MODELING THE NATURAL HISTORY OF YOUNGER-ONSET INSULIN-DEPENDENT DIABETIC RETINOPATHY

by

Bruce A. Craig   Dennis G. Fryback
Purdue University   University of Wisconsin

and

Ronald and Barbara E. K. Klein
University of Wisconsin

Technical Report #98–11

◇

MODELING THE NATURAL HISTORY OF YOUNGER-ONSET INSULIN-DEPENDENT DIABETIC RETINOPATHY

by

Bruce A. Craig   Dennis G. Fryback

Purdue University   University of Wisconsin

and

Ronald and Barbara E. K. Klein

University of Wisconsin

Abstract

The Wisconsin Epidemiologic Study of Diabetic Retinopathy is a population-based study to investigate prevalence, incidence, and progression of diabetic retinopathy. To investigate the natural history of the younger-onset insulin-dependent subpopulation, we propose and fit a nonhomogeneous discrete-time Markov model. We also account for treatment intervention and death. Bayesian inference is used to combine the data with the model. Predictive distributions are discussed as a prognostic tool to assist researchers in evaluating costs and benefits of treatment protocols. While our analysis focuses on diabetic retinopathy, we believe this methodology can have wide application.

1 Introduction

A serious complication associated with diabetes is a progressive disorder of the retina known as diabetic retinopathy. Late stages of this condition are characterized by a proliferation of abnormal new retinal blood vessels which may bleed into the vitreous gel of the eye leading to a decrease in vision. This condition is an important cause of severe vision loss (visual acuity of 5/200 or less) in working-age Americans and accounts for a large percentage of the new cases of legal blindness each year.

Beginning in 1959, the use of light therapy to burn the retina, thereby causing regression of these vessels and/or decreasing the leakage, was studied as a treatment for this condition. In 1976, the results of an extensive randomized controlled clinical trial revealed that the risk of severe visual loss can be reduced with panretinal photocoagulation, a random scatter laser application. Timely application of this treatment prevented severe vision loss associated with hemorrhaging of the new retinal vessels approximately 50% of the time[1,2]. However, without proper detection, the condition may progress to a point where this treatment is less efficacious. Because of this and the risks associated with the treatment, there is great interest in modeling the natural progression of diabetic retinopathy to investigate the cost-benefit of treatment under different screening and treatment programs.

As with many chronic diseases, a subject's "history" of diabetic retinopathy can be thought of as a progression through distinct severity states. The Markov process has been a convenient method to describe this series[3–6]. Both Garg et al.[4] and Dasbach et al.[5] used a homogeneous discrete-time Markov chain to describe the natural history from year to year. Often with these discrete-time models, estimation of these transition probabilities is difficult because the severity states are observed at several separated and unequally spaced time points. For example, the Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR) data[7], used in Dasbach et al.[5], now includes three observations per surviving participant with inter-observation periods of four and six years.

In this paper, we analyze the updated WESDR data set by proposing a flexible method to obtain estimates of annual transition probabilities when the observed data are at multiyear and unequal intervals. Unlike the other discrete-time Markov models,

we incorporate covariate information, thereby allowing estimation of a nonhomogeneous Markov process. In addition, we include descriptions of treatment intervention and death in our model to explain these occurrences during the study. This is an extension of the Craig and Newton paper[6] in which the model lumped together important screening level states thereby making treatment program comparisons impossible.

1.1 The Data

The WESDR is a population-based observational cohort study conducted in southern Wisconsin. The data were collected at three time points, four and six calendar years apart. For the remainder of the paper, we label these time points year 1 (baseline), year 5, and year 11. The natural history of diabetic retinopathy may vary for different types of diabetes. For this paper, we focus on the younger-onset insulin-dependent subpopulation.

For each subject at year 1, 5, and 11, a trained grader, masked to any information regarding the subject, rated the presence and severity of several characteristic lesions associated with diabetic retinopathy in seven standard 30° stereoscopic color fundus photographs using the modified Airlie House Classification scheme[8]. An algorithm[9] was used to determine retinopathy levels (RL) ranging from 10 (no disease) to 85 (end-stage proliferative retinopathy). Presence of panretinal photocoagulation scars and best corrected visual acuities (VA) in each eye were determined using standardized protocols.

The severity states for this analysis aggregate some of the retinopathy levels and also account for photocoagulation treatment status and visual acuity. These states, shown in Table 1, are mutually exclusive. All states, except state 5, are based on one visual acuity level, better or worse than 20/200 (legal blindness). State 5 is defined to represent sequelae of retinopathy resulting in either severe visual impairment or severe anatomic changes in the retina which can lead to severe visual impairment. The current analysis considers only information for the right eye.

4

Table 1. Severity States of Diabetic Retinopathy

State	Description
1. No retinopathy present	VA better than 20/200 and RL < 15
2. Nonproliferative retinopathy	VA better than 20/200 and $15 \leq$ RL ≤ 37
3. Mild nonproliferative retinopathy	VA better than 20/200 and $38 \leq$ RL ≤ 59
4. Active proliferative retinopathy	VA better than 20/200 and $60 \leq$ RL ≤ 84 with no signs of treatment intervention
T. Treatment intervention	VA better than 20/200 and $60 \leq$ RL ≤ 84 with signs of treatment intervention
5. Severe visual impairment	VA worse or equal to 20/200 or RL = 85

Complete information is not available on all subjects. Only subjects whose severity can be classified for at least one of the two observation times after baseline are included in our analysis. Table 2.1 and 2.2 summarize the data consisting of 970 subjects. In addition, several covariates were collected at each survey year including the subject's gender, age, duration of diabetes at baseline, and glycosylated hemoglobin level. Several of these will be incorporated into our analysis.

Our primary concern in the modeling of the natural history is from the standpoint of screening to identify a subject at the optimal stage for treatment with panretinal photocoagulation. As a result, this disease classification and subsequent analysis ignores several additional concerns outside this scope. First, we ignore macular edema (swelling of the retina in an area most critical for acute vision) which can occur in the presence of nonproliferative and proliferative stages of the disease prior to panretinal photocoagulation consideration and can also lead to severe vision loss. Also, subjects who have reached the severe visual impairment state can receive other treatments or in some cases regress naturally back to the active proliferative state. These subjects, however, would be under an ophthalmologist's care after the first occurrence of this state and thus for our purposes, their vision status will be considered to remain in this final state once it is entered.

The paper is organized as follows: Section 2 develops our stochastic model and Section 3 outlines our inference method. In Section 4, we discuss the results of

5

our Bayesian analysis and present some of the prognostic tools associated with this model. A discussion follows. In Appendix A we discuss the details of the posterior calculations.

2 The Markov Model

We propose that each subject's severity state from year to year is a Markov chain. By observing the state at three unevenly separated times, estimation of the year to year transition probabilities is numerically difficult. Although a subsampled Markov chain is still a Markov chain, the model parameters are often more naturally formulated on a yearly basis because this allows for a better understanding of the disease progression and possible treatment interventions at an interval commonly used for periodic physician visits. The decoupling of the observation intervals and the analytic intervals is a major feature of the proposed analysis process.

2.1 The Natural History

The natural history of the disease (i.e., the course of the disease without treatment) is described by transitions among the five severity states, labeled 1, 2, 3, 4, and 5 in Table 1. Assuming diabetic retinopathy is a semi-progressive disease, with regression possible between the first three states, the transition matrix from one year to the next is,

$$\begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} & \theta_{15} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} & \theta_{25} \\ \theta_{31} & \theta_{32} & \theta_{33} & \theta_{34} & \theta_{35} \\ 0 & 0 & 0 & \theta_{44} & \theta_{45} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where θ_{ij} represents the probability of moving from state i to state j during one year.

Unlike the earlier models of this disease, these transition probabilities may vary across time and subject. Using the proportional-odds model for each row in the transition matrix[10], we allow the subject's duration of diabetes, $\mathbf{d}$, to affect the chance of progression. For example, the first row of transition probabilities, $\{\theta_{1j}\}$, are determined from the following relationship:

$$\text{logit}\,(P(S \leq s)) = \eta_s + \alpha\mathbf{d} \qquad s = 1, 2, 3, 4 \quad,$$

6

where S is the subject's severity state, $\mathbf{d}$ is the subject's duration of diabetes at year's end, $\eta_1 \leq \eta_2 \leq \eta_3 \leq \eta_4$ are cut parameters, and α is a nonstationarity parameter. Similarly, for the other rows of the transition matrix, we restrict $\eta_5 \leq \eta_6 \leq \eta_7 \leq \eta_8$ and $\eta_9 \leq \eta_{10} \leq \eta_{11} \leq \eta_{12}$. The transition probabilities can also be viewed as areas under standard logistic curves as shown in Figure 1. If α is negative, all the cut points z slide to the left for positive duration, thereby increasing the chance of progression.
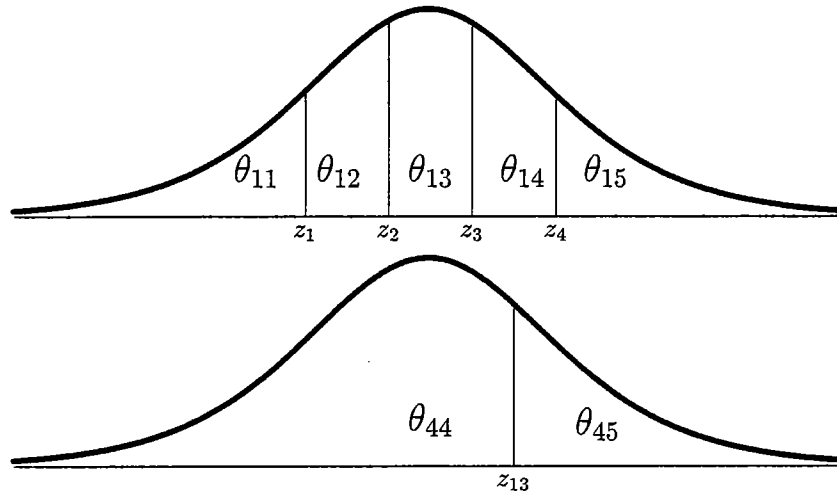

Figure 1: Relationship of transition probabilities with model parameters for the first and fourth row of the transition matrix. $z_i = \eta_i + \alpha\mathbf{d}$.

Other covariates, such as the glycosylated hemoglobin level, can be included in this proportional-odds model. We chose to use only duration because it is the only covariate known for every year in the study. The glycosylated hemoglobin level is known only at the observed years and thus additional modeling would be needed to include this or similar covariates.

2.2 Treatment Intervention and Death

Because WESDR is an observational study, treatment intervention and death occur for some subjects. As a result, the data do not directly represent the natural history of the condition. Rather than remove subjects from the analysis, or somehow adjust the outcome[5], we explicitly model the possibility of these occurrences. This allows us

to utilize as much subject information as possible in estimating the natural history.

2.2.1 Treatment Intervention

Limited information is available on subjects receiving photocoagulation treatment. For example, we do not know the year of treatment intervention, this being interval censored. Given the cost, the complications of the procedure, the period of the study, and the fact that state 4 was identified to be the state at which treatment is most likely to be effective[1,2], it is reasonable to suppose that photocoagulation is given only to those subjects in the active proliferative state. To disentangle treatment intervention from the natural history, we introduce a parameter ψ equal to the proportion of state 4 subjects who receive treatment. Once subjects receive treatment, they are considered to be in a new state T. Of these treated subjects, a fraction γ will respond to treatment. In ophthalmologist's terms, this means progression of retinopathy is stopped or there is a regression of the abnormal vessels without the development of severe vision impairment. For the fraction of cases that do not respond, we assume progression is equal in distribution to that of untreated cases.

A subject's transition matrix can be expanded to include this new state T. Since a treated eye will either respond successfully to treatment or not, there are two matrices defining the subject's progression. The only difference between these matrices is the fifth row, the transitions from the treated state. The first matrix is for a subject in whom the retinopathy responds favorably to treatment, and thus, once treated, will not progress to the vision impairment state:

$$
\begin{pmatrix}
\theta_{11} & \theta_{12} & \theta_{13} & (1-\psi)\theta_{14} & \psi\theta_{14} & \theta_{15} \\
\theta_{21} & \theta_{22} & \theta_{23} & (1-\psi)\theta_{24} & \psi\theta_{24} & \theta_{25} \\
\theta_{31} & \theta_{32} & \theta_{33} & (1-\psi)\theta_{34} & \psi\theta_{34} & \theta_{35} \\
0 & 0 & 0 & (1-\psi)\theta_{44} & \psi\theta_{44} & \theta_{45} \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\theta_{11} & \theta_{12} & \theta_{13} & (1-\psi)\theta_{14} & \psi\theta_{14} & \theta_{15} \\
\theta_{21} & \theta_{22} & \theta_{23} & (1-\psi)\theta_{24} & \psi\theta_{24} & \theta_{25} \\
\theta_{31} & \theta_{32} & \theta_{33} & (1-\psi)\theta_{34} & \psi\theta_{34} & \theta_{35} \\
0 & 0 & 0 & (1-\psi)\theta_{44} & \psi\theta_{44} & \theta_{45} \\
0 & 0 & 0 & 0 & \theta_{44} & \theta_{45} \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\tag{1}
$$

$$\text{w.p. } \gamma \qquad\qquad\qquad \text{w.p. } 1-\gamma$$

2.2.2 Death

Decreased survival in persons with proliferative retinopathy is associated with the presence of other life-threatening diseases, such as kidney and cardiac disease. This is

especially true when the subject reaches the vision impairment state[11]. We incorporate this fact into a hazard function for a diabetic individual. The estimated hazard function is based on the age/gender specific mortality rates for the general Wisconsin population (1992), the duration of diabetes, and whether or not the subject has reached the absorbing state.

Mortality rates for the general Wisconsin population are available at five year intervals. We assume the hazard function for the general population is constant within each five year interval. We incorporate the duration of diabetes as a covariate affecting the hazard. Other data have shown excess mortality to be low initially and then increasing in persons with younger-onset diabetes. For a subject in this study, it is assumed there is a constant addition to the general hazard until 15 years after diagnosis, after which time, the general hazard increases at a linear rate. Lastly, we include a constant addition to the hazard if the subject is currently in state 5. To summarize, for subject n at year t, the hazard is,

$$
\lambda_{n,t} = \begin{cases} \lambda(\,,\,) & t < t_d \\ \lambda(\,,\,) + \beta_1 + \beta_2 I_{S_{n,t-1}} & t_d \leq t < t_d + 15 \\ \lambda(\,,\,) + \beta_1(t - t_d - 14) + \beta_2 I_{S_{n,t-1}} & t \geq t_d + 15 \end{cases} , \qquad (2)
$$

where t_d is the year subject n was diagnosed with diabetes, $\lambda(\,,\,)$ is the general age/gender population hazard, and $I_{S_{n,t-1}} = 1$ if the subject has previously reached the severe vision impairment state.

By labeling the death state, D, and adding it to our transition matrix as an absorbing state, the transition matrix $M(n,t)$ for subject n from year $t-1$ to year t can be written as a sum of two matrices. The first matrix represents the natural progression of retinopathy combined with treatment intervention. The second represents the death process described above. If $p_{n,t}$ is the probability that subject n dies between year $t-1$ and year t, $M(n,t)$ is,

$$
(1 - p_{n,t}) \begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} & \delta_{14} & \delta_{1T} & \delta_{15} & 0 \\ \delta_{21} & \delta_{22} & \delta_{23} & \delta_{24} & \delta_{2T} & \delta_{25} & 0 \\ \delta_{31} & \delta_{32} & \delta_{33} & \delta_{34} & \delta_{3T} & \delta_{35} & 0 \\ 0 & 0 & 0 & \delta_{44} & \delta_{4T} & \delta_{45} & 0 \\ 0 & 0 & 0 & 0 & \delta_{TT} & \delta_{T5} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} + p_{n,t} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (3)
$$

where the δ_{ij} depend on the subject's response to treatment (equation 1).

3 Inference

We are interested in the joint posterior of the model parameters that define the natural history. Calculation of this distribution by analytic or numeric integration is very difficult due to the additional model parameters, which describe the occurrences of treatment intervention and death, and the numerous unobserved severity states.

3.1 Unobserved States

Even subjects with complete information have unobserved severity states (the years at which no survey is done). The distribution of these depends on the full set of model parameters, $\mathcal{P} = \{\eta_1, ..., \eta_{13}, \alpha, \beta_1, \beta_2, \psi, \gamma\}$, the covariate information, and the three observed states. While in some cases, due to our semi-progressive model, the unobserved states are determined by the observed states, often they can take on a multitude of values. For example, if the observed states at baseline and year 5 were 2 and 4 respectively, the following sequences of missing states are possible at years 2, 3, and 4:

(1, 1, 1)	(1, 1, 2)	(1, 1, 3)	(1, 1, 4)	(1, 2, 1)	(1, 2, 2)	(1, 2, 3)	(1, 2, 4)
(1, 3, 1)	(1, 3, 2)	(1, 3, 3)	(1, 3, 4)	(2, 1, 1)	(2, 1, 2)	(2, 1, 3)	(2, 1, 4)
(2, 2, 1)	(2, 2, 2)	(2, 2, 3)	(2, 2, 4)	(2, 3, 1)	(2, 3, 2)	(2, 3, 3)	(2, 3, 4)
(3, 1, 1)	(3, 1, 2)	(3, 1, 3)	(3, 1, 4)	(3, 2, 1)	(3, 2, 2)	(3, 2, 3)	(3, 2, 4)
(3, 3, 1)	(3, 3, 2)	(3, 3, 3)	(3, 3, 4)	(1, 4, 4)	(2, 4, 4)	(3, 4, 4)	(4, 4, 4)

When determining these series, besides restricting a nontreated subject from entering the treated state, we also force all treated subjects to enter the treated state the year they receive treatment. Without this condition, a treated subject could jump from severity state 1, 2, or 3 to the severe vision impairment state in one year. We would not be able to completely untangle the natural history from treatment intervention since $\theta_{i,5}$ would depend on both treated and untreated subjects.

3.2 The Prior

The prior distribution represents our information about the set of model parameters before analyzing the data. For simplicity, we assume prior independence of all model parameters, except those η's that must be ordered to determine a row of transition probabilities.

First, consider the natural history transition matrix. This is determined by $\eta_1, \eta_2, ..., \eta_{13}$ and the nonstationarity parameter α. Within a row, the joint prior has these parameters equal to the order statistics from a standard logistic distribution, thereby producing a uniform distribution over transition matrices when $\alpha = 0$. For the parameter α, since we do not assume to know whether duration increases or decreases the chance of progression, we use a normal prior centered at zero. Because α is multiplied by the number of years the subject has had diabetes, it does not have to be far from zero to have a large effect on the transition probabilities. We set the variance to be .0625.

We use a uniform(0,1) prior for both the treatment intervention parameter ψ and response parameter γ. Lastly, we assume that a subject with diabetes has a lower life expectancy than a similar non-diabetic subject. Thus we restrict β_1 to be positive. In addition, since we assume a subject having reached the severe vision impairment state is more likely to die, we restrict β_2 to be positive. We use gamma priors centered at .005 with a large variance of .25 and centered at 1 with a variance of 1 respectively.

3.3   Computing the Posterior

There are eighteen model parameters, the response to treatment is unknown for approximately 200 subjects, and more than 6000 unobserved severity states are created by year to year modeling. We use Markov chain Monte Carlo[12,13] to examine the joint posterior of the model parameters. We run a Markov chain over the full set of model parameters, $\mathcal{P}$ and unobserved variables, $\mathcal{U}$.

$$(\mathcal{U}^1, \mathcal{P}^1), (\mathcal{U}^2, \mathcal{P}^2), (\mathcal{U}^3, \mathcal{P}^3), ...$$

Each step or "cycle" of the chain is the result of smaller steps, each being a Metropolis-Hastings update which modifies a component of the larger state. One step modifies the unobserved severity states. Another modifies the unknown treatment responses. The remaining steps update the model parameters. Appendix A details these steps. By its construction, this Markov chain is irreducible with the equilibrium distribution equal to the posterior

$$\pi(\mathcal{U}, \mathcal{P}|\{\text{data}\}, \{\text{covariates}\}) \propto \pi(\mathcal{U}, \{\text{data}\}|\mathcal{P}, \{\text{covariates}\})\pi(\mathcal{P}) \qquad (4)$$

where the first factor on the right comes from the Markov model and is a product of transition probabilities (equation 3) based on the series of severity states. The second factor on the right is the prior.

A chain of 5,000 cycles was run on an IBM RISC/6000 computer, taking approximately 48 CPU hours. Estimates of the marginal posterior summaries were based on subsampling every other cycle after an initial burn in of 1000; these dimensions were chosen by informal assessment of time-series plots under various starting positions. Marginal posterior densities are approximated using normal kernel densities applied to the simulated parameter values.

4 Results

Because we incorporate specific covariate information into the natural history transition matrix, there are only a few model parameters that are meaningful marginally. We will discuss these and then present several posterior distributions that are functions of the parameters. Lastly, we will discuss posterior predictive distributions that are useful for assessing the fit of the model and making prognostic inferences.

4.1 Marginal Posterior Distributions

Two interesting parameters to examine marginally are γ and α. While the exact value of α is difficult to interpret, its sign tells us whether duration of diabetes increases or decreases the chance of retinopathy progression. The left panel of Figure 2 shows the prior and resultant marginal posterior of α. The posterior mean and median are negative, implying that duration of diabetes increases the chance of progression. There is not, however, very strong evidence that this parameter is negative as the posterior probability of α being negative is approximately 90%.

The parameter γ is the fraction of subjects responding to treatment. Controlled clinical trials have shown this to be around 50%. The right panel of Figure 2 shows the marginal prior and posterior for this parameter. The data provide little information about this parameter, except that it is likely to be below 50%. The probability of success in the WESDR observational study is between 2.1% and 55.0% with 95% posterior probability.

The interpretation of the remaining model parameters is dependent on specific
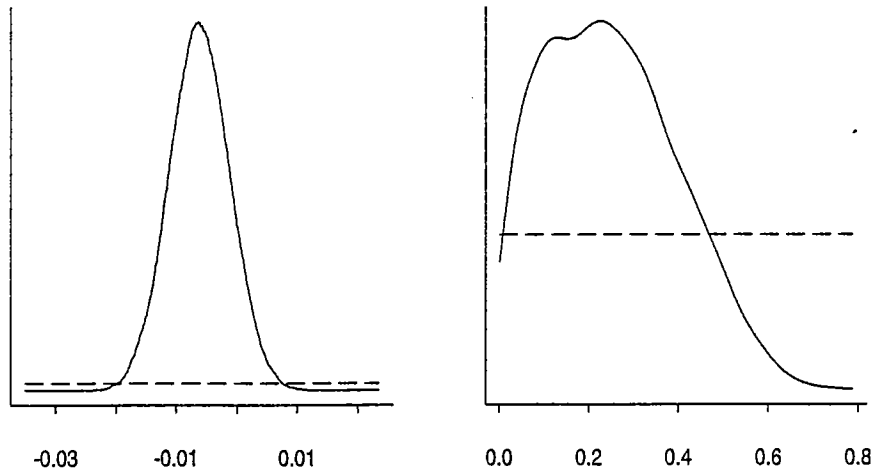
12

Figure 2: Two parameters of interest. The left panel shows the marginal prior (dashed) and posterior distribution (solid) for α. The parameter α adjusts the transition matrix based on the number of years the subject has had diabetes. The right panel shows the prior and posterior distribution for γ, the probability of successful treatment.

covariate information. Two subject-specific functions of the model parameters that are of interest are the probability that a subject is in a certain severity state after t years, and the remaining years before vision impairment or death. To calculate the remaining expected number of years before vision impairment or death, we need to know the probability that the subject dies or enters state 5 after t years. These probabilities can be obtained directly from the t_o to year $t_o + t$ transition matrix $M(n, t_o + t)M(n, t_o + t - 1) \cdots M(n, t_o + 1)$, where t_o is the current year.

For example, suppose we have a 30 year old male subject who is in state 2 and was diagnosed with insulin dependent diabetes 8.3 years earlier. Ignoring the possibility of treatment, there is 95% posterior probability that the subject's expected number of years before vision impairment or death is between 25.6 and 29.5 years (Figure 3, left panel). This is about 19.2 years less than the remaining life expectancy of a similar person without diabetes.

A subject may be more interested in the probability of vision impairment in the following year if not treated. This probability can be found directly from the transition matrix $M(n, t_o + 1)$. The right panel of Figure 3 shows the posterior distribution of

13

this probability for a 35 year old male diagnosed with insulin dependent diabetes 12.0 years earlier who is currently in the active proliferative state. The probability is between 2.6% and 6.0% with 95% posterior probability.
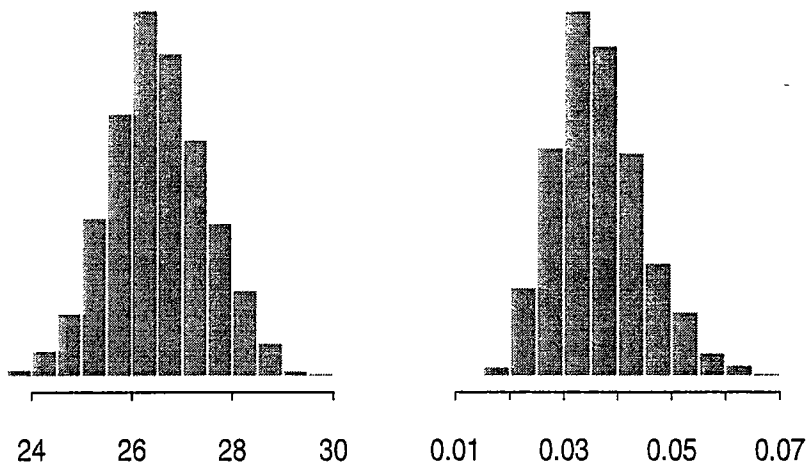

Figure 3: Posterior distribution for the expected number years before vision impairment or death without treatment intervention (left panel) and the posterior probability that a subject in the active proliferative state would enter the vision loss state.

We can extend this to compute the probability that a subject starting in state i will be in state j after t years. If we do this for each subject in a cohort, we can compute the expected number of subjects in each state after t years. For example, starting with an identical cohort (same initial covariates and severity states as WESDR), we calculated the expected number of subjects in each severity state 4 and 6 years later, the observed years of WESDR. The posterior mean and standard deviation for each of these expected counts are shown in Table 2.1 and 2.2.

4.2 Predictive Distributions

Posterior predictive distributions describe other data that may reasonably be expected to have been observed, and can be used to assess the fit of the model. Again starting with an identical cohort, for each saved cycle from the Markov chain, we generate, for each subject, a new sample path up to year 5 and a sample path from year 5 to year 11. Combining the information for year 5 and 11, we obtain an expected count

For each cell in table:

Observed	Expected
	S.D. (Mean)
	S.D. (Pred)

Table 2.1 First to Fifth Year Transitions.

Init. State	Future State						
	1	2	3	4	T	5	D
1	144 142.0	148 137.8	8 18.7	3 2.5	0 1.2	1 1.3	3 3.5
	6.9	6.0	1.8	0.5	0.4	0.8	0.4
	10.9	10.5	4.5	1.7	1.2	1.4	1.9
2	14 13.5	224 229.6	60 67.4	17 14.6	15 8.2	4 1.7	14 13.1
	2.4	5.4	4.2	1.4	1.1	0.6	1.5
	4.3	10.3	8.4	4.0	3.0	1.4	3.8
3	0 0.8	20 13.2	28 38.3	18 20.7	18 19.3	6 3.0	10 4.7
	0.4	2.2	2.6	1.7	1.7	0.6	0.5
	1.0	4.0	5.6	4.5	4.3	1.8	2.2
4				29 17.7	21 31.8	3 7.4	9 5.1
				1.7	1.8	1.3	0.5
				4.0	4.3	2.8	2.2
T					54 58.5	8 6.8	9 5.7
					1.6	1.3	0.6
					3.5	2.8	2.3
5						44 39.3	19 23.7
						2.6	2.6
						4.6	4.6

Table 2.2 Fifth to Eleventh Year Transitions.

Init. State	Future State						
	1	2	3	4	T	5	D
1	47 45.5	81 74.1	15 15.7	0 3.2	0 2.1	0 1.0	2 3.4
	3.2	2.8	1.2	0.4	0.4	0.5	0.4
	6.5	6.6	4.0	1.8	1.5	1.1	1.8
2	17 15.9	200 206.8	75 78.5	36 24.8	23 21.1	1 4.3	22 22.6
	2.7	6.8	4.8	2.3	2.3	1.0	2.8
	4.7	11.7	9.3	5.3	5.0	2.3	5.8
3	0 0.9	12 13.5	29 23.8	23 17.3	23 26.8	0 5.1	7 6.6
	0.4	2.1	2.2	1.6	2.1	0.9	0.7
	1.0	4.0	4.8	4.0	4.9	2.3	2.5
4				19 9.3	29 34.3	2 8.9	11 8.7
				1.3	1.8	1.4	0.8
				3.2	4.2	3.0	2.8
T					85 76.3	8 12.6	12 16.1
					3.2	2.4	1.5
					5.5	4.1	3.9
5						32 30.9	31 32.1
						3.1	3.1
						5.0	5.0

and predictive standard deviation for each cell in Table 2.1 and 2.2. The mean of each cell will be the same as that of the posterior distribution of expected counts. The standard deviation, however, will be larger due to the increased uncertainty in the generated paths.

Both the posterior predictive distribution and the posterior distribution incorporate the current treatment intervention process as the simulated decision process. In practice, the treatment process is much more involved and that is reflected in our tables. While is appears from these tables that this process explains the occurrences adequately, the poorest fit cells involve the treatment process. This will be discussed in more detail in the following section.

5 Discussion

Our primary objective is to describe a method to estimate the natural history of a chronic disease when the observed states are at multiple separated time points and the desired model is on a less coarse scale. While we feel we have adequately described the natural history, we realize that this model may be overly simplified. Cross-validation using the left eye and other model fitting techniques are currently being pursued. Unlike many other estimation techniques, our method can easily accommodate changes in the discrete-time Markov model. For example, one could incorporate the proportional-hazards model (extreme value distribution) or any other distribution to determine the transition probabilities. To disentangle treatment from the natural history, in the current model we assume that unsuccessfully treated subjects progress as if they were never treated. If a researcher were uncomfortable with this assumption, the progression could be modeled differently. Inference would still be possible but the number of model parameters may increase. Different models for the hazard and treatment intervention as well as the inclusion of additional covariates can also be investigated.

Our method has allowed us to estimate the natural history from an observational study with treatment intervention and death by incorporating simple analytic descriptions of the processes. This allows us to include all the subjects in the analysis. Since the treatment process and death were built around the natural history model, they can easily be stripped away leaving the natural history. This is important in the

16

comparison of treatment and screening protocols.

The main prognostic tool of this analysis is the predictive distribution. If a researcher wants to investigate different screening or treatment protocols, each modeled program could be combined with the natural history model to generate a posterior predictive distribution (for a specific cohort) of expected counts. For each realization from this distribution, a summary statistic, such as a cost-effectiveness ratio, can be calculated to obtain a summary distribution of each program. With the increasing demand for efficient health care, these summary distributions can play an important role in program comparisons. Not only can we get an estimate for each program, but we are also able to assess the variability of summary statistics. Lastly, since we sample from the posterior distribution of parameter values, our estimates are no longer conditional on particular parameter values.

Due to the coarseness of sampling times, the description of treatment intervention is very limited. But apparently it does an adequate job of explaining the current data. Why is our inferred treatment success probability less than that observed in clinical trials? The WESDR study was initiated in 1979, three years after the publication of the controlled clinical trial which showed the efficacy of treatment. Some subjects in WESDR may have been treated in suboptimal stages of retinopathy because the disease was not detected in a timely fashion or in some cases because the physician was unaware of the dependence on severity state for response to panretinal photocoagulation. It may be that the probability of a successful response is increasing over the duration of WESDR, however we have chosen to model treatment success as a constant. We also describe the natural history of diabetic retinopathy from a screening standpoint for panretinal photocoagulation. An important aspect of the disease that we did not consider is macular edema, another cause of visual loss. It is likely that some of the subjects in this study were affected by this condition, reducing the apparent success of treatment with respect to visual impairment.

The Bayesian inference approach described here to model a nonhomogeneous discrete-time Markov chain appears to be a useful analytic tool. The results agree in large part with generally accepted views of the natural history of diabetic retinopathy, and the posterior predictive distributions derived here are useful for model building. While our analysis was limited to younger-onset insulin dependent diabetes, we believe this method can have wide application.

Appendix

Metropolis-Hastings Steps

We run a Markov chain over the set of model parameters and unobserved variables with the equilibrium distribution equal to the posterior distribution π of this set (equation 4). One complete step in the Markov chain is produced by a sequence of smaller Metropolis-Hastings (MH) steps which modify different aspects of this set. Each MH algorithm is defined by a proposal distribution, having density $q(s, s')$, which indicates the probability density of sampling the new state s' given the current state s. The MH ratio,

$$r = \frac{\pi(s')q(s', s)}{\pi(s)q(s, s')}$$

is calculated and with probability $\min(r, 1)$ the Markov chain moves to s', otherwise it stays put. The following sections describe the proposal distribution for each of the MH steps and the form of the MH ratio.

Proposing new health states

This is done on a subject-by-subject basis. For each subject, a new series of states is proposed differing by only one state. This results in a simple and computer efficient algorithm. We cycle through each subject ten times.

1. One of the subject's unobserved years, t, is chosen randomly using a discrete uniform.

2. A potential state, j', is chosen using a discrete uniform of the states between i, the state at year $t-1$ and k, the state at year $t+1$. Adjustments to the discrete uniform are due to:

   - State i and/or k is less than or equal to 3 (regression possible).
   - A non-treated subject cannot move to the treated state.
   - A treated subject must enter the treated (T) state.
   - For those subjects with unknown treatment information, treatment or no treatment is possible.

- A successfully treated subject cannot enter state 5.

3. Reject or accept the proposed state based on the MH ratio.

Because the unobserved year and potential state are chosen using discrete uniforms, the proposal density is symmetric, $q(s, s') = q(s', s)$. The remainder of the MH ratio greatly simplifies, because a single severity state is changed. For example,

$$
\begin{aligned}
\frac{\pi(s')}{\pi(s)} &= \delta^t_{ij'}\delta^{t+1}_{j'k}/\delta^t_{ij}\delta^{t+1}_{jk} \quad j,j' < 5, k \le 5 \\
&= \delta^t_{ij'}/\delta^t_{ij} \qquad\qquad j,j' < 5, k = D
\end{aligned}
$$

where δ^t_{ij} represents the probability of moving from state i to state j during year t when ignoring death (equation 3). The death process only enters this ratio when the current or potential state is death or the vision impairment state.

Changing the Subjects' Response to Treatment

For each treated subject, response is unknown unless the subject enters the vision impairment state. We update each subject's response using a deterministic flip. For each subject with unknown response and a series that does not enter state 5, the response opposite to the current response is proposed. The proposed state is accepted with probability

$$
r = \prod_{t=2}^{11} \frac{\delta^t_{ij} \mid \text{proposed response}}{\delta^t_{ij} \mid \text{current response}}.
$$

Changing the Probability of Treatment and Successful Response

Because ψ and γ must be between 0 and 1, a beta proposal distribution centered at the current value is used. For these updates (in this example ψ),

$$
\frac{q(s', s)}{q(s, s')} = \left(\frac{\Gamma(\alpha' + V)\Gamma(\alpha)\psi^{\alpha'-1}}{\Gamma(\alpha + V)\Gamma(\alpha')\psi'^{\alpha-1}} \right) \left(\frac{1 - \psi}{1 - \psi'} \right)^{V-1}
$$

$$
\alpha = V\left(\frac{\psi}{1 - \psi} \right) \qquad \alpha' = V\left(\frac{\psi'}{1 - \psi'} \right)
$$

where V controls the variance. We set $V=500$ for ψ and $V = 300$ for γ to maintain acceptance rates around 70%. The remainder of the MH ratio is a product of transition probabilities from the Markov model and the marginal prior.

Changing β_1 and β_2 in the Hazard Function

Because each β is assumed positive, we use a gamma proposal distribution with mean being the current value. Thus,

$$\frac{q(s', s)}{q(s, s')} = \left(\frac{\beta}{\beta'}\right)^{2V-1} e^{V(\frac{\beta'}{\beta} - \frac{\beta}{\beta'})}$$

where V controls the variance and is set at 75 and 50, for β_1 and β_2 respectively. Since either parameter only changes the hazard, the remainder of the ratio is a product of the $p_{n,t}$'s and $(1 - p_{n,t})$'s (equation 2) and the marginal gamma prior.

Changing the η's in the transition matrix

We update each η individually using a folded normal centered at the current value. We must fold because of the cutpoint order restrictions. The folded normal is a symmetric proposal distribution so the MH ratio is simply the ratio of the likelihood and prior. The likelihood reduces down to only those transitions for which the probability depends on the cutpoint parameter being updated and the death process can be ignored.

Acknowledgements

References

1 Diabetic Retinopathy Study Group. 'Report No. 6, Design, methods, and baseline results', *Investigative Ophthalmology and Visual Science* 21, 149–209 (1981).

2 Diabetic Retinopathy Study Group. 'Photocoagulation treatment of proliferative diabetic retinopathy: Clinical application of diabetic retinopathy study (DRS) findings. (DRS report No. 8)', *Ophthalmology* 88, 583–600 (1981).

3 Beck, J.R. and Pauker, S.G. 'The Markov process in medical prognosis', *Medical Decision Making* 3, 419–458 (1983).

4 Garg, S.K., Marshall G., Chase, H.P., Jackson, W.E., Archer, P., and Crews, M. 'The use of the Markov process in describing the natural course of diabetic retinopathy', *Archives of Ophthalmology* 108, 1245–1247 (1990).

5 Dasbach, E.J., Fryback, D.G., Newcomb, P.A., Klein, R. and Klein, B.E.K. 'Cost-effectiveness of strategies for detecting diabetic retinopathy', *Medical Care* 29(1), 20–39 (1991).

6 Craig B.A. and Newton, M.A. 'Modeling the history of diabetic retinopathy' *in Case Studies in Bayesian Statistics III*, C. Gatsonis et al. eds, New York: Springer-Verlag, 305–323 (1997).

7 Klein, R., Klein, B.E.K., Moss, S.E., Davis, M.D., and Demets, D.L. 'The Wisconsin epidemiologic study of diabetic retinopathy. II. Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 Years', *Archives of Ophthalmology* 102, 520–526 (1984).

8 Diabetic Retinopathy Study Group. 'Report No. 7, A modification of the Airlie house classification of diabetic retinopathy', *Investigative Ophthalmology and Visual Science* 21, 210–226 (1981).

9 Klein, R., Klein, B.E.K., Moss, S.E., and Cruickshanks, K.J. 'The Wisconsin epidemiologic study of diabetic retinopathy. XIV. Ten-year incidence and progression of diabetic retinopathy', *Archives of Ophthalmology* 112(9), 1217–1228 (1994).

10 McCullagh, P. and Nelder, J.A. *Generalized Linear Models*, Chapman and Hall, London, 149–155 (1983).

11 Klein, R., Moss, S.E., Klein, B.E.K., DeMets, D.L. 'Relation of ocular and systemic factors to survival in diabetes', *Archives of Internal Medicine* 149, 266–272 (1989).

12 Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. 'Equation of state calculations by fast computing machines', *Journal of Chemical Physics* 21, 1087–1092 (1953).

13 Smith, A.F.M. and Roberts, G.O. 'Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with Discussion)', *Journal of the Royal Statistical Society, Series B* 55, 3–24 (1993).