A prediction model for colon cancer surveillance data

Norm M. Good,[a] Krithika Suresh,[b] Graeme P. Young,[c] Trevor J. Lockett,[d] Finlay A. Macrae[e] and Jeremy M. G. Taylor[b*†]

Dynamic prediction models make use of patient-specific longitudinal data to update individualized survival probability predictions based on current and past information. Colonoscopy (COL) and fecal occult blood test (FOBT) results were collected from two Australian surveillance studies on individuals characterized as high-risk based on a personal or family history of colorectal cancer. Motivated by a Poisson process, this paper proposes a generalized nonlinear model with a complementary log–log link as a dynamic prediction tool that produces individualized probabilities for the risk of developing advanced adenoma or colorectal cancer (AAC). This model allows predicted risk to depend on a patient's baseline characteristics and time-dependent covariates. Information on the dates and results of COLs and FOBTs were incorporated using time-dependent covariates that contributed to patient risk of AAC for a specified period following the test result. These covariates serve to update a person's risk as additional COL, and FOBT test information becomes available. Model selection was conducted systematically through the comparison of Akaike information criterion. Goodness-of-fit was assessed with the use of calibration plots to compare the predicted probability of event occurrence with the proportion of events observed. Abnormal COL results were found to significantly increase risk of AAC for 1 year following the test. Positive FOBTs were found to significantly increase the risk of AAC for 3 months following the result. The covariates that incorporated the updated test results were of greater significance and had a larger effect on risk than the baseline variables. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: colonoscopy; cancer surveillance; interval censored; adenoma; complementary log-log link; Poisson process

There is a well-established model for colorectal cancer (CRC) development [1]. CRC begins as benign polyps that grow over time in the lining of the colon. Individuals are more susceptible to developing polyps if they are aged 50 years or older, have a personal or family history of polyps, and/or have inherited gene mutations known to be associated with colon polyps. Most of these polyps, characterized as hyperplastic polyps, will remain benign for the lifetime of the individual, with a very low chance of becoming cancer. Some polyps, known as adenomas, may become cancerous if not removed. The risk of an adenoma developing into cancer increases with adenoma size and the period it has been growing in the colon. A person is said to have advanced adenoma if an adenoma is large ($\geqslant 10$ mm), if an adenoma has certain characteristics (tubulovillous, villous histology, or high-grade dysplasia), or if the person has three or more adenomas. Advanced adenoma is a surrogate for present and future CRC risk [2]. Malignant adenomas that are not removed in the early stages can develop into invasive CRC. As the cancer grows, it may infiltrate neighboring structures, which if undetected and untreated would lead to the individual experiencing symptoms. CRC is one of the most common cancers, and treatments are only moderately effective, particularly for more advanced disease. CRC survival rates are improved with early diagnosis; thus, there is a great interest in methods for early detection and prevention. There are a number of

[a]*CSIRO Mathematical and Information Sciences/Australian e-Health Research Centre, Royal Brisbane and Women's Hospital, Herston, QLD, 4029, Australia*
[b]*Department of Biostatistics, University of Michigan, Ann Arbor, MI, U.S.A.*
[c]*Flinders Centre for Innovation in Cancer, Flinders University, Bedford Park, SA, 5042, Australia*
[d]*CSIRO Preventative Health Flagship and Animal, Food and Health Sciences, Riverside Corporate Park, North Ryde, NSW, 2113, Australia*
[e]*Colorectal Medicine and Genetics, The Royal Melbourne Hospital, VIC, 3050, Australia*
*Correspondence to: Jeremy M. G. Taylor, Department of Biostatistics, University of Michigan, 1415 Washington Heights, Ann Arbor, MI, 48109, U.S.A.
[†]*E-mail: jmgt@umich.edu*

tests for detecting early signs of CRC, such as a colonoscopy (COL), a fecal occult blood test (FOBT), a sigmoidoscopy, many variations of these tests, and new biomarker-based tests. These tests have properties, risks, and expenses associated with them, so an important medical and public health problem is the judicious use of these tests.

A colonoscopy is a procedure by which adenomas and cancers are visually looked for by inserting a tube into the colon. As part of the procedure, any adenomas that can be found are resected, removed, and subsequently examined to see if they contain cancer. Thus, a colonoscopy is both a detection and a prevention method because it removes adenomas that are or may become malignant. A FOBT is a test of whether blood is detected in the stool. It is not specific to adenomas or cancers, but for mechanical reasons, larger adenomas and cancers are more likely to be bleeding, and hence more likely to give rise to a positive FOBT. In practice, a positive FOBT is usually quickly followed up by a COL. When these tests are applied in the general low-risk population, it is called screening. When the tests are applied in a population who are considered high risk for CRC, it is called surveillance. There are a number of published guidelines for the use of these tests in the screening and surveillance setting [3–8].

In this paper, we will be developing a dynamic individual prediction model from a data set of high-risk individuals who are in a surveillance program. The increasing availability of longitudinal patient information enables individualized dynamic predictions [9, 10]. For a given individual, dynamic prediction models predict the probability of an event occurring in an interval by conditioning on the individual's longitudinal patient history of tests. In addition to varying between individuals, the risk of developing CRC changes over the trajectory of an individual's life. The benefit of these models is the ability to update patient risk at each colonoscopy and FOBT while still incorporating the effect of previous test results.

This paper presents an approach to analyzing data from studies of serial colonoscopies in individuals who are at high risk for colon cancer. The primary goal is to provide a method that gives individual predictions based on a person's characteristics and the results from their sequence of FOBT and COL tests. Specifically, the aim is to estimate the distribution of times to the development of the disease for a person at the time of a colonoscopy or at any specified age. Because cancers are relatively rare, we characterize the disease, or event of interest, as the development of an advanced adenoma or cancer. The prediction, which will be in the form of a distribution function, may be a useful ingredient for individual recommendations or for an overall surveillance policy. For example, someone who is estimated to have a very low risk of developing a new advanced adenoma or cancer in the next 5 years would likely not be scheduled for a COL in the immediate future, compared with someone with a high risk who should be scheduled for a COL much sooner. The prediction model we present will also use information on simpler covariates, such as gender and family history of CRC, as they are also likely to influence the estimates of someone's risk.

While the data are rich in information, there are complicating aspects associated with it that require careful and potentially non-standard statistical approaches. These data were collected over a number of years, and the guidelines for when COLs and FOBTs should be scheduled have changed during that period. The number of colonoscopies per patient is quite variable, and the time interval between colonoscopies is heterogeneous both within and between people. The development of advanced adenoma or cancer is something that happens in continuous time, but observations about that process are only made at the time of COLs, leading to interval-censored data. A person can have more than one observed occurrence of advanced adenoma or cancer during their follow-up. The COL and FOBT results can be regarded as time-dependent covariates that provide information about an individual's risk. There are more subtle issues too, such as the frequency of colonoscopies, which itself is a measure of risk and is impacted by prior results. We will not attempt to address all the issues in this paper, but rather limit ourselves to providing a statistical framework that can be used for prediction purposes.

The statistical approach we develop regards the time of development of the event as arising in continuous time from a Poisson process, as has been used by others in colon cancer research [11]. Treating the event times as interval-censored, we formulate a likelihood function and proceed with maximum likelihood estimation. As a consequence of the Poisson process assumption, we show that our binary outcome data can be modeled using a generalized nonlinear model with a complementary log–log link, which facilitates estimation using standard software.

The paper is organized as follows. We describe the data in Section 1, and the model and the estimation method in Section 2. The results including parameter estimates and goodness-of-fit are presented in Section 3. An example of individualized predictions of risk is demonstrated in Section 4. We finish with a discussion of model limitations and alternative strategies for future model adaptations.

1. Colorectal cancer surveillance data

1.1. Data

The data come from two surveillance programs in Adelaide (7670 participants) [12–14] and Melbourne (2829 participants) [15, 16], and are also described in other literature [17]. The data were collected over the period 1 January 1976 to 3 August 2010. The total number of colonoscopies recorded during this period is 20,056. Patients with only one recorded colonoscopy, of which there are 4550 (3818 from

Table I. Descriptive statistics for the study cohort used in analysis.

Variable	Adelaide	Melbourne	Combined
Patients	3499	1059	4558
Average Age at index (SD)	63.7 (11.4)	54.0 (8.8)	61.4 (11.6)
Female (%)	1612 (46.1)	683 (64.5)	2295 (50.4)
Family History (%)†			
R0	2430 (69.4)	46 (4.3)	2476 (54.3)
R1	150 (4.3)	108 (10.2)	258 (5.7)
R2	224 (6.4)	173 (16.3)	397 (8.7)
R3	234 (6.7)	107 (10.1)	341 (7.5)
R4	407 (11.6)	273 (25.8)	680 (14.9)
CCC	0 (0)	83 (7.8)	83 (1.8)
HNPCC-suspected	23 (0.7)	126 (11.9)	149 (3.3)
HNPCC-definite	31 (0.9)	143 (13.5)	174 (3.8)
Colonoscopies	9,704	3,723	13,427
Colonoscopies per patient (%)			
2	1961 (56.0)	345 (32.6)	2306 (50.6)
3	881 (25.2)	268 (25.3)	1149 (25.2)
4	365 (10.4)	177 (16.7)	542 (11.9)
5+	292 (8.3)	269 (25.4)	561 (12.3)
Reason for Colonoscopy (%)			
Surveillance (<9 months since previous COL)	2101 (21.7)	525 (14.1)	2626 (19.6)
Surveillance (≥9 months since previous COL)	4120 (42.5)	2560 (68.8)	6680 (49.8)
Symptoms	2005 (20.7)	50 (1.3)	2055 (15.3)
Positive FOBT	1255 (12.9)	381 (10.2)	1636 (12.2)
Abnormal CT/Other	223 (2.3)	207 (5.6)	430 (3.2)
Colonoscopy Result (%)			
Normal	4160 (42.9)	2011 (54.0)	6171 (46.0)
Hyperplastic	766 (7.9)	369 (9.9)	1135 (8.5)
Adenoma	1795 (18.5)	267 (7.2)	2062 (15.4)
Advanced adenoma	1497 (15.4)	275 (7.4)	1772 (13.2)
CRC	453 (4.7)	24 (0.6)	477 (3.6)
Other	1033 (10.6)	777 (20.9)	1810 (13.5)
FOBTs	10,944	8999	19,943
FOBT result (%)			
Positive	1036 (9.5)	762 (8.5)	1798 (9.0)

†R0 – No family history but a personal history of adenoma or cancer.
R1 – CRC both parents; CRC two siblings; CRC parent and sibling; CRC twins.
R2 – CRC parent and grandparent; One first degree relative plus other family members with CRC; CRC parent and parent's sibling.
R3 – CRC first degree relative under 55 years of age.
R4 – CRC first degree relative over 55 years of age; CRC second or third degree relative.
CCC - Cluster of common colorectal cancers, that is, three or more first degree relatives with CRC all over 50 years at diagnosis.
HNPCC-suspected – suspected HNPCC, that is, two Amsterdam criteria.
HNPCC-definite – definite HNPCC Individuals in families meeting Amsterdam criteria, or carrying a mismatch repair mutation. SD, standard deviation; HNPCC, hereditary nonpolyposis colorectal cancer; CRC, colorectal cancer, COL, colonoscopy; FOBT, fecal occult blood test.

Adelaide and 732 from Melbourne), are not included in the data analysis because of the lack of follow-up. The two surveillance programs have different recruitment strategies. The Melbourne study recruits high-risk individuals based on a family history of CRC, and thus has a younger average age at first COL compared with Adelaide study, which recruits individuals with both a family and a personal history of CRC. The Melbourne study also has a larger average number of COLs per person. The eligibility criteria for the studies were that participants had to be at high risk for CRC based on a personal history of CRC or adenomas, or a family history of CRC. For the purpose of this paper, individuals younger than 40 years old at their first COL were excluded and only 20 years of follow-up data were retained for each patient. Summary statistics for the cohort used in our analysis are given in Table I.

The data consist of serial results of COLs and FOBTs, the dates at which these were measured, the reason for the test, and other information about the individual (age, gender, family history of colorectal cancer, and date of death). The result of each COL is a categorical variable taking one of six values: Normal, Other, Hyperplastic, Adenoma, Advanced adenoma, or CRC. For many patients, the scheduling of colonoscopies was dynamic, with the planned time of the next colonoscopy generally dependent on the result of the current colonoscopy. For someone with a clear (normal) colonoscopy, the typical interval to the next colonoscopy was 5 years; if a small lesion was found the typical interval was 3 years; if a serious lesion was found, the typical interval was 1 year. However, there are substantial departures from this pattern. The average time between successive COLs in the studied group is 3 years (SD = 1.8 years). The result of the FOBT is binary, either positive or negative. The timing of the FOBT was more variable. In some patients, it was every year, whereas in others, it was seldom. If a patient had a positive finding on a FOBT, it was nearly always followed by a COL within 3 months. For 101 colonoscopies (87 from Adelaide, 14 from Melbourne) that were characterized as being conducted because of a positive FOBT result, there was no prior positive FOBT result in the data set. For these cases, a prior positive FOBT result was assumed to have occurred and was coded into the data as occurring 1 month prior to the corresponding colonoscopy.

2. Statistical model and estimation

2.1. Notation

Let τ_{ij} be the age of patient i ($i = 1, \dots, n$) at their j-th colonoscopy (COL) ($j = 1, \dots, m_i$) where patient i has $m_i \geqslant 2$ COLs at ages $\tau_{i1}, \dots, \tau_{im_i}$. Let τ_{i0} be the youngest possible age for a COL, which we take as 40 years old.

The primary outcome variable of interest is the occurrence of advanced adenoma or cancer, denoted by AAC. The results of the COL for this primary endpoint are denoted by $Y_{i1}, \dots, Y_{im_i}$, where $Y_{ij} = 1$ indicates AAC found at age τ_{ij} and $Y_{ij} = 0$ indicates a finding of no AAC at age τ_{ij}.

Let X_i denote the baseline covariates associated with each individual (e.g., gender, family history, and city) and X_{ij} denote time-varying covariates associated with COL at age τ_{ij} (e.g., a binary indicator that the reason for the COL is symptoms).

Let $Z_i(a)$ denote a vector of time-dependent covariates. Each component of $Z_i(a)$ is constructed to represent some aspect of the knowledge about individual i at age a, for example, summary of FOBT history or summary of past COL results. Each $Z_i(a)$ will be defined at all ages throughout the individuals' follow-up. In the formulation described in the succeeding texts, we will assume that every Z_i is binary, but the methodology would allow them to be continuous time-dependent variables, for example, the total number of adenomas found in all previous colonoscopies. There is much flexibility in how the Z's are constructed to match various hypotheses that may be of interest. Specific choices for the Z's will be discussed later. See Figure 1 for an illustration of the described notation using the COL and FOBT history of a sample patient.

2.2. Statistical model

Assume that the hazard of an AAC event at age a (in years) for patient i is given by $\lambda_i(a)$. The specific time at which the AAC event occurs is the first age at which an advanced adenoma or cancer would be found if a colonoscopy were to be performed at that age. The occurrence of one AAC does not prevent the occurrence of later AAC events. We will assume that the events, if observable, follow a Poisson process for each person, with intensity denoted by $\lambda_i(a)$. This process is for when a new advanced adenoma or cancer would appear if the monitoring were continuous. In this formulation, the actual event, if observable,
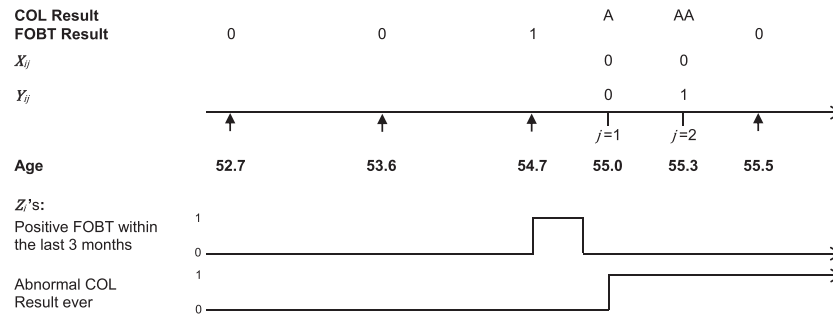
Figure 1. Illustration of the described notation using the colonoscopy (COL) and FOBT history of sample patient with $m_i = 2$ COLs (at $j = 1, 2$) and four fecal occult blood tests (FOBTs) (at ↑s). COL results are indicated by A (adenoma) and AA (advanced adenoma). FOBT results are indicated by 0 (negative) and 1 (positive). The X_{ij} corresponds to the time-varying COL-specific covariate 'binary indicator that the reason for the COL is symptoms.' The Y_{ij} corresponds to the binary outcome of AAC. Two possible Z_i's are indicated here: (1) A binary indicator of whether there was a positive FOBT in the past 3 months, and (2) A binary indicator of whether the patient ever had an abnormal COL. For (1) we see the corresponding Z jump to 1 at age 54.7 years, corresponding to a positive FOBT, and stay there for 3 months, after which it returns to 0. For (2) we see Z jump to 1 at age 55.0 years, corresponding to an abnormal COL, and stay that for the rest of the individual's follow-up.

could occur many times. So, for each person, $\lambda_i(a)$ changes with age, and the value at age a represents the chance of getting an AAC event in the next unit of time. From any age going forward, $\lambda_i(a)$ can be modified by new information you learn about the person. For example, if a new test were performed at age τ and the results suggest this person is at high risk, we would expect $\lambda_i(a)$ to increase at $a = \tau$.

Monitoring for AAC is not carried out continuously, and data are collected about AAC only at the time of each COL. Based on the model, the number of occurrences of AAC in (τ_{ij-1}, τ_{ij}) has a Poisson distribution with mean $(\int_{\tau_{ij-1}}^{\tau_{ij}} \lambda_i(a)da)$; and hence, the probability of at least one AAC in the interval (denoted by $Y_{ij} = 1$) is

$$P(Y_{ij} = 1) = P\left\{\text{event in interval}(\tau_{ij-1}, \tau_{ij})\right\} = 1 - \exp\left\{-\int_{\tau_{ij-1}}^{\tau_{ij}} \lambda_i(a)da\right\} \tag{1}$$

The rationale for integrating the hazard over the interval (τ_{ij-1}, τ_{ij}) is to accumulate the risk of AAC during intervals between COLs and model the probability of an event being higher following longer intervals. This reflects a unique data feature that a longer period between COLs represents a longer time during which advanced adenoma or cancer can develop, thus increasing the risk of AAC at the end of a longer interval between COLs.

A convenient form is to assume $\lambda_i(a)$ has a proportional hazards structure given by

$$\lambda_i(a) = \lambda_0(a) \exp\{\beta X_i + \theta Z_i(a)\} \tag{2}$$

We expect $\lambda_0(a)$ to be some smooth increasing function of age, and we will assume

$$\lambda_0(a) = \exp(\alpha_0)(a - 35)^{\alpha_1} \tag{3}$$

for ages greater than 40 years. We would expect $\alpha_1 > 0$ because adenoma rate tends to be higher in older people. Other parametric forms for $\lambda_0(a)$ are possible. For identifiability purposes, βX does not include an intercept term because it is captured by α_0.

Combining Equations (1) and (2), and letting $p_{ij} = P(Y_{ij} = 1)$, it can be seen that

$$\log\left\{-\log\left(1 - p_{ij}\right)\right\} = \beta X_i + \log\left[\int_{\tau_{ij-1}}^{\tau_{ij}} \lambda_0(a) \exp\left\{\theta Z_i(a)\right\} da\right]$$

which is a generalized nonlinear model for binary responses with a complementary log–log link.

The model is further extended to accommodate time-dependent covariates, X_{ij}, that can vary at different COLs.

$$\log\left\{-\log\left(1 - p_{ij}\right)\right\} = \beta_1 X_i + \beta_2 X_{ij} + \log\left[\int_{\tau_{ij-1}}^{\tau_{ij}} \lambda_0(a)\exp\{\theta Z_i(a)\}da\right] \qquad (4)$$

This model allows the probability of the COL test being positive to be modified by COL specific covariates, such as the reason for the test being patient symptoms.

2.3. Maximum likelihood estimation

Each patient has $m_i - 1$ intervals, each of which contributes to the likelihood function. Assuming independence between intervals, the likelihood contribution for patient i is

$$L_i(\beta, \theta, \alpha) = \prod_{j=2}^{m_i} \left\{P\left(Y_{ij} = 1\right)\right\}^{Y_{ij}} \left\{1 - P\left(Y_{ij} = 1\right)\right\}^{(1-Y_{ij})} \qquad (5)$$

where the expression for $P(Y_{ij} = 1)$ comes from Equation (4), and β, θ, and α are parameters, where α are the parameters in the equation for $\lambda_0(a)$. Note that $j = 1$ is not included in the product for the likelihood; thus, the results of each person's first COL are not used as an outcome variable.

The observed data likelihood is then

$$L(\beta, \theta, \alpha) = \prod_{i=1}^{n} L_i(\beta, \theta, \alpha) \qquad (6)$$

The form of the likelihood in Equations (5) and (6) is identical to that of a binary response model with $\Sigma_{i=1}^{n}(m_i - 1)$ observations and assuming independent Bernoulli distributions for each observation. Thus, it is possible to use software for generalized nonlinear models to obtain the parameter estimates.

Note that the parameters β always enter the model in a linear way; however, the parameters θ and those in $\lambda_0(a)$ enter into the model in a nonlinear way. For variables Z that do not change between intervals, there is some simplification. Specifically, their coefficients can be extracted from the integral and will be included as linear terms θZ. However, if the Z is a variable that can change between COLs, the parameters associated with that Z will remain in the nonlinear part of the model.

In this approach, the unit of analysis is the interval between COLs, rather than the person. Thus, for the purposes of obtaining point estimates, each interval is regarded as conditionally independent. However, because of the potential correlation between intervals for the same person, the usual information matrix is not expected to provide valid standard errors. Thus, we will use the bootstrap to obtain standard errors, where patients are resampled with replacement.

2.4. Discrete parametrization

Because of the large number of time-dependent covariates, to use available software to implement the integrations in Equation (4), we discretized time into 1-month unit. The integrals in the given model and likelihood functions are thus approximated by sums. The baseline hazard becomes a step function that has a jump discontinuity at each month. The ages at which individuals had test results were rounded to the nearest twelfth of a year. The covariate $Z_i(a)$ then represents the value of covariate Z at age a for individual i, where a is a multiple of 1/12. Every Z is represented by a matrix, where each column corresponds to a month and each row corresponds to an interval between colonoscopies. The discretized integrals for each colonoscopy interval are computed by parsing through the columns of the Z matrices. All statistical analyses and model building were conducted using R 3.0.2 software. Specifically, it is possible to use the 'gnm' package in R to fit the described generalized nonlinear model. However, depending on the complexity of the functions $\lambda_0(a)$ and $Z_i(a)$, the integral on the right hand side of Equation (4) may not have an analytic closed-form expression. Thus, the development of the model covariates for input into the 'gnm' function was challenging. An outline of the R code used is presented in the Supporting Information.

2.5. *Incorporation of someone's historical and prior information into the model*

For construction of the likelihood and estimation of the parameters in the model, information about the person that does not change over time, such as gender and family history, would be included as one of the fixed covariates X_i.

To capture the differential risk associated with a 'for reason' colonoscopy compared with a colonoscopy scheduled for surveillance, a colonoscopy-varying indicator variable indicating whether the reason for the colonoscopy was 'symptoms' was included in the model. Individuals with a colonoscopy conducted because of the symptoms are expected to have an increased probability of having an AAC result. This covariate contributes to the model as a fixed covariate X_{ij}, but varies for each colonoscopy interval rather than by individual.

The information about a person's history of COL and FOBTs and their results could be captured by $Z_i(a)$. We will call this their historical record, which includes everything that is known about a person's COL/FOBT testing both during and before the study. A sensible model will probably require a vector of time-dependent covariates, that is, $Z_{1i}(a), Z_{2i}(a), \ldots, Z_{pi}(a)$ to allow for different aspects of their history to be important. There is much flexibility in this approach, and there are many different choices that can be made for the Z's.

In the analysis of the data, we were particularly interested in incorporating prior COL and FOBT results and their impact on risk, and allowing for differences in the impact on risk if the prior COL/FOBT result was in the recent past compared with in the distant past. The specific choices for $Z_i(a)$ we considered were of the form 'has the patient had a test with result M in the last P years'. The set of M's we considered were hyperplastic/other on a COL, adenoma on a COL, advanced adenoma on a COL, cancer on a COL, and a positive FOBT. The time frames for P for the COL results we considered were the last 1 year, the last 1.5 years, the last 5 years and ever during the study period. The time frames considered for P for a positive FOBT were the last 3 months, the last 6 months, the last 9 months, and the last 1 year. With these definitions each Z is a time-dependent binary variable. For example, the Z corresponding to having an advanced adenoma ever during the study period would be coded as zero up to the age of the first finding of an advanced adenoma on a COL, and then as one at all later ages. The Z corresponding to having a positive FOBT within the last 3 months is illustrated in Figure 1 and would be coded as zero up to the time of the first positive FOBT, would jump to one at that age and remain at one for the next 3 months, then would change to zero at the next age for which there was not a positive FOBT within the last 3 months. It could then subsequently jump up to one again if the patient had a later positive FOBT.

Another intricacy to be considered in this particular situation is the overlapping of the effect of the Z covariates corresponding to abnormal COL results. Due to the binary nature of the identified Z covariates and the corresponding positive coefficients, an individual with multiple colonoscopies with various abnormal results in a short period (e.g., adenoma at the first colonoscopy, followed by an advanced adenoma within 6 months) would have an inflated predicted risk compared with someone with multiple colonoscopies but with the same, and possibly more severe, abnormal results at all colonoscopies (e.g., advanced adenoma at the first colonoscopy, followed by an advanced adenoma within 6 months). In such a situation we would expect that consecutive advanced adenoma results would increase an individual's risk of an AAC finding compared with an adenoma followed by an advanced adenoma. Thus, the effect of the identified Z matrix covariates was modified to last for the minimum of their chosen P time frame and the time until the individual's next colonoscopy result.

The $Z_i(a)$'s described previously capture many aspects of a person's current and past results of FOBT and COL tests. The dynamic nature of a person's sequence of tests is very complex, and it would certainly be possible to capture more aspects of the timing and results by defining additional Z's. For example, if the relevance of a test result differed with age, that could be captured by an interaction of the test result with a discrete version of age. The set of possible Z's one could consider is large.

Model selection was conducted based on minimization of model Akaike information criterion (AIC), which is a goodness-of-fit measure that is defined as AIC= $2\times$(number of parameters in the model)$-2\times$(maximized log likelihood). All considered models included the fixed effects of location, gender, risk group, and the colonoscopy-varying covariate of test reason being symptoms. The Z's considered to be of interest were those corresponding to having a colonoscopy result of hyperplastic/other, adenoma, advanced adenoma, and cancer within the past 1, 1.5, and 5 year(s), or ever while under surveillance. A Z corresponding to ever having any abnormal colonoscopy result, defined here as adenoma, advanced adenoma, or cancer, during the study period was also included. To account for the effect of FOBTs, we

considered Z's corresponding to a positive FOBT within the last 3, 6, and 9 months, or 1 year. Including multiple Z's corresponding to the same abnormal colonoscopy result (e.g., adenoma within the past 1 year and adenoma within the past 1.5 years) in the same model produced a negative coefficient for the corresponding covariate because of multiple covariates trying to describe the same effect. Thus, the selection of Z's was conducted by varying the time frame P for a single abnormal result in each model and then selecting the best combination of time frames for the abnormal COL results based on the model with the lowest AIC.

2.6. Interpretation of parameters

The β's are risk ratios associated with a unit change in the baseline covariates X_i or the reason for test covariate X_{ij}. The θ's are the risk ratios associated with unit changes in Z. Thus, for example, if one of the Z's was set up to represent any prior abnormal FOBT, then a large positive θ would suggest that people who had prior positive FOBTs anytime in the past are at an increased risk of AAC. Similarly, if another Z was set up to jump from zero to one if a person had an adenoma COL result, then an estimate of θ near zero for this Z would suggest that an adenoma result on a COL does not change the assessment of the subsequent risk of AAC for that person. If we had set up Z to represent at least two prior normal COLs, this might indicate the person was at lower than normal risk, in which case you would expect the estimate of θ to be negative.

The previous discussion illustrates the subtleties in interpretation and possible over-interpretation of the θ's. A unique feature of a COL is that it is also an intervention. The procedure attempts to remove any adenomas that are found, thus decreasing the risk of an adenoma developing into AAC. The described model specifies the risk of an event as a function of $Z(t)$, where $Z(t)$ represents the results of past testing. If $Z(t)$ captures colonoscopy testing and results, a person's actual risk may change as a result of having a COL. However, even though the COL procedure may reduce that person's risk because of the removal of an adenoma, we might still expect to see a positive estimate of θ because this person is now considered at higher long-term risk because of what has been learned about them compared with what was known about them before. In contrast, an FOBT is purely for detection purposes and a positive FOBT indicates that the person is in a subset of the population with possible increased risk of AAC, but because it is just a test it does not change that person's specific probability of advanced adenoma or cancer developing. Thus, if $Z(t)$ represents FOBTs and their results, the person's hazard of the event does not change as a result of the test; yet, we would expect the corresponding θ to be positive because this person is now considered at higher risk because of what has been learned about them compared with what was known about them before. In both of these cases, the result of the test has provided information about the person's risk, and this information is being represented by θ and is useful for making accurate predictions for the individual. Because our main purpose is prediction, the aforementioned difficulties in interpreting the θ's is not an over-riding limitation.

The parameters α_0 and α_1 define how the intensity of experiencing an AAC changes with age. The specific shape for $\lambda_0(a)$ should not be over-interpreted because it is simply a convenient functional form. Algebraically, it is the intensity for people who have either had no tests completed or had tests completed and all the test results are normal or negative. But, because the program enrolled high-risk people who had previous abnormal findings or were tested during the study, there are very few such people. The expected increasing nature of $\lambda_0(a)$ is simply a way to capture the idea that if the interval between COLs is long and the person has an AAC finding at the COL, it is more likely to have occurred towards the latter part of the interval because the occurrence of cancer and adenomas is more common in older than in younger people.

2.7. Standard errors

Standard error estimates were obtained using bootstrap methods. Three hundred bootstrap samples were generated by randomly selecting n (the total number of patients) individuals with replacement from the original data set. When an individual is selected for a bootstrap sample, all of the individual's COL and FOBT data are selected. Bootstrap standard errors account for model overdispersion and for some of the correlation between the observations on the same individual.

3. Results

3.1. Parameter estimates

The results are given in Table II. The negative parameter estimate for women indicates that a risk of AAC is decreased for women compared with men. There is no significant difference in the risk of AAC between the Melbourne and Adelaide groups, which is a notable finding given that these two studies had very different recruitment strategies. The coefficient for α_1 is greater than 0, indicating that risk significantly increases with age. Using the estimate for α_1, we compute that a 70-year old person has 2.5 times the hazard of AAC compared with a 50-year old if all other risk factors were the same. In comparison with those with a family history of R0, those in the hereditary nonpolyposis colorectal cancer (HNPCC) definite group have a significant increased risk of AAC. The R3, CCC, and HNPCC-suspected groups are also of increased risk, but the risk is not significantly different compared with those in R0. The R1 group has the lowest risk, but it is not statistically significantly different from the R0 group. If the reason for the COL is because the patient had symptoms, there is a significant increased risk of AAC. All of the abnormal colonoscopy results provided the best fit when the effect was considered 1 year past the observed result. A prior colonoscopy result of advanced adenoma produced the greatest increase in risk, followed by cancer, adenoma, and hyperplastic/other. The Hyperplastic and Other results were grouped together because they exhibited a similar pattern of AAC event rates when examining the event rate by previous COL result and time until the next COL. The estimates for all the prior colonoscopy results were all positive. Thus, the long-term risk associated with an abnormal COL result is greater than the benefit from having adenomas removed as a result of the COL procedure. A positive FOBT significantly affected risk when the effect was considered for 3 months past the positive result, and the positive coefficient indicates a significant increased risk for 3 months subsequent to a positive FOBT result.

Table II. Parameter estimates.

Variable	Estimate	Model SE	Bootstrap SE	*p*-value†		Hazard Ratio
Age						
Intercept (α_0)	−7.93	0.38	0.44			
α_1	1.10	0.11	0.13	< 0.001	***	
β						
Women	−0.13	0.07	0.07	0.08		0.88
Melbourne	−0.10	0.11	0.13	0.45		0.90
Symptoms	0.21	0.10	0.10	0.04	*	1.23
Family history‡						
R1	−0.26	0.18	0.21	0.23		0.77
R2	−0.10	0.17	0.18	0.55		0.90
R3	0.27	0.15	0.17	0.10		1.31
R4	−0.09	0.12	0.13	0.49		0.91
CCC	0.09	0.27	0.27	0.73		1.10
HNPCC-suspected	0.42	0.19	0.26	0.11		1.52
HNPCC definite	0.53	0.18	0.20	< 0.01	**	1.70
θ						
Adenoma < 1 yr	1.59	0.16	0.15	< 0.001	***	4.90
Advanced adenoma < 1 yr	2.46	0.12	0.13	< 0.001	***	11.8
Cancer < 1 yr	1.73	0.18	0.18	< 0.001	***	5.63
Hyperplastic/Other < 1 yr	1.42	0.14	0.15	< 0.001	***	4.15
Abnormal COL ever during the study period	0.35	0.11	0.11	< 0.01	**	1.41
Positive FOBT < 3 months	1.85	0.19	0.21	< 0.001	***	6.37

†The *p*-values are determined using the Bootstrap SEs in a Wald test for significance.
*$0.01 < p \leqslant 0.05$; **$0.001 < p \leqslant 0.01$; ***$p \leqslant 0.001$.
‡Family history covariates are relative to R0. HNPCC, hereditary nonpolyposis colorectal cancer; SE, standard error.

Of note from the results is that the risk ratios and significance levels of the updated COL and FOBT results are much stronger than for the baseline variables. Thus, more recent test results are a much more important determinant of a person's risk than family history and gender.

The standard deviations of the bootstrap estimates were similar, although sometimes slightly larger, than those from the fitted model that uses data from all patients. This suggests that the impact of overdispersion and within-subject correlation are not large in these data.

3.2. Goodness-of-fit

Calibration plots were used to compare the observed proportion of events with the probability of AAC predicted by the model. The estimated values for the probability that the COL was AAC (i.e., the estimates of p_{ij} from Equation 4) were categorized into bins of observations and the proportion of events observed within these groups was compared with the median of the bin's predicted probabilities. Model fit was evaluated by comparing the agreement between the two. Calibration plots were created for subsets of the data (e.g., by interval length, patient's age at observation, and surveillance status) to examine whether inconsistencies in agreement between predicted and observed proportions are associated with a particular covariate. Figure 2 shows the overall calibration plot and depicts a generally good agreement between the predicted probabilities and the observed proportions. The final dot represents fewer COL intervals with a wider range of prediction probabilities, and thus, the median predicted probability of AAC appears to be lower than the observed proportion of AAC events for those intervals. In addition, the proposed model tends to slightly underestimate the risk when the estimated risk of AAC is low. This phenomenon is more apparent in the examination of calibration plots by interval length, shown in Figure 3. For the intervals between COLs of less than 1 year, the observed proportion of AAC events is noticeably higher than the predicted probability of occurrence. The selected model does not fully capture the clinical relevance of short time between COLs. There are many reasons why a COL may be performed within 1 year of a previous COL. Some of these are included in the model, but there are others that may not be. For example, clinical judgment or factors not in the data set may suggest the patient is at higher risk and hence should return soon for another COL. This would likely lead to a higher probability of a positive COL finding than what is predicted by the model. Another reason is that the first COL may have been incomplete, or have given ambiguous or suspicious results, which would then require a 'follow-up' colonoscopy. This would also lead to a higher probability of a positive COL finding than what is predicted by the model. Although, in constructing the data set considerable effort was given to collapse such 'follow-up' COLs into a single COL, so this may be a minor reason for the higher observed rate.
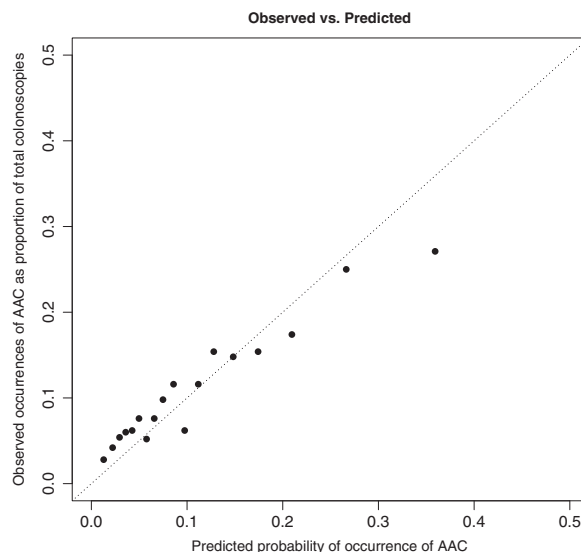

Figure 2. Overall calibration plot to compare the model's predicted probability with the proportion of observed events. Each dot represents 500 colonoscopy intervals, except for the last bin, which represents 369 intervals. Dots are centered at the median-predicted probability for the corresponding bin.
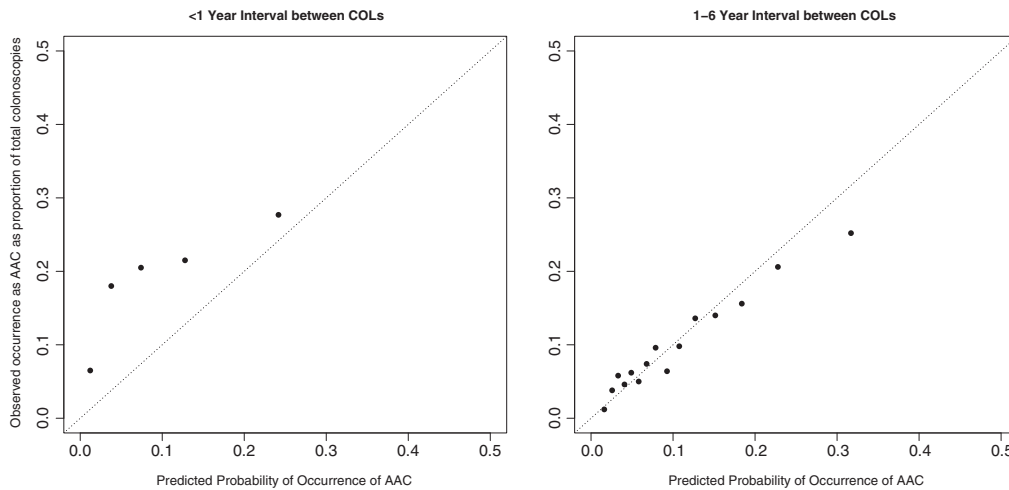
Figure 3. Calibration plot by interval length to compare the model's predicted probability with the proportion of observed events. (Left) Each dot represents 200 colonoscopy intervals, except for the last bin, which represents 130 intervals. (Right) Each dot represents 500 colonoscopy intervals, except for the last bin, which represents 51 intervals. Dots are centered at the median-predicted probability for the corresponding bin.

4. Individualized predicted survival distribution of time to a future AAC event

The model can be used to give the individualized survival distribution for time T_i, when the next AAC event occurs for patient i, given all that is known about that patient up to age A. Specifically, assume the prediction is to be made for person i at age A. Then, the probability of no new AAC within the next t years is

$$S(t) = P\{T_i > A + t | age = A, X_i, Z_i(A)\}$$

$$= \exp\left[-\int_A^{A+t} \lambda_i(a)da\right] \tag{7}$$

which is estimated by

$$\exp\left[-\exp(\hat{\beta}X_i)\int_A^{A+t} \hat{\lambda}_0(a)\exp\{\hat{\theta}Z_i(a)\}da\right] \tag{8}$$

This calculation requires specification of $Z_i(a)$ from the current age A to age $A + t$. Thus, we are making predictions under the assumption that the person will not have any COLs or other tests between age A and age $A + t$. Also, note that because the integral is from A to $A + t$ the actual values of $Z_i(a)$ for ages before A do not enter into this calculation; however, someone's record of COLs and FOBTs prior to age A may get captured depending on how $Z_i(a)$ is defined. Hazard and survival curves can be generated for individuals following their last follow-up using model estimates. Individuals are considered to be at increased risk if they have a positive colonoscopy result within the last 1 year or ever during the study period, depending on the covariate corresponding to that positive result in the model. The reason for the subsequent colonoscopies is assumed to be not because of symptoms. To represent the risk graphically, we plot one minus the survival distribution versus time from the current age. The risk (p) of an AAC can be read from the graph at any desired follow-up time (t). For a person, this risk can be interpreted in the following way: If there are 1000 people like you, with exactly the same age, family history, and series of test results, if the next colonoscopy for each of these people is t years in the future, without any intermediate tests, then $1000 \times p$ of them are expected to have a positive finding of AAC on that colonoscopy.

As an illustration, Table III describes the colonoscopy and FOBT follow-up data for four hypothetical patients of varying risk. Patients 1 and 2 are male Melbourne patients with R1 family risk and based on their test results can be categorized as being low-risk. Patients 3 and 4 are male Melbourne patients with the more risky HNPCC definite family history. Patient 1 has a different test sequence and all four patients

Table III. Example of patient historical record for low-risk male Melbourne patients with R1 family risk (Patients 1 and 2) and high-risk male Melbourne patients with HNPCC definite family risk (Patient 3 and Patient 4).

Age (years)	Test	Test result			
		Patient 1	Patient 2	Patient 3	Patient 4
52.65	FOBT	Negative	Negative	Negative	Negative
53.55	FOBT	Negative	Negative	Negative	Negative
54.72	FOBT	Positive	Positive	Positive	Positive
55.04	COL	Normal	Advanced adenoma	Adenoma	Adenoma
55.29	COL	—	Adenoma	Advanced adenoma	Advanced adenoma
55.58	FOBT	—	Negative	Negative	Positive

HNPCC, hereditary nonpolyposis colorectal cancer; FOBT, fecal occult blood test; COL, colonoscopy.
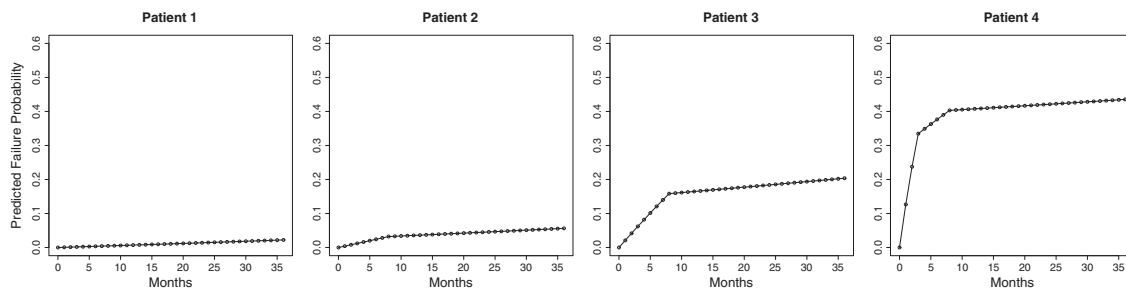

Figure 4. Predicted failure probability $(1 - \hat{S}(t))$ for the next 3 years past the current age of 55.58 years for four individuals with varying risk, for whom historical records of colonoscopies and fecal occult blood tests are given in Table III.

have differing test results. In Figure 4, we show the individualized predicted event probability curves for these four patients, assuming no symptoms and no additional FOBTs for the prediction period. Reading from the graphs, the 3-year risk for the patients 1, 2, 3, and 4 are 0.02, 0.06, 0.20, and 0.44, respectively. These graphs were produced using the parameter estimates in Table II and Equation 8. The rate of the continuous increase of the predicted event probability is determined by the effect of age and the fixed effects of gender, location, and family risk. Because of the final selected model, a positive FOBT only has an effect for the subsequent 3 months, and a particular abnormal COL only has an effect for the following 1 year. In addition, Patients 2, 3 and 4 have increased risk because of the positive effect of the 'abnormal COL ever during the study period' coefficient. Because of the dynamic nature of the prediction model, Patients 2 and 3 have different predicted risk trajectories based on the difference in their history of test results. This would not have been reflected by standard prediction tools that only incorporate the effect of the most recently observed event.

In the long run, one could imagine a website where a person specifies a horizon of time t and inputs their values for their current age A, their baseline variables X_i, and their history of COL and FOBT results. The website will then calculate and output $1 - \hat{S}(t)$, the chance of an AAC within the next t years. This could be a useful aid in helping to decide when to schedule the next COL or FOBT.

5. Discussion

The proposed model produces prediction of risk in a dynamic manner. As new information becomes available about a person's COL and FOBT results, the model incorporates its effects and updates the predicted risk of developing the outcome of interest. This type of model is particularly useful for the surveillance data studied in this paper because it incorporates the longitudinal COL and FOBT results from the frequent testing of this high-risk group. A variety of covariates could have been chosen to capture the effect of updated test information. Abnormal COL results were found to have a significant effect on increased AAC risk for 1 year past the abnormal finding. A positive FOBT was found to significantly increase risk for 3 months following the test. The covariates that incorporated the updated information aspect of the model were of greater significance than the baseline variables.

The timing of the COLs themselves perhaps carries some information about the disease development. While many of the reasons a person comes in for a COL are included in the model, such as positive FOBT, age, symptoms, and regular surveillance, there may be other reasons, such as type of symptoms, other lab tests, and change in family history, which were not included. These might provide information about the level of risk, suggesting that if the process of the timing of the COLs had been modeled, additional information may have been obtained about a person's risk. Developing such a model would require more information than what is available in this data set.

The estimation method presented here does not utilize any of the information that may be known about a person after their last COL. Specifically, we may know the date of death for many people. Quite a number of people have one or more FOBT after their last COL, and all the FOBT results are negative. For people such as these, updated predictions like those shown in Figure 4 could be made after the latest FOBT test, as was illustrated by Patients 2, 3, and 4 in the Figure.

The initial formulation of the model was based on a Poisson process, which lead in a natural way to a model for the observed binary response data. The binary response model, together with an assumption of independence gave a likelihood, from which point estimates were obtained. Measures of uncertainty were obtained using the bootstrap. Thus, the validity of all the aspects of the Poisson process assumptions is not being relied on. We regard it simply as a good starting point to develop a model.

Some sensitivity analyses were undertaken. The analysis was run separately for the Adelaide and Melbourne cohorts. We found that the parameter estimates were quite close to each other, and the location-specific 95% confidence intervals had considerable overlap for nearly all the parameters. This, together with the small estimated location effect in Table II, provides justification for combining the data sets in one model. A second sensitivity analysis was to discretize the intervals into 2 months, instead of 1 month intervals. This gave very similar estimates to those shown in Table II and Figure 4.

For the calibration plots in Figures 2 and 3, the predictions are for the same patients which were used to derive the parameter estimates. As a more honest assessment, leave-one-out cross-validation was conducted. For each patient, the model was fit to the data set excluding the patient's data, and the estimated coefficients were applied to the left out patient to give predicted probabilities of AAC for that patient's intervals. Calibration plots were created using the resulting predicted probabilities and were found to be very close to those produced by the fitted model. A different form of validation is external validation, in which predictions are applied to an external data set. We mimicked this by fitting the model to the Melbourne data and using the estimated coefficients to compute the predicted probabilities of AAC for the Adelaide data. Calibration plots for the Adelaide data were found to be similar to those in Figures 2 and 3. This provides evidence of generalizability of the fitted model to an independent data set.

There are many ways to consider adapting the presented model or to develop alternative approaches. For example, including a subject-specific frailty to better accommodate within and between-person heterogeneity would be a natural consideration. Such a model formulation would only include frailty terms in the hazard model and not include the Z's, but would then require models in which the Z's are outcome variables that depend on the frailty. Such a model could be challenging to formulate because it would require modeling over time the development of all five levels of the COL results and the FOBT results. In the model we have presented, the Z_i terms can be regarded as a substitute for a frailty term, and one can think of the Z_i's as they are observed and measured as providing information about the frailty. This is the rationale for not including frailty in the model.

An alternative method for developing dynamic predictions is through landmarking [18]. In this approach, at each age, a regression model of the residual time to the next AAC event is developed. The covariates in this model would need to be chosen to represent the current and past history of the person's tests results and their baseline characteristics. Developing such covariates and implementing this method using available dynamic model software would be challenging. The regression coefficients for each covariate themselves depend on age, and these would need to be linked. Overall, this would be a viable, although complex, approach, and it would be interesting to compare the predictions from such an approach with the predictions from the stochastic process model developed in this paper. Because the outcome being modeled is binary, it is also possible to explore a simpler linear logistic model approach. The challenge with such an approach would be what covariates to use. The length of time since the last COL could be included to capture the increased risk of AAC for longer intervals between COLs that results from having more time for advanced adenoma or cancer to develop. However, we think this is less satisfactory than capturing the effect of the interval width through integrating the intensity function over the time interval, as in Equations 1 and 4.

Aspects of the Z_i's can be thought of as 'internal' covariates in the Cox model jargon, that is, they themselves reflect disease progression. There is a reasonably large literature on joint modeling of longitudinally measured covariates together with event times [19, 20]. There are potential biases in estimates of hazard model parameters if one does not take into account the longitudinal data. This could require a complex model for the development and progression of adenomas and how they relate to what one can actually measure. Because the Z's are internal variables, which can be regarded as outcome variables, interpretation of the θ's may be tricky; however, for the purpose of future prediction, this is less of a concern.

The individualized predictions presented in this paper are intended to be an aid to the patient and physician to help schedule the next colonoscopy or FOBT. Currently, there are guidelines that a population can follow. The work presented here would facilitate a more personalized approach. For example, a COL might be scheduled at the time at which the risk is 10% or at 5 years, whichever is sooner, and a FOBT might be scheduled at the time at which the risk is 5%.

References

1. Day DW, Morson BC. The adenoma-carcinoma sequence. *Major Problems in Pathology* 1977; **10**:58–71.
2. Winawer SJ, Zauber AG. The advanced adenoma as the primary target of screening. *Gastrointestinal Endoscopy Clinics of North America* 2002; **12**(1):1–9.
3. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, Kuntz KM. Evaluating test strategies for colorectal cancer screening: a decision analysis for the US Preventive Services Task Force. *Annals of Internal Medicine* 2008; **149**(9):659–669.
4. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. *Journal of the American Medical Association* 2000; **284**(15):1954–1961.
5. Lieberman DA. Cost-effectiveness model for colon cancer screening. *Gastroenterology* 1995; **109**(6):1781–1790.
6. O'Leary BA, Olynyk JK, Neville AM, Platell CF. Cost-effectiveness of colorectal cancer screening: comparison of community-based flexible sigmoidoscopy with fecal occult blood testing and colonoscopy. *Journal of Gastroenterology and Hepatology* 2004; **19**(1):38–47.
7. Allison JE, Tekawa IS, Ransom LJ, Adrain AL. A comparison of fecal occult-blood tests for colorectal-cancer screening. *New England Journal of Medicine* 1996; **334**(3):155–160.
8. Levin B, Lieberman DA, McFarland B, Smith RA, Brooks D, Andrews KS, Dash C, Giardiello FM, Glick S, Levin TR, Pickhardt P, Rex DK, Thorson A, Winawer SJ. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US multi-society task force on colorectal cancer, and the American College of Radiology. *CA: A Cancer Journal for Clinicians* 2008; **58**(3):130–160.
9. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* 2011; **67**(3):819–829.
10. Taylor JMG, Park Y, Ankerst DP, Proust-Lima C, Williams S, Kestin L, Bae K, Pickles T, Sandler H. Real-time individual predictions of prostate cancer recurrence using joint models. *Biometrics* 2013; **69**(1):206–213.
11. Rutter CM, Yu O, Miglioretti DL. A hierarchical non-homogenous Poisson model for meta-analysis of adenoma counts. *Statistics in Medicine* 2007; **26**(1):98–109.
12. Bampton PA, Sandford JJ, Cole SR, Smith A, Morcom J, Cadd B, Young GP. Interval faecal occult blood testing in a colonoscopy based screening programme detects additional pathology. *Gut* 2005; **54**(6):803–806.
13. Bampton PA, Sandford JJ, Young GP. Achieving long-term compliance with colonoscopic surveillance guidelines for patients at increased risk of colorectal cancer in Australia. *International Journal of Clinical Practice* 2007; **61**(3):510–513.
14. Lane JM, Chow E, Young GP, Good N, Smith A, Bull J, Sandford J, Morcom J, Bampton PA, Cole SR. Interval fecal immunochemical testing in a colonoscopic surveillance program speeds detection of colorectal neoplasia. *Gastroenterology* 2010; **139**(6):1918–1926.
15. Dowling DJ, St John DJB, Macrae FA, Hopper JL. Yield from colonoscopic screening in people with a strong family history of common colorectal cancer. *Journal of Gastroenterology and Hepatology* 2000; **15**(8):939–944.
16. Brown GJE, St John DJB, Macrae FA, Aittomäki K. Cancer risk in young women at risk of hereditary nonpolyposis colorectal cancer: implications for gynecologic surveillance. *Gynecologic Oncology* 2001; **80**(3):346–349.
17. Good N, Macrae F, Young G, O'Dywer J, Slattery M, Venables W, Lockett T, O'Dywer M. Ideal colonoscopic surveillance intervals to reduce incidence of advanced adenoma and colorectal cancer. *Journal of Gastroenterology and Hepatology* 2015. DOI: 10.1111/jgh.12904.
18. van Houwelingen H, Putter H. *Dynamic Prediction in Clinical Survival Analysis*. CRC Press, Inc.: Boca Raton, FL, 2011.
19. Yu M, Law NJ, Taylor JMG, Sandler HM. Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica* 2004; **14**(3):835–862.
20. Wang Y, Taylor JMG. Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* 2001; **96**(455):895–905.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.