Optimal Screening Schedules For Disease Progression, With Application to Diabetic Retinopathy

IONUT BEBU*, JOHN M. LACHIN
The Biostatistics Center, The George Washington University
6110 Executive Blvd., Rockville MD 20852, USA ibebu@bsc.gwu.edu

October 4, 2016

Abstract

Clinical management of chronic diseases requires periodic evaluations. Subjects transition between various levels of severity of a disease over time, one of which may trigger an intervention that requires treatment. For example, in diabetic retinopathy, patients with type 1 diabetes are evaluated yearly for either the onset of proliferative diabetic retinopathy (PDR) or clinically significant macular edema (CSME) that would require immediate treatment to preserve vision. Herein we investigate methods for the selection of personalized cost-effective screening schedules and compare them with a fixed visit schedule (e.g. annually) in terms of both cost and performance. The approach is illustrated using the progression of retinopathy in the DCCT/EDIC study.Optimal screening schedule; Markov models; undetected time; diabetic retinopathy.

1 Introduction

Periodic evaluations are common in the management of chronic diseases, and informed evaluation of their frequency is of interest. For example, the current guidelines for screening for retinopathy in type 1 diabetes (T1D) recommend yearly visits. The Diabetes Control and Complications Trial (DCCT) and its follow-up Epidemiology of Diabetes Interventions and Complications (EDIC) provide a unique opportunity to re-evaluate this recommendation in a well characterized cohort with available fundus photography for over thirty years [1, 4].

The health status at each visit is typically categorized into one of several possible states. For example, diabetic retinopathy status is assessed on the ETDRS scale [5], which is an ordered scale. Based on this, five clinically meaningful retinopathy states can be defined: 1) no diabetic retinopathy (DR) and no clinically significant macular edema (CSME); 2) mild non-proliferative diabetic retinopathy (MiNPDR) and no CSME; 3) moderate non-proliferative diabetic retinopathy (MoNPDR) and no CSME; 4) severe proliferative diabetic retinopathy (SNPDR) and no CSME; and 5) proliferative diabetic retinopathy (PDR) or CSME, either of which requires treatment to preserve vision. Thus, State 5 (PDR/CSME) is the state that

[0]To whom correspondence should be addressed.

triggers clinical intervention, and is considered as an absorbing state in the Markov models employed herein. Also note that a subject cannot jump from state 1 (no DR) to state 3 (MoNPDR) without first going through state 2 (MiNPDR). It is important to note that the health status is observed only at the scheduled visits, and therefore the exact transition times are unknown. Moreover, a subject can either progress from state 2 to state 3 or regress to state 1. To account for these possible transitions, a Markov model in continuous time [2, 7] can be employed to estimate the cumulative incidence functions for each state, and to then focus on transitions to the absorbing state, at which point interventions are required.

Multi-state Markov models have been employed to identify risk factors for diabetic retinopathy [9, 10]. Fixed schedules (e.g., annual visits) have also been considered [3]. However, our goal is to take this one step further and to provide rational examination schedules based on the patient's risk profile.

Herein screening schedules are compared in terms of two cost components of particular interest in this context. The first is the undetected time, defined as the elapsed time from the actual onset of progression (assumed to occur in continuous time) and the next visit at which it is detected. The less frequent the visits, the longer progression will go undetected and the harder or more costly it may be to treat when finally detected. The second component is the number of visits or screening examinations up to and including when the progression is detected. The more frequent the visit schedule, the greater the number of negative visits before progression is detected.

The goal is to introduce a screening schedule that minimizes both the undetected time and the number or frequency of examinations, and the associated costs, and at the same time is practical and easy to explain. Both fixed screening schedules (e.g. annual visits) and personalized screening schedules based on the current health state are investigated and illustrated using the progression of retinopathy in the DCCT/EDIC study.

2 The Model

Let $S = \{u_1, \dots u_m\}$ be the set of all health states, where m is the number of states ($m \geq 2$). A subject is followed over time, and at each time point can be in any of the m health states. Denote by $X(t)$ the state of the subject at time t, $t \geq 0$. The transition probability matrix $P(s,t)$ with entries

$$p_{uv}(s,t) = P(X(t) = v | P(X(s) = u) \ ,$$

can be defined in terms of the transition intensity matrix $Q(t)$ with elements

$$q_{uv}(t) = lim_{\Delta t \to 0} p_{uv}(t, t + \Delta t)/\Delta t \ , \quad u \neq v.$$

Then $P(s,t)$ and $Q(t)$ are related by the Kolmogorov differential equations (forward/backward) [2]. A homogeneous Markov process (i.e., $p_{uv}(s,t) = p_{uv}(t-s)$ or constant intensities $\{q_{uv}\}$) is employed,

$$P(\Delta t) = \exp\{Q \cdot \Delta t\} \, , \tag{1}$$

with elements $p_{uv}(\Delta t)$, $u, v \in S$.

The state requiring the intervention is denoted by u^* and it is assumed absorbing (i.e., $q_{u^*v} = 0$, for all $v \in S$).

The focus is on the time to the absorbing state, in other words to estimate the cumulative distribution function for time from state u to state u^*, denoted by F_u ($u \in S$, $u \neq u^*$). A closed form expression for F_u is available for homogeneous Markov processes. Briefly, using the spectral decomposition of the intensity matrix Q,

$$Q = L \cdot D \cdot R\,,$$

where $R = L^{-1}$ and D is a diagonal matrix with elements λ_i, then one has

$$F_u(t) = \sum_{v \neq u^*} f_{uv} \cdot e^{\lambda_v \cdot t}\,, \tag{2}$$

where $f_{uv} = L_{uv} \cdot R_{vu^*}$.

The intensity matrix of the Markov model can incorporate covariates, and their effect is expressed as hazard ratios,

$$q_{uv}(X) = q^0_{uv} \cdot \exp(\beta'_{uv} \cdot X)\,,$$

where X denotes the covariates, and β_{uv} the corresponding hazard ratios. This can be taken into account for developing personalized screening schedules.

All parameters are estimated by maximizing the likelihood [8].

Let L denote the time horizon, say $L = 20$ years. The total cost over this follow-up period is based on the number of visits N_v, and the undetected time T_U. With c_v denoting the cost of each visit, and c_u the cost of one unit (e.g., year) of undetected time, the expected total cost becomes:

$$E(C) = c_v \cdot E(N_v) + c_u \cdot E(T_U).$$

When comparing two screening schedules, one is preferable if it performs better (i.e., lower T_U) and is less costly (i.e., lower $E(C)$).

3 Fixed Visit Schedules

Consider the case of a single sub-clinical state not requiring treatment that can lead to progression to a clinical state requiring treatment. Let F denote the cdf of T, the time from the sub-clinical state to the clinical state. A screening schedule, denoted by $\tau = (\tau_k)_{k=1,\ldots K}$ is a partition of the interval $[0, L]$, so that $0 = \tau_1 < \ldots < \tau_K = L$, where L is the time horizon and $K \geq 2$.

The expected value of the undetected time is

$$\begin{aligned} E(T_U) &= \sum_{k=1}^{K-1} \int_{\tau_k}^{\tau_{k+1}} (\tau_{k+1} - t)\ \mathrm{d}F(t) \\ &= \sum_{k=1}^{K-1} \tau_{k+1} \cdot [F(\tau_{k+1}) - F(\tau_k)] - \int_0^L t\,\mathrm{d}F(t)\,. \end{aligned} \tag{3}$$

One can see that, for a different partition $\xi = (\xi_j)_{j=1,\ldots J}$ of the interval $[0, T]$ ($0 = \xi_1 < \ldots < \xi_J = L$, the difference in undetected time between the two visit schedules is

$$E(T_U(\tau)) - E(T_U(\xi)) = \tau'_{2:K} \cdot \Delta F(\tau) - \xi'_{2:K} \cdot \Delta F(\xi)\,, \tag{4}$$

where $\tau_{2:K} = (\tau_2, \ldots, \tau_K)'$ and $\Delta F(\tau) = (F(\tau_2) - F(\tau_1), \ldots, F(\tau_K) - F(\tau_{K-1}))'$ (and similar notations for $\xi_{2:J}$ and $\Delta F(\xi)$).

The expected number of visits before the onset $N_v(\tau)$ is

$$E(N_v(\tau)) = 1 + \sum_{k=1}^{K-1} k \cdot [F(\tau_{k+1}) - F(\tau_k)] + (K-1) \cdot [1 - F(L)]\,, \tag{5}$$

and therefore

$$\begin{aligned} E(N_v(\tau) - N_v(\xi)) &= (1:(K-1))' \cdot \Delta F(\tau) - (1:(J-1))' \cdot \Delta F(\xi) \\ &+ (K-J) \cdot (1 - F(L))\,. \end{aligned}$$

The difference in expected costs between two visit schedules τ and ξ is

$$E(C(\tau)) - E(C(\xi)) = c_v \cdot [E(N_v(\tau)) - E(N_v(\xi))] + c_u \cdot [E(T_U(\tau)) - E(T_U(\xi))]$$

Assuming $\xi \subseteq \tau$, the visit schedule τ is more effective, i.e., $E(T_U(\tau)) \leq E(T_U(\xi))$. It is also less costly iff

$$\frac{c_u}{c_v} \geq \frac{E(N_v(\tau)) - E(N_v(\xi))}{E(T_U(\xi)) - E(T_U(\tau))}\,. \tag{6}$$

4 Personalized Screening Schedules

Two personalized screening approaches are presented. First, given the unit costs c_u and c_v, a screening schedule can be obtained as the solution of an optimization problem, illustrated by minimizing the expected cost. The second approach is to select the time to the next visit such that the risk of reaching the absorbing state is below an acceptable risk (say 5%).

4.1 Minimal Expected Cost

Assume that at visit k ($k \geq 1$), a subject is in state u ($u \in S$), denoted by $X_k = u$. The decision regarding the time to the next visit is based on the retinopathy level at the current visit, namely the state u. The action at visit k is to follow up the patient in a years, denoted by $A_k = a$. The probability of the state at the next visit depends on the current state and the action a,

$$P(X_{k+1} = v | X_k = u,\, A_k = a) = p_{u,v}(a)\,,$$

where $p_{u,v}(a)$ is defined in (1).

The total cost associated with the action a for a subject in state u is computed first conditionally on the state v reached after the a years. There are two costs associated with transitioning from u to v under action a. The first one is the cost per year associated with the timing of the next visit, namely

$$C_1(a|u,v) = \frac{c_v}{a}\,,$$

which does not depend of the state v. The second type of cost $C_2(u,v|a)$ is due to the (expected) undetected time. Notice that this cost is zero unless the patient is in the absorbing state at the next visit (i.e., $C_2(u,v|a) = 0$ for $v \neq u^*$), while

$$C_2(a|u,u^*) = c_u \cdot \int_0^a (a-t)\mathrm{d}F_u(t)\,,$$

where $F_u(\cdot)$ denotes the cdf of the time to the absorbing state starting from state u.

The total cost of action a when in state u conditional on reaching the state v is

$$C(a|u,v) = \begin{cases} \frac{c_v}{a} + c_u \cdot \int_0^a (a-t)\mathrm{d}F_u(t)\,, & \text{if } v = u^* \quad \text{(with probability } p_{uu^*}(a))\,, \\ \frac{c_v}{a} & \text{if } v \neq u^* \quad \text{(with probability } 1 - p_{uu^*}(a))\,. \end{cases}$$

Then, unconditionally over the set of possible states reached, the expected total cost is

$$\begin{aligned} C(a|u) &= \frac{c_v}{a} + c_u \cdot \int_0^a (a-t)\mathrm{d}F_u(t) \cdot p_{u,u^*}(a) \\ &= \frac{c_v}{a} + c_u \cdot \int_0^a (a-t)\mathrm{d}F_u(t) \cdot F_u(a)\,. \end{aligned} \tag{7}$$

Since the time to the next visit is bounded (i.e., a is bounded), there is a value $a^*(u)$ that minimizes the total cost $C(a|u)$, and the values so obtained define the optimal screening schedule $A^* = (a^*(1), \ldots, a^*(m))$. Notice that, for example, one could define the optimal screening schedule employing a weighted sum of C_1 and C_2. However, this would lead to a similar objective function but with different values for c_u and c_v.

4.2 Limiting Risk of Undetected Time

One difficulty that arises when using the previous approach is the need to specify values for c_v and c_u. While the visit cost c_v can be readily obtained (e.g., the cost of the ophtalmoscopy and fundus photography in the retinopathy example), it is rather difficult to elicit values for the undetected time cost c_u. More importantly, while intuitive from a payer's perspective, it is less clear how meaningful the previous cost-benefit analysis is to a particular patient who is more likely interested in a screening schedule that minimizes the undetected time. Thus, a different approach is to choose the time to the visit such that the probability of reaching the absorbing state (which requires an intervention) is below a certain cutoff value, deemed as an acceptable risk (e.g., 5% or 10%). In applications, other considerations (such as practical constraints) may play a role as well in choosing a certain screening policy.

For each state u, the time to the next visit $a(u)$ is specified, so that the time to the next visit vector $A = (a(u_1), \ldots, a(u_m))$ defines a screening schedule. This leads to a Markov chain (in discrete time) $\tilde{X}$ with the same states and with transition matrix $\tilde{P}$ obtained as follows. Transition probabilities from state u are given by the state probabilities of the Markov process X after $a(u)$ years starting from state u, or equivalently the u-th row of $P(a(u))$ (see Eq. (1)), namely

$$\tilde{P}_{u,v} = p_{u,v}(a(u))\,,\ u, v = u_1, \ldots, u_m\,.$$

Notice that $\tilde{P}$ further depends on A, but for simplicity this is suppressed in the notation.
As before, two screening schedules A and A' are compared with respect to expected number of visits $N_v(A)$ and $N_v(A')$ and the undetected times $T_U(A)$ and $T_U(A')$. However, the undetected time is 0 unless the subject reaches the absorbing state at the next visit, and therefore it is more meaningful for both the patients and the clinicians to estimate the undetected time conditional on reaching the absorbing state at the next visit.

Let k ($k \geq 2$) denote the visit when this occurred, so $\tilde{X}_k = u^*$, and denote by γ_u the probability of reaching the absorbing state *from* state u under regime A. Bayes' formula gives

$$\begin{aligned} \gamma_u &= \tilde{P}(u, u^*) \Big/ \sum_v \tilde{P}(v, u^*) \\ &= F_u(a(u)) \Big/ \sum_v F_v(a(v))\,,\ u, v \neq u^*\,. \end{aligned}$$

The expected undetected time conditional on reaching the absorbing state is:

$$E(T_U | S_k = u^*) = \sum_{u \neq u^*} E(T_U | S_{k-1} = u,\ S_k = u^*) \cdot \gamma_u\,, \tag{8}$$

where

$$E(T_U | S_{k-1} = u,\ S_k = u^*) = \frac{1}{F_u(a(u))} \cdot \int_0^{a(u)} (a(u)) - t) dF_u(t)\,.$$

Another measure of interest is the probability that the unobserved time will exceed a clinically meaningful length of time t^*. A natural choice for the retinopathy example might be 0.5 years, which is the expected unobserved time under annual screening assuming a uniform distribution. One has:

$$\begin{aligned} P(T_U > t^* | S_k = u^*) &= \sum_{u \neq u^*} P(T_U > t^* | S_{k-1} = u,\ S_k = u^*) \cdot \gamma_s \\ &= \sum_{u \neq u^*} \frac{F_u\left(a(u) - t^*\right)}{F_u(a(u)))} \cdot \gamma_u\,. \end{aligned} \tag{9}$$

The number of visits within a L-year horizon under a specific screening schedule is obtained using an imbedding approach [6], see Appendix for details, or through simulations under this Markov model.

5 Illustration: Screening for Retinopathy in DCCT/EDIC

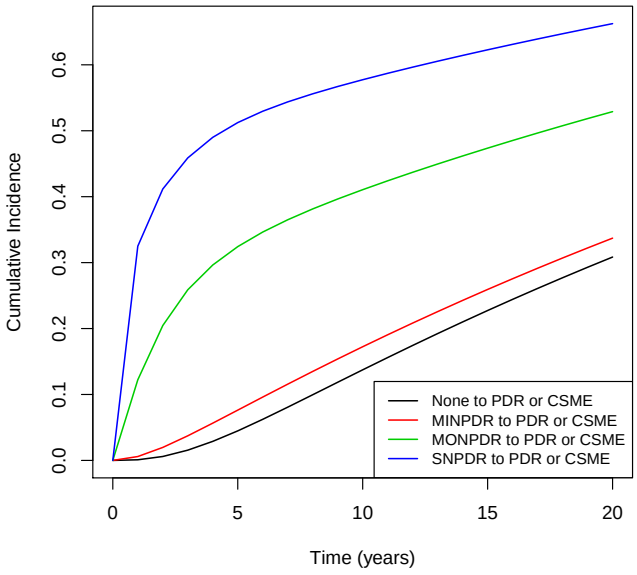
From 1983 to 1989, the DCCT enrolled 1441 participants with type 1 diabetes, and standard seven-field fundus photographs were obtained every 6 months during DCCT and every fourth year during EDIC, and in the complete cohort during EDIC years 4 and 10.

The maximum likelihood estimator for the intensity matrix Q is used to estimate the transition matrix $P(t)$ in (1). The cumulative incidence functions of reaching the absorbing state (PDR/CSME) are obtained using the eigenvalues and eigenvectors of Q and are depicted in Figure 1. As expected, these are lower when starting from states 1 and 2 compared to starting from states 3 and 4.

5.1 Fixed Schedule

For illustration, we compare annual visits versus biennial (every other year) visits. Using (3), the expected undetected time for annual visits is 0.153 years, while for biennial visits is 0.303 years. Annual screening leads to approximately 18.37 expected visits over 20 years, while for biennial visits yields approximately 9.76 visits. Using (6), annual visits are less costly than visits every two years if the ratio c_u/c_v is 33.86 or higher.

Figure 1: Incidence functions for PDR/CSME, stratified by the initial state.



5.2 Personalized Schedule

The first approach is illustrated using $c_v = 1$, and $c_u = 10, 30$. The total costs are computed using Eq. (7), and are depicted in Figure 2 for various values of the time to the next visit a. The optimal screening is given by the values that minimize the total cost in Figure 2, namely (5, 4, 1.5, 1) for $c_u = 10$ and (4.5, 3.5, 1.25, 0.63) for $c_u = 30$. More frequent visits are required as the severity of the current retinopathy state increases. Furthermore, the larger the ratio c_u/c_v, the shorter the optimal time to the next visit.

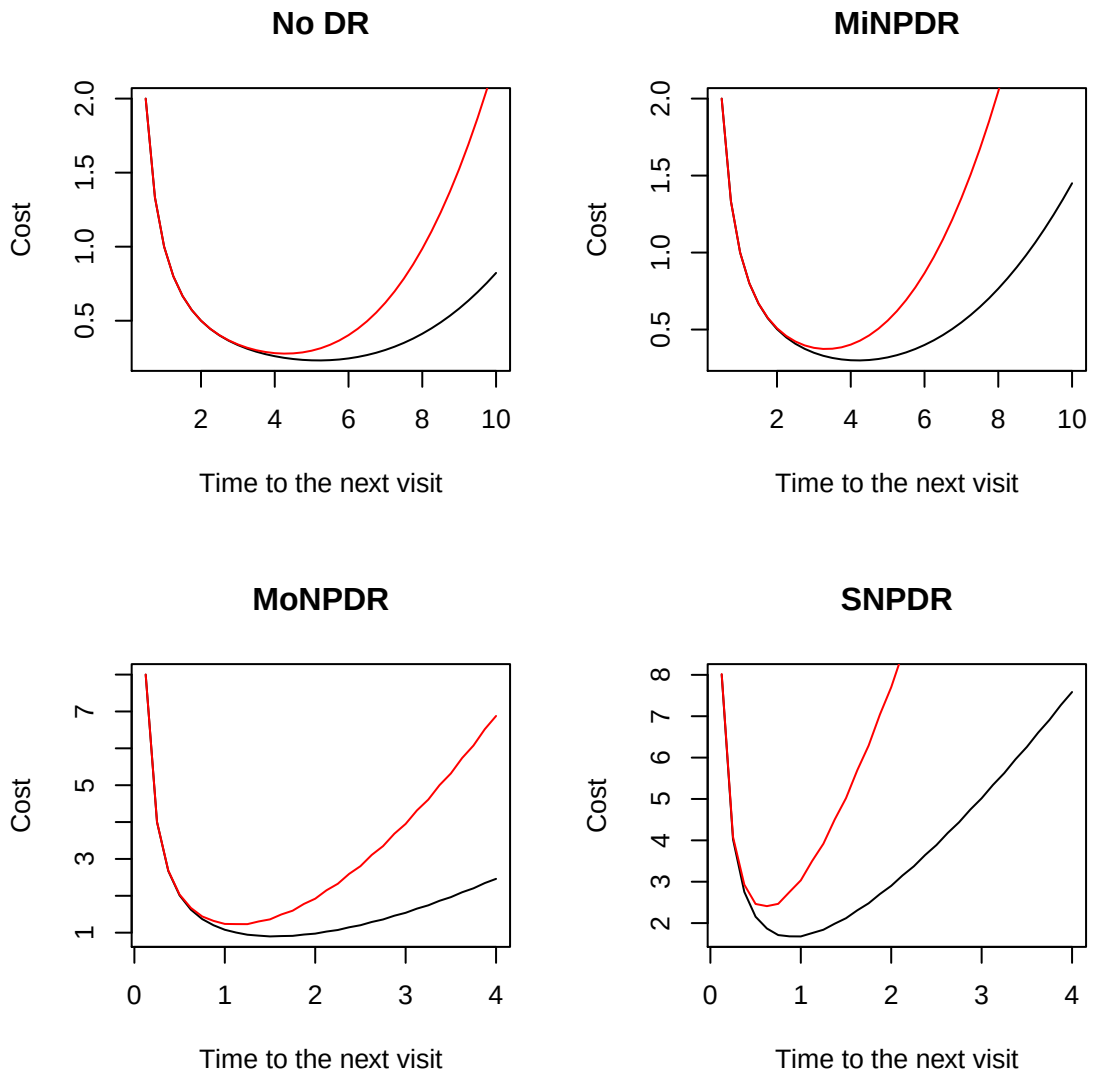
To illustrate the second approach, first notice (Figure 1) there is less than 5% chance of reaching PDR or CSME within 4 years when starting from no DR, and less than 5% chance within 3 years when starting from MiNPDR. One option is then to schedule the next visit in 4 years for a subject with no retinopathy, in 3 years for a subject with mild retinopathy, and (as per current guidelines) every year for those with MoNPDR or SNPDR, which leads to the screening schedule A =(4, 3, 1, 1). This schedule will be compared with annual screening.

The conditional probabilities γ's of reaching the PDR/CSME state under the (4, 3, 1, 1) screening are given by (0.056, 0.073, 0.238, 0.633), while for annual visits these are (0.002, 0.013, 0.270, 0.715).

The undetected time among those reaching the absorbing state under a (4, 3, 1, 1) screening schedule is 0.684 years, while for annual visits it is 0.606 years. Since most transitions to the absorbing state occur when subjects currently have MoNPDR and SNPDR, more frequent visits from those states will lead to lower T_U. For example, a (4, 3, 0.5, 0.25) schedule leads to an expected undetected time of 0.415 years.

Using (9), there is 67.43% chance the undetected time will exceed 0.5 years for a (4, 3, 1,

Figure 2: Total cost as a function of the time to the next visit for $(c_v, c_u) = (1, 10)$ (black line) and $(c_v, c_u) = (1, 30)$ (red line), stratified by the current retinopathy state. Note that the axes differ for each intermediate state to best display the change in costs as the screening interval changes.



1) schedule, 65.60% chance for annual screening, but only 18.15% chance for a (4, 3, 0.5, 0.25) schedule.

The average number of visits with a 20-year horizon for a $(4, 3, 1, 1)$ schedule is 6.73 visits, for $(1, 1, 1, 1)$ is 18.38, while a $(4, 3, 0.5, 0.25)$ schedule leads to an average of 7.65 visits.

Notice that the $(4, 3, 0.5, 0.25)$ schedule dominates the $(1, 1, 1, 1)$ schedule both in terms of effectiveness with an expected 0.19 years lower average undetected time, and costs with an expected 10.7 fewer number of visits over up to 20 years of follow-up.

Given an acceptable probability of progressing to PDR/CSME, the time to the next visit is determined based on the current retinopathy level such that the probability of reaching PDR/CSME is below (approximately) that threshold value. For illustration, using 5% as the acceptable risk, Table 1 presents the time to the next visit based on the current retinopathy state.

Table 1: Time to next visit (years) as a function of the current state such that the risk of reaching PDR/CSME is approximately 5%.

Current State	Time to Next Visit	Probability
No retinopathy	5.250	0.04912
MiNPDR	3.583	0.04847
MoNPDR	0.333	0.04493
SNPDR	0.083	0.05662

5.3 Covariates

The effect of covariates can also be incorporated, and this is illustrated using HbA1c (Table 2).

Table 2: Time to next visit (years) as a function of the current state such that the risk of reaching PDR/CSME is approximately 5% for various values of HbA1c, along with the average undetected time T_U (years) and the expected number of visits N_v.

HbA1c	No retinopathy	MiNPDR	MoNPDR	SNPDR	$E(T_U)$	$E(N_v)$
6	14.417	11.917	0.583	0.417	3.056	2.112
8	6.083	4.333	0.417	0.083	1.126	5.779
10	3.167	2.083	0.250	0.083	0.421	14.461

More frequent visits are required for larger values of HbA1c. With a 5% acceptable risk of reaching the PDR/CSME state, the next visit for a patient with no retinopathy at the current visit is scheduled in 14.4 years for an HbA1c of 6% and in 3.2 years for an HbA1c of 10%, while for a patient with SNPDR they are scheduled in 0.42 years for an HbA1c of 6% and in 0.08 years for an HbA1c of 10%.

Other covariates were also considered (gender, age, duration of diabetes, hypertension), but did not have a significant role in determining a screening schedule.

6 Discussion

A Markov model in continuous time is employed to describe transitions among health states over time, with particular interest in the incidence of an absorbing state which requires intervention. The goal is to investigate personalized cost-effective screening schedules. The cost associated with a given schedule has two components the number of visits and the length of time between the onset of the treatable condition and the visit when it is actually diagnosed (the undetected time). A more frequent screening schedule leads to more visits (more costly) and lower undetected time (less costly).

We first investigated fixed screening schedules (e.g., annual vs. biennial visits), and then considered tailoring the time to the next visit based on the current health state. The methods were illustrated using the progression of retinopathy in the DCCT/EDIC cohort, and the proposed approaches compared these personalized screening schedules with the current guidelines which recommend annual visits. It was shown that having more frequent visits for subjects at high risk and fewer visits for subjects at low risk leads to cost-effective screening schedules. In our example, a (4, 3, 0.5, 0.25) schedule for a subject starting in the no retinopathy state resulted in a 58% reduction in the number of visits compared to annual visits over a 20-year follow-up, while at the same time reducing the undetected time by 31%. The screening schedule can also take into account the effect of various risk factors, which is illustrated using HbA1c.

Besides the clinical benefit of early detection of progression, adopting the personalized screening schedule for retinopathy may lead to important savings. The US population of type 1 diabetes is approximately 1 million. Assume that 10% have already reached the PDR/CSME state and of the remaining 90%, approximately 25% have no retinopathy, 25% have MiNPDR, 30% MoNPDR, and 20% SNPDR. Assuming $200 per fundus photography, annual visits will require approximately $2.46B over 20 years, while the (4,3,0.5,0.25) schedule requires $1.39B, for a saving of $1.07B, or 43%.

In the DCCT/EDIC retinopathy example, the standard therapy for PDR is photocoagulation and for CSME is either photocoagulation or anti-VEGF, the costs of which are well established. There is no literature, however, on the costs associated with delayed detection and treatment (the untreated time, c_u), or the savings associated with early detection and treatment. Thus, these costs may be more or less fixed, at least over short periods of a few months that PDR/CSME might go undetected. However, for retinopathy and other conditions, there could be substantial costs associated with an increase in the period that progression goes undetected. Retinopathy progression doesn't stop when PDR/CSME occurs. Rather, if untreated subjects can progress to more severe levels, such as the development of what are termed "high risk characteristics" that place a subject at markedly higher risk of blindness. Thus, delayed detection could increase the costs associated with the initial treatment (e.g., photocoagulation) owing to a more severe condition upon detection, and could increase the costs associated with the treatment of a more rapid deterioration in the patient's condition. Thus, the above models could be further extended to incorporate temporal and other covariates into the cost function c_u.

The setup considered for determining the optimal screening schedule by minimizing the expected total cost (Section 4.1) is somewhat similar to a discrete-time Markov decision process (MDP) with deterministic history-dependent decision rules [12]. However, in the retinopathy application, the time horizon (L) refers to the time elapsed since the first visit, and therefore

the number of visits up to and including L is random. In contrast, an MDP with finite time-horizon typically refers to a fixed (finite) number of transitions, visits in our case. Another difference is that the ultimate goal is to obtain screening schedules that are meaningful for each patient, rather than cost-effective from a payer perspective, which is addressed by limiting the risk of undetected time (Section 4.2).

It should be noted that the calculation of undetected time only depends on the cumulative distribution of the time to the absorbing state. Therefore the results presented here apply to non-homogeneous Markov models as well, where the transition probabilities are obtained as solutions of the nonlinear differential equations corresponding to the Kolmogorov equations [13].

Other authors have considered the problem of optimal scheduling examinations [14, 11]. Three health states were considered (no disease, pre-clinical state and the clinical state), and the goal was to determine screening schedules that lead to early detection of disease. Unlike the retinopathy application considered herein where the status can either improve or worsen over time, the disease (e.g., cancer) was assumed progressive (i.e., no transitions from the pre-clinical state to the healthy state), and these methods are not directly applicable to our problem.

Clinical management of a chronic disease such as diabetes is a complex task requiring periodic evaluation of various potential complications, such as retinopathy, nephropathy and neuropathy. It is important to help provide informed guidelines on the screening for these complications which are tailored to the risk profile of the individual patient. The present work shows that personalized screenings have the potential to be cost-effective relative to fixed screenings.

7 Appendix

The number of visits with a time horizon L is obtained by discretizing the problem based on the smallest interval of time which can be considered between two consecutive visits. In the retinopathy progression example, these are monthly intervals, and the time scale is changed accordingly (e.g., L is in months). Since we are interested in the number of visits with an L-year follow up, we define an imbedded Markov chain which captures information on both the clinical state and the follow up time. Given the clinical states $u_1, \dots u_m$, where u_m is the absorbing state, the imbedded Markov model will have states $S^{\mathcal{I}} = \{u_{11}, \dots u_{1L}, \dots, u_{m-1,1}, \dots u_{m-1,L}, u_{m1}\}$, where a subject is in state u_{ij} if the clinical state is u_i $(i = 1, \dots m-1)$ at time j $(j = 1, \dots L)$, while u_{m1} is the absorbing state of the imbedded state defined as either reaching the absorbing clinical state u_m or the follow up time exceeding the time horizon L.

Transitions over time in the imbedded model occur with probabilities obtained based on the actions $a(u_{ij}) = a(u_i)$, since the actions only depend on the clinical state. Transitions to a non-absorbing state ($i\prime < m$) are given by

$$p^{\mathcal{I}}_{u_{ij},u_{i\prime j\prime}} = \begin{cases} p_{u_i,u_{i\prime}}(a(u_i)) & , \text{ if } j + a(i) \le L \ \&\ j\prime = j + a(i) \\ 0 & , \text{ otherwise}, \end{cases} \tag{10}$$

while the transition probability to the absorbing state is

$$p^{\mathcal{I}}_{u_{ij},u_{m1}} = 1 - \sum_{i\prime \neq m} \sum_{j\prime} p^{\mathcal{I}}_{u_{ij},u_{i\prime j\prime}} . \tag{11}$$

The expected number of visits before absorption in the initial Markov model (S, P) with a time horizon L is the time to absorption in the imbedded chain with state set $S^{\mathcal{I}}$ and transition probabilities matrix $P^{\mathcal{I}}$ with components defined in (10) and (11). It follows that the expected number of visits within the first L years is given by $N \cdot \mathbf{e}$, where $\mathbf{e}$ is a vector of ones and N is the fundamental matrix of the imbedded chain, with $N = (I - T)^{-1}$ and $T = P^{\mathcal{I}}_{1:(m-1)L,1:(m-1)L}$ [7].

Funding

This work was supported with funding from NIDDK, NIH for the study of the Epidemiology of Diabetes Interventions and Complications (EDIC) through grant U01-DK-094176, John M. Lachin, PI.

References

[1] Aiello L.P., Sun W., Das A., Gangaputra S., Kiss S., Klein R., Cleary P.A., Lachin J.M., Nathan D.M. and DCCT/EDIC Research Group. (2015). Intensive diabetes therapy and ocular surgery in type 1 diabetes. *New England Journal of Medicine*, **372**, 1722–1733.

[2] Cox D.R., Miller H.D. (1965). The Theory of Stochastic Processes, Methuen, London.

[3] Dasbach E.J., Fryback D.G., Newcomb P.A., Klein R., Klein, B.E.K. (1991). Cost-Effectiveness of Strategies for Detecting Diabetic Retinopathy, *Medical Care*, **29**, 20–39.

[4] The DCCT/EDIC Research Group (DCCT). (2015). Effect of intensive diabetes therapy on the progression of diabetic retinopathy in patients with type 1 diabetes: 18 years of follow-up in the DCCT/EDIC, *Diabetes*, **64**: 631–642.

[5] Early Treatment Diabetic Retinopathy Study Research Group (ETDRS). (1991). Early photocoagulation for diabetic macula retinopathy: ETDRS report number 9, *Ophthalmology*, **98**, 766–785.

[6] Fu, J.C., Lou, W.Y.W. (2003). Distribution Theory of Runs and Patterns and its Applications: A Finite Markov Chain Imbedding Approach, World Scientific, New Jersey.

[7] Iosifescu M. (1980). Finite Markov Processes and Their Applications, Wiley, New-York.

[8] Kalbfleisch J.D., Lawless J.F. (1985). The analysis of panel data under a Markov assumption, *Journal of the American Statistical Association* 1985, **80**, 863–871.

[9] Liu Y., Wang M., Morris A.D., Doney A.S., Leese G.P., Pearson E.R., Palmer C.N. (2013). Glycemic exposure and blood pressure influencing progression and remission of diabetic retinopathy: a longitudinal cohort study in GoDARTS, *Diabetes Care*, **36**, 3979–3984.

[10] Marshall G., Jones R.H. (1995). Multi-state Models and Diabetic Retinopathy, *Statistics in Medicine* 1995, **14**, 1975–1983.

[11] Parmigiani G. (1997). Timing medical examinations via intensity functions. *Biometrika*, **84**, 803–816.

[12] Puterman ML. (2005). Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley.

[13] Titman AC. (2011). Flexible nonhomogeneous markov models for panel observed data, *Biometrics*, **67**, 780–787.

[14] Zelen M. (1993). Optimal scheduling of examinations for the early detection of disease, *Biometrika*, **80**, 279–293.