *Biostatistics*. 2009; 10:535–549. [PubMed: 19369642]
15. Duffy SW, Chen HH, Tabar L, Day N. Estimation of Mean Sojourn Time in Breast Cancer Screening Using a Markov Chain Model of Both Entry to and Exit from the Preclinical Detectable Phase. *Statistics in Medicine*. 1995; 14:1531–1543. [PubMed: 7481190]
16. Lin H, McCulloch CE, Turnbull BW, Slate EH, Clark LC. A Latent Class Mixed Model for Analysing biomarker Trajectories with Irregularly Scheduled Observations. *Statistics in Medicine*. 2000; 19:1303–1318. [PubMed: 10814979]
17. Lin H, Turnbull BW, McCulloch CE, Slate EH. Latent Class Models for Joint Analysis of Longitudinal Biomarker and Event Process Data: Application to Longitudinal Prostate-Specific Antigen Readings and Prostate Cancer. *Journal of the American Statistical Association*. 2002; 97:53–65.
18. Tsodikov AD, Asselain B, Fourque A, Hoang T, Yakovlev AY. Discrete strategies of cancer post-treatment surveillance. Estimation and Optimization problems. *Biometrics*. 1995; 51:437–447. [PubMed: 7662836]
19. Kent DL, Shachter R, Sox HC, Hui NS, Shortliffe LD, Moynihan S, Torti FM. Efficient Scheduling of Cystoscopies in Monitoring for Recurrent Bladder. *Cancer Med Decis Making*. 1989; 9:26–37.
20. Parmigiani G. On optimal screening ages. *Journal of the American Statistical Association*. 1993; 88:622–628.
21. Taha, HA. *Operations Research An Introduction*. Eighth. Pearson Prentice Hall; 1997.
22. Skates SJ, Pauler DK, Jacobs JJ. Screening Based on the Risk of Cancer Calculation from Bayesian Hierarchical Change-point and Mixture Models of Longitudinal Markers. *Journal of the American Statistical Association*. 2001; 96:429–439.
23. Parmigiani, G. *Modeling in Medical Decision Making A Bayesian Approach*. John Wiley and Sons, LTD; 2002.
24. D'Amico AV, Hanks GE. Linear Regressive Analysis Using Prostate-Specific Antigen Doubling Time for Predicting Tumor Biology and Clinical Outcome in Prostate Cancer. *Cancer*. 1993; 72:2638–2643. [PubMed: 7691393]
25. Parker C. Active Surveillance of Early Prostate Cancer: Rationale, Initial Results and Future Developments. *Prostate Cancer and Prostatic Disease*. 2004; 7:184–187.
26. Shih JH, Louis TA. Inferences on association parameter in copula models for bivariate survival model. *Biometrics*. 1995; 51:1384–1399. [PubMed: 8589230]
27. Glidden DV. A two-stage estimator of the dependence parameter for the Clayton-Oakes model. *Lifetime Data Analysis*. 2000; 6:141–156. [PubMed: 10851839]
28. Andersen EW. Composite likelihood and two-stage estimation in family studies. *Biostatistics*. 2004; 5:15–30. [PubMed: 14744825]
29. Jain D, Bass FM, Chen Y. Estimation of Latent Class Models with Heterogeneous Choice Probabilities: An Application to Market Structuring. *Journal of Marketing Research*. 1990; 27:94–101.

30. Reboussin BA, Miller ME, Lohman KK. Latent class models for longitudinal studies of the elderly with data missing at random. *Appl Statist.* 2002; 51:69–90.
31. Putter H, Vos T, Haes H, Houwelingen H. Joint Analysis of Multiple Longitudinal Outcomes: Application of a latent Class Model. *Statistics in Medicine.* 2008; 27:6228–6249. [PubMed: 18816496]
32. Vermunt JK. Latent Class Modeling with Covariates: Two Improved Three-step Approaches. *Political Analysis.* 2010; 18:450–469.
33. Liang Y, Lu W, Ying Z. Joint Modeling and Analysis of Longitudinal Data with Informative Observation Times. *Biometrics.* 2009; 65:377–384. [PubMed: 18759841]
34. Leisch, F. Exploring the Structure of Mixture Model Components. *Physica Verlag; Heidelberg:* 2004. p. 1405-1412.
35. Arminger, G.; Clogg, CC.; Sobel, ME. *Handbook of Statistical Modeling for the Social and Behavioral Sciences.* Plenum Press; 1995.
36. Loken E. Using Latent Class Analysis to Model Temperament Types. *Multivariate Behavioral Research.* 2004; 4:625–652.
37. Harel O, Miglioretti D. Missing Information as a Diagnostic Tool for Latent Class Analysis. *Journal of Data Science.* 2007; 5:269–288.
38. Chung H, Flaherty BP, Schafer J. Latent Class Logistic Regression: Application to Marijuana Use and Attitudes among High School Seniors. *Journal of the Royal Statistical Society Series A.* 2006; 169:723–743.
39. Little RJA, Rubin DB. *Statistical Analysis with Missing Data.* 2002 Statistical Analysis with Missing Data.
40. Zelen M, Lee S. Models and the early detection of disease: methodological considerations. *Cancer treatment and research.* 2002; 113:1–18. [PubMed: 12613347]
41. LAPCW Lions and Australian and Prostate and Cancer and Website. Prostate Cancer: Monitoring After Treatment. 2007. <http://www.prostatehealth.org.au>
42. Hiranmanek N. Breast Cancer Recurrence: Follow Up after Treatment for Primary Breast Cancer. *Postgrad Med J.* 2004; 80:172–176. [PubMed: 15016942]
43. Tuxen MK, Soletormos G, Dombernowsky P. Serum Tumour Marker CA125 in Monitoring of Ovarian Cancer during First-Line Chemotherapy British. *Journal of Cancer.* 2001; 84:1301–1307.
44. Koch MO, Foster RS, Bell B, Beck S, Cheng L, Parekh D, Jung S. Characterization and Predictors of Prostate Specific Antigen Progression Rates after Radical Retropubic Prostatectomy. *The Journal of Urology.* 2000; 164:749–753. [PubMed: 10953139]
45. Blumenstein BA, Ellis WJ, Ishak LM. The Relationship between Serial Measurements of the Level of a Bladder Tumor Associated Antigen and the Potential for Recurrence. *The Journal of Urology.* 1999; 161:57–61. [PubMed: 10037368]
46. Arlen PM, Bianco F, Dahut WL, D'Amico A, Figg WD, Freedland SJ, Gulley JL, Kantoff PW, Kattan MW, Lee A. Prostate-Specific Antigen Working Group Guidelines on PSA Doubling Time. *J Urol.* 2008; 179:2181–2186. [PubMed: 18423743]
47. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning.* Springer; 2001.
48. Molenberghs G, Burzykowski T, Alonso A, Buyse M. A Perspective on Surrogate Endpoints in Controlled Clinical Trials. *Statistical Methods in Medical Research.* 2004; 13:177–206. [PubMed: 15198486]
49. Dixon MR, Haukoos JS, Udani SM, Naghi JJ, Arnell TD, Kumar RR, Stamos MJ. Carcinoembryonic Antigen and Albumin Predict Survival in Patients with Advanced Colon and Rectal. *Cancer Arch Surg.* 2003; 138:962–966.
50. Cooper BC, Sood AK, Davis CS, Ritchie JM, Sorosky JI, Anderson B, Buller RE. Preoperative CA 125 Levels: An Independent Prognostic Factor for Epithelial Ovarian Cancer. *The American College of Obstetricians and Gynecologists.* 2002; 100:59–64.
51. Bairey O, Blickstein D, Stark P, Prokocimer M, Nativ HM, Kirgner I, Shaklai M. Serum CA 125 as a Prognostic Factor in Non-Hodgkin's Lymphoma. *Leukemia and Lymphoma.* 2003; 44:1733–1738.

52. Bender DP, Sorosky JJ, Buller RE, Sood AK. Serum CA 125 Is an Independent Prognostic Factor in Cervical Adenocarcinoma. *Am J Obstet Gynecol*. 2003; 189:113–117. [PubMed: 12861147]
53. Munstedt K, Krisch M, Sachsse S, Vahrson H. Serum CA 125 levels and Survival in Advanced Ovarian Cancer. *Arch Gynecol Obstet*. 1997; 259:117–123. [PubMed: 9187463]
54. Proust-Lima C, Joly P, Dartigues J, Jacqmin-Gadda H. Joint Modeling of Multivariate Longitudinal Outcomes and a Time-to-event: A nonlinear latent class approach. *Computational Statistics and Data Analysis*. 2009; 53:1142–1154.
55. Bandeen-Roche K, Miglioretti DL, Zeger SL, Rathouz PJ. Latent Variable Regression for Multiple Discrete Outcomes. *Journal of American Statistical Association*. 1997; 92:1375–1386.
56. Teicher H. Identifiability of Finite Mixtures. *The Annals of Mathematical Statistics*. 1963; 34:1265–1269.
57. Gilbert SM, Cavallo CB, Kahane H, Lowe FC. Evidence Suggesting PSA Cutpoint of 2.5 ng/ml for Prompting Prostate Biopsy: Review of 36316 Biopsies. *Urology*. 2004; 65:549–553. [PubMed: 15780374]
58. Klein JP, Klotz JH, Grever MR. A Biological Marker Model for Predicting Disease Transitions. *Biometrics*. 1984; 40:927–936. [PubMed: 6598390]
59. Parmigiani G, Skates S, Zelen M. Modeling and Optimization in Early Detection Programs with a Single Exam. *Biometrics*. 2002; 58:30–36. [PubMed: 11890325]

Figure 1. Histogram Using Data Up to 24 Months

Figure 2. Monitoring Schedules

Table 1

Parameter Estimates for the Joint Model

Class	Model	Parameter	Parameter True Value	Full Data 60 Months est. (se)	Simulated Data up to 41 Months est. (se)	Simulated Data up to 45 Months est. (se)	
1	longitudinal	intercept	0.26 (3e-03)	0.26 (3e-03)	0.258 (3e-03)	0.258 (3e-03)	
		γ_{11}	5e-04 (3e-04)	5.1e-04 (3.3e-04)	5e-04 (3.1e-04)	5e-04 (3e-04)	
		γ_{12}	1.3e-05 (3e-06)	1.33e-05 (3.2e-06)	1.29e-05 (3.1e-06)	1.3e-05 (3e-06)	
	latent class	intercept	-	35.31 (1.9)	35.31 (1.9)	35.31 (1.9)	
		ω_{11}	-	-0.57 (0.32)	-0.57 (0.32)	-0.57 (0.32)	
		ω_{12}	-	-4.8 (1.48)	-4.8 (1.48)	-4.8 (1.48)	
		ω_{13}	-	-1.07 (0.72)	-1.07 (0.72)	-1.07 (0.72)	
2	longitudinal	intercept	0.4 (0.003)	0.4 (0.003)	0.4 (0.003)	0.4 (0.003)	
		γ_{21}	5e-02 (3e-03)	5e-02 (3e-03)	5e-02 (3.6e-03)	5e-02 (3e-03)	
		γ_{22}	2.86e-03 (1e-04)	2.84e-03 (1e-04)	2.84e-03 (8e-05)	2.84e-03 (8.4e-05)	
	latent class	intercept	-	121.48 (2.4)	121.48 (2.4)	121.48 (2.4)	
		ω_{21}	-	-8.82 (2.4)	-8.82 (2.4)	-8.82 (2.4)	
		ω_{22}	-	2.7 (1.2)	2.7 (1.2)	2.7 (1.2)	
		ω_{23}	-	-2.5 (1.16)	-2.5 (1.16)	-2.5 (1.16)	
	hazard	δ_{200}	0.027	0.0277 (5e-03)	0.025 (0.02)	0.024 (0.009)	
	3	longitudinal	intercept	0.5 (0.01)	0.498 (0.014)	0.5 (0.01)	0.5 (0.01)
			γ_{31}	2.5e-02 (3e-03)	2.46e02 (3e-03)	2.45e-02 (3e-03)	2.45e-02 (3e-03)
γ_{32}			6e-03 (1e-04)	6e-03 (1e-04)	6e-03 (1.3e-04)	6e-03 (1.2e-04)	
latent class		intercept	-	-	-	-	
		ω_{31}	-	-	-	-	
		ω_{32}	-	-	-	-	
		ω_{33}	-	-	-	-	
hazard		δ_{300}	0.048	0.0461 (0.005)	-	-	
1-sensitivity		β_{10}	-	0.258 (0.18)	0.49 (0.28)	0.59 (0.23)	

Table 2
The Number of Months Earlier of Detecting Cancer Recurrence by the Proposed Strategy
Comparing with the Routine Strategy for Class 2 and 3

Number of Months Earlier	Simulated Data (80) (Number of Patients)	New Patients (40) (Number of Patients)
0	7	2
1	30	17
3	33	15
4	7	6

Table 3
The Frequency of Monitoring Tests for 40 New Patients in Class 2 and 3

Test Frequency	Routine Schedule (Number of Patients)		Proposed Strategy (Number of Patients)	
	Class 2	Class 3	Class 2	Class 3
5				3
6				9
7			5	8
8		3	11	
9		9	2	
10		8	2	
11	5			
12	13			
13	2			