STATISTICS IN MEDICINE
Statist. Med. 2002; **21**:3369–3382 (DOI: 10.1002/sim.1277)

Assessing chronic disease progression using non-homogeneous exponential regression Markov models: an illustration using a selective breast cancer screening in Taiwan

Hsin-Ju Hsieh[1], Tony Hsiu-Hsi Chen[2,*,†] and Shu-Hui Chang[1]

[1] *Biostatistics Unit, Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan*
[2] *Institute of Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan*

SUMMARY

Previous research on estimation of the progression of chronic disease, from the normal preclinical screen-detectable phase (PCDP) to the final clinical phase, has usually assumed constant transition rates and has rarely addressed how relevant covariates affect multi-state transitions. The present study proposes two non-homogeneous models using the Weibull distribution and piecewise exponential model, together with covariate functions of the proportional hazard form, to tackle these problems. We illustrate the models by application to a selective breast cancer screening programme. The results of the Weibull model yield estimates of scale and shape parameters for annual preclinical incidence rate as 0.0000058 (SE = 0.0000019) and 2.4755 (SE = 0.1153), the latter being significantly higher than 1. Annual transition rate was estimated as 0.3153 (SE = 0.1385). Relative risks for the effects of late age at first pregnancy (AP) and high body mass index (BMI) on preclinical incidence rate were 1.98 and 2.59, respectively. The corresponding figures on the transition from the PCDP to clinical phase were 1.56 and 1.99, respectively. Non-homogeneous Markov models proposed in this study can be easily applied to rates of progression of chronic disease with increasing or decreasing rates with time and to model the effect of relevant covariates on multi-state transition rates. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: non-homogeneous Markov process; exponential regression model; breast cancer screening

1. INTRODUCTION

Progression of chronic diseases is often depicted by three states: normal, preclinical screen-detectable phase (PCDP), and clinical phase [1]. Some malignancies, in particular breast cancer, are considerably more successfully treatable when diagnosed early. In order to detect the disease as early as possible in the PCDP, mass screening has been carried out, using mammography. However, screening programmes are not always equally effective. For instance,

* Correspondence to: Tony Hsiu-Hsi Chen, Institute of Preventive Medicine, College of Public Health, National Taiwan University, Room 207, No. 19, Hsuchow Road, Taipei, Taiwan.
† E-mail: stony@episerv.cph.ntu.edu.tw

Received December 2000
Accepted January 2002

screening for breast cancer with mammography for women under 50 years of age showed rather heterogeneous results [2]. As a matter of fact, the efficacy of screening for chronic disease is usually dependent on how quickly the disease may progress from the preclinical screen-detectable phase (PCDP) (asymptomatic) to the clinical phase (symptomatic). A clear understanding of progression from the PCDP to the clinical phase is therefore important for chronic disease screening. For instance, the long duration of the PCDP for prostate cancer [3] provides an opportunity to detect cancers at an early stage and also lowers the likelihood of false negative cases provided that an optimal screening interval is deployed.

There are a variety of approaches for estimation of parameters with respect to the above three-state disease model. These include simple prevalence/incidence ratio [4], a non-parametric lead-time method [5] and a mathematical model based on prevalent cancer plus interval cancer cases [6]. To render estimation stable, the constraint of preclinical incidence rate to the clinical incidence rate in the absence of screening was usually required. Recently, a three-state Markov model was proposed to explicitly estimate the preclinical incidence rate and the transition rate from the PCDP to clinical phase [7–9]. There are two drawbacks to this approach. First, there is a constant preclinical incidence rate as a result of the time-homogeneous property of the Markov model. This may be unrealistic. For example, the Armitage–Doll multi-stage model [10] with respect to carcinogenesis proposes that the age-specific incidence rate of many human adult carcinoma increases roughly as a power function of age, which may be approximated by a Weibull form. It is therefore worthwhile developing a three-state model dispensing with the assumption of constant preclinical incidence. Secondly, allowing for relevant risk factors, or covariates, was rarely addressed in previous Markov models partly because the transition from no detectable disease to the PCDP is usually unobservable and partly because the disease natural history from the PCDP to the clinical phase is also unobservable due to the interruption of medical regime. Addressing the association between risk factors and the disease natural history may be even more important for those removable variables such as obesity or smoking. Medical consultation at regular intervals for different characteristics among women can be suggested based on this knowledge. Few studies, to the best of our knowledge, have been carried out in the context of breast cancer screening. In this study, non-homogeneous Markov exponential regression models were proposed to accommodate the above two problems.

2. AN EXAMPLE OF SELECTIVE BREAST CANCER SCREENING IN TAIWAN

We are faced with a particular example of the above two problems in evaluating selective screening for relatives of breast cancer index cases. Details of the study design for this high-risk group screening project, Taiwan Multicentre Cancer Screening (TAMCAS), have been published elsewhere [11]. In brief, this was a multi-centre project to identify early cancer for a population at high risk via different screening tools. Here the high-risk group of breast cancer is identified as female relatives of breast cancer cases, including mothers, daughters, sisters and grandmothers. Since 1994, relatives over 35 years old of breast cancer index cases from 12 hospitals were invited to annual screening by a combination of mammography, ultrasound and physical examination. Those with either one positive result on the above three tests were further confirmed by histological biopsy. Once an individual is affirmed as a breast cancer case, she will be properly treated with medical care. Up to 1997, a total of 4280 women received

their first screen and their age ranged from 34 years old to 86 years old (with an average age of 46.67 (SD = 9.004)). However, it should be noted that about 75 per cent of the subjects were under 50 years old; 50 breast cancer cases (which are called prevalent cases) were found among them. A total of 1584 women received the second screening one year after their first screen, and 8 cases (incidence cases) were found. No interval cases were obtained in this study.

Since the overall screening efficacy of the high-risk approach has been demonstrated in a previous study [11], a series of subsidiary questions will be asked. First, does the incidence rate of preclinical breast cancer among this high-risk group increase, or decrease (non-homogeneous property) or stabilize (homogeneous property) with age? To gain a better understanding of this information is very helpful when women with a family history of breast cancer ask for a medical consultation to pre-empt their own risk of getting breast cancer. In addition, if the force of entering into preclinical breast cancer is age-dependent, the health authority can set up the priority for selecting a target population in the light of this information under limited resources. Secondly, in order to make such a selective screening as efficacious as possible for each woman, one may ponder how screening policy can be established with the incorporation of relevant covariates into three-state disease progression, from normal, the PCDP phase and finally to clinical phase as described above.

Modelling the impact of relevant covariates on three-state transitions has a significant implication for prevention of breast cancer. A risk factor for breast cancer may act either as an initiator for onset of preclinical screen-detectable breast cancer (PCDP), or as a promoter accelerating the progression from PCDP to the clinical phase or both. If it acts as an initiator, primary prevention by removing the factor should be addressed. If it behaves as a promoter, different screening policies, such as more frequent screening for people carrying this factor, might be required.

In the TAMCAS breast cancer data, there are three suspected risk factors of interest, including late age at first full-term pregnancy, obesity and early age at menarche. These factors are referred to as covariates and abbreviated as AP, BMI and MF in this paper. The definition of these covariates are described as follows:

$\mathrm{AP} = 1$ if age at first full-term pregnancy is older than 25 years old, 0 otherwise;
$\mathrm{BMI} = 1$ if one's body mass index ($\mathrm{kg/m^2}$) is greater than 23, 0 otherwise;
$\mathrm{MF} = 1$ if one's age at menarche is earlier than 13 years old, 0 otherwise;

Throughout the following text, the positive cases with respective to AP, BMI and MF are defined as $\mathrm{AP} = 1$, $\mathrm{BMI} = 1$ and $\mathrm{MF} = 1$, respectively, otherwise negative cases.

The rest of this paper is organized as follows. Section 3 delineates model specification for a non-homogeneous Markov model taking covariates into account. Likelihood function and estimation of parameters are also derived. Section 4 illustrates how the proposed model is applied to the data.

3. MODEL SPECIFICATION

To model the natural history of a three-state Markov process, transitions from normal (state 0) to the PCDP (state 1) and the PCDP (state 1) to clinical phase (state 2) at time t

are depicted as follows:

$$\begin{array}{ccccc} \text{(State 0)} & \xrightarrow{q_{01}(t)} & \text{(State 1)} & \xrightarrow{q_{12}(t)} & \text{(State 2)} \\ \text{Normal} & & \text{PCDP} & & \text{Clinical phase} \end{array}$$

$q_{01}(t)$ and $q_{12}(t)$ are defined as instantaneous rates (transition intensities) at time t from state 0 to state 1 and from state 1 to state 2, respectively. The mathematical property of transition intensities is shown in Appendix A.

To take account of the effect of covariates on $q_{01}(t)$ and $q_{12}(t)$, exponential regression models are used to model the different characteristics of random process between individuals. For the transition from normal (state 0) to the PCDP (state 1), the effects of AP (X_1), BMI (X_2) and MF (X_3) are incorporated in $q_{01}(t)$ by the following form:

$$q_{01}(t) = q_{010}(t) \times \exp(\beta_{11}X_1 + \beta_{12}X_2 + \beta_{13}X_3) \tag{1}$$

where $q_{010}(t)$ denotes for the baseline intensity when three covariate values are equal to 0. β_{11}, β_{12} and β_{13} stand for regression coefficients corresponding to X_1, X_2 and X_3, respectively. A similar expression is applied to $q_{12}(t)$ from the PCDP (state 1) to clinical phase (state 2)

$$q_{12}(t) = q_{120}(t) \times \exp(\beta_{21}X_1 + \beta_{22}X_2 + \beta_{23}X_3) \tag{2}$$

It should be noted that covariates accounting for transition intensity $q_{01}(t)$ may be different from those for $q_{12}(t)$.

3.1. A non-homogeneous three-state Markov process with Weibull distribution

As pointed out earlier, preclinical incidence may not be constant across all ages. That is, the transition time from state 0 to state 1 or even state 1 to state 2 may not follow an exponential distribution. Here a Weibull distribution with scale parameter λ and shape parameter k is considered to accommodate this non-homogeneous Markov property ($k > 1$, increasing with age; $k < 1$, decreasing; $k = 1$, constant). The baseline intensity at age t, $q_{010}(t)$ in expression (1) is defined as

$$q_{010}(t) = \lambda_1 k_1 t^{k_1 - 1} \tag{3}$$

$q_{120}(t)$ in expression (2) is

$$q_{120}(t) = \lambda_2 k_2 t^{k_2 - 1} \tag{4}$$

Expressions (1) to (4) enable one to develop transition probabilities between age t_1 and t_2 of the three-state progressive Markov process as shown in Table I.

$p_{00}(t_1, t_2)$ and $p_{11}(t_1, t_2)$ represent the probabilities of remaining state 0 and state 1 between age t_1 and t_2. Transition probabilities from state 0 to state 1, from state 0 to state 2 and from state 1 to state 2 between age t_1 and t_2 are denoted by $p_{01}(t_1, t_2)$, $p_{02}(t_1, t_2)$ and $p_{12}(t_1, t_2)$, respectively. A detailed derivation of these transition probabilities in the light of expressions (1)–(4) is shown in Appendix B. Fitting these transition probabilities to data on selective breast cancer screening can throw light on whether the incidence rate of preclinical breast

Table I.

Previous state	Current state		
	Normal (state 0)	PCDP (state 1)	Clinical phase (state 2)
Normal (state 0)	$p_{00}(t_1,t_2)$	$p_{01}(t_1,t_2)$	$p_{02}(t_1,t_2)$
PCDP (state 1)	0	$p_{11}(t_1,t_2)$	$p_{12}(t_1,t_2)$
Clinical phase (state 2)	0	0	1

cancer varies with age and the effects of AP, BMI and MF on transitions from normal to the PCDP and from the PCDP to the clinical phase.

3.2. A non-homogeneous three-state Markov process with piecewise exponential model

As the likelihood function based on Weibull models may be complicated and estimation would be time-consuming, an alternative is therefore proposed using a piecewise exponential model. Since the exponential distribution is a special case of Weibull distribution for the shape parameter k equals to 1, the expressions of baseline intensity $q_{010}(t)$ and $q_{120}(t)$ in (3) and (4) are reduced to $q_{010}(t)=q_{010}=\lambda_1$ and $q_{120}(t)=q_{120}=\lambda_2$. Thus according to a cut-off point time t^*, the transition intensity $q_{01}(t)$ for a piecewise exponential model can be defined as

$$\begin{cases} q_{01}=\lambda_1\times\exp(\beta_{11}X_1+\beta_{12}X_2+\beta_{13}X_3) & \text{if } t\leqslant t^* \\ q'_{01}=\lambda'_1\times\exp(\beta'_{11}X_1+\beta'_{12}X_2+\beta'_{13}X_3) & \text{if } t>t^* \end{cases}$$

Similarly, $q_{12}(t)$ can be defined as

$$\begin{cases} q_{12}=\lambda_2\times\exp(\beta_{21}X_1+\beta_{22}X_2+\beta_{23}X_3) & \text{if } t\leqslant t^* \\ q'_{12}=\lambda'_2\times\exp(\beta'_{21}X_1+\beta'_{22}X_2+\beta'_{23}X_3) & \text{if } t>t^* \end{cases}$$

For example, the age of 50 years (around menopause) is applied as t^* in our following piecewise exponential models.

3.3. Likelihood function and parameter estimation

The maximum likelihood estimation method is used to estimate parameters. Let $\theta=(\lambda_1,\lambda_2,k_1,k_2,\beta_{11},\beta_{12},\beta_{13},\beta_{21},\beta_{22},\beta_{23})$. Suppose n_1 individuals attend the first screen, the likelihood based on the prevalent screen is

$$L_1(\theta)=\prod_{i=1}^{n_1}\left[\frac{p_{01}(0,t_{1_i})}{p_{00}(0,t_{1_i})+p_{01}(0,t_{1_i})}\right]^{v_{1_i}}\left[\frac{p_{00}(0,t_{1_i})}{p_{00}(0,t_{1_i})+p_{01}(0,t_{1_i})}\right]^{1-v_{1_i}} \tag{5}$$

where $v_1=1$, if individual i is detected as a positive case, $v_1=0$, otherwise; t_1 is the age at first screen.

Note that this likelihood is a product of conditional probabilities, which are conditional on no clinical cancers at first screen, because those with previous clinical breast cancer will be excluded at first screen.

The likelihood based on the second screen is

$$L_2(\theta) = \prod_{j=1}^{n_2} [p_{01}(t_{1_j}, t_{2_j})]^{v_{2j}} [p_{00}(t_{1_j}, t_{2_j})]^{1-v_{2j}} \tag{6}$$

where n_2 is the total number of individuals attending the second screen, t_2 denotes age of the second screen; $v_2 = 1$, if individual j is detected as a positive case, $v_2 = 0$, otherwise.

The total likelihood function

$$L(\theta) = L_1(\theta) \times L_2(\theta) \tag{7}$$

is a function of the desired parameters θ.

The solution of score function and information matrix based on first and second derivatives of the logarithm of the likelihood function gives the maximum likelihood estimates (MLE) and standard errors. An SAS program using PROC IML is written to estimate the parameters. The program is available from the authors (T.H.H. Chen) upon request.

3.4. Model selection and diagnostics

Likelihood ratio test was used for selecting the parsimonious model among a series of nested models. In order to demonstrate the versatility of our exponential regression models, model selections not only include the addition of significant covariates or the deletion of superfluous covariates but also compare the models with covariates affecting both types of transitions (state 0 to state 1 and state 1 to state 2) with those that only include the transition from state 0 to state 1 or from state 1 to state 2. To assess the goodness-of-fit of a model, we may compare the observed number of transitions between particular states with the expected, and a Pearson χ^2 test statistics can be used to judge whether there is a good fit for the model.

4. RESULTS

Data on number of cases and relevant transition types at first and second screens of the data described in Section 2 is shown in Table II. The application of expression (7) to Table II

Table II. Transition type, transition probabilities and number of transition applied for the three-state Markov model, breast cancer screening data from TAMCAS.

Screening rounds and detection mode	Transition type	Transition probability	Number
First screen			
Prevalent case	$0 \to 1$	$\frac{p_{01}(0,t_1)}{p_{00}(0,t_1)+p_{01}(0,t_1)}$	50
Normal	$0 \to 0$	$\frac{p_{00}(0,t_1)}{p_{00}(0,t_1)+p_{01}(0,t_1)}$	4230
Second screen			
Incident case	$0 \to 1$	$p_{01}(t_1, t_2)$	8
Normal	$0 \to 0$	$p_{00}(t_1, t_2)$	1576

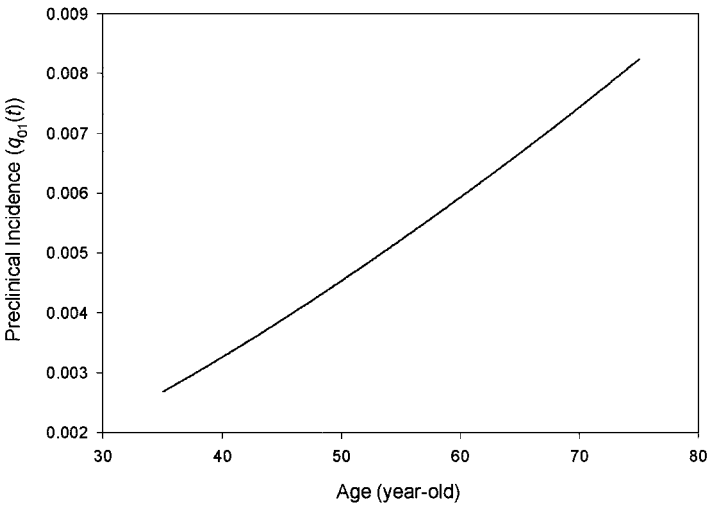

Figure 1. Preclinical incidence estimated from a Weibull model.

yields the total likelihood

$$L(\theta)=\prod_{i=1}^{4230}\frac{p_{00}(0,t_{1_i})}{p_{00}(0,t_{1_i})+p_{01}(0,t_{1_i})}\prod_{j=1}^{50}\frac{p_{01}(0,t_{1_j})}{p_{00}(0,t_{1_j})+p_{01}(0,t_{1_j})}\prod_{m=1}^{1576}p_{00}(t_{1_m},t_{2_m})\prod_{n=1}^{8}p_{01}(t_{1_n},t_{2_n})$$

To fit the selective breast cancer screening data, we consider transition time from normal to PCDP follows a non-homogeneous distribution such as the Weibull or the piecewise exponential. However, since previous studies [6] have shown a good fit of exponential distribution for sojourn time in the PCDP and the majority of women are under 50 years, to simplify estimation, we apply an exponential distribution to model the transition from the PCDP to clinical phase.

4.1. Weibull models

The Weibull model without covariates (the null model) gives estimate of λ_1, k_1 in expression (3) as 0.0000058 (SE = 0.0000019), 2.4755 (SE = 0.1153), respectively, and $\hat{q}_{12}=0.3153$ (SE = 0.1385). As the 95 per cent confidence interval of k_1 is 2.2496–2.7015, this suggests that preclinical incidence significantly increase with age as depicted in Figure 1.

Table III shows the results of model selection based on the likelihood ratio test. In models with single covariate, both transitions, state 0 to state 1 and state 1 to state 2, are highly associated with BMI ($x^2_{(2)}=24.1$, $P<0.01$) and AP ($x^2_{(2)}=59.36$, $P<0.01$), as compared with the null model. The effect of MF on both transitions is lacking of statistical significance ($x^2_{(2)}=3.50$, $P=0.17$). When models addressing both transitions are compared with those only highlighting one transition, only the model addressing the effect of BMI on state 0 to

Table III. Results of −2 log-likelihood for each Weibull model, with a combination of covariates.

Covariates	−2 log-likelihood	Degree of freedom
1. Null model	640.12	127
2. One covariate		
(1) Both transitions		
AP	580.76	125
BMI	615.98	125
MF	636.62	125
(2) One transition		
AP on state 0 to state 1	581.28	126
BMI on state 0 to state 1	622.43	126
MF on state 0 to state 1	636.63	126
AP on state 1 to state 2	582.06	126
BMI on state 1 to state 2	617.8	126
MF on state 1 to state 2	636.64	126
3. Two covariates		
(1) Both transitions		
AP, MF	578.34	123
BMI, MF	612.52	123
AP, BMI	557.48	123
(2) One transition for AP		
AP on state 0 to state 1		
BMI on both transitions	557.66	124
AP on state 1 to state 2		
BMI on both transitions	558.30	124
(3) One transition for BMI		
BMI on state 0 to state 1		
AP on both transitions	557.98	124
BMI on state 1 to state 2		
AP on both transitions	558.82	124
4. Three-covariate		
AP, BMI, MF	554.04	122

state 1 shows statistical significance as compared with the corresponding model addressing the effect of BMI on both transitions. This suggests that BMI may be an important covariate for the transition from state 0 to state 1. For models with two covariates, similarly MF is not a significant factor whereas both AP and BMI may be significantly responsible for the occurrence and progression of breast cancer. However, after including AP and BMI into the same model, there is a lack of substantial differences between models addressing both transitions (state 0 to state 1 and state 1 to state 2) and models only taking one transition. Results for parameter estimates and standard errors of models with covariates of AP and BMI simultaneously are listed in Table IV. Relative risks for the effects of late AP and high BMI on preclinical incidence rate were 1.98 (exp(0.6875)), and 2.59 (exp(0.9530)), respectively. (The corresponding figures on the transition from the PCDP to clinical phase were 1.56 (exp(0.4468)) and 1.99 (exp(0.6918)), respectively.)

Table IV. Estimated results for Weibull models with two covariates, breast cancer screening data from TAMCAS.

Parameter	Covariate	Estimate	SE
λ_1		0.0000061	0.00000044
k_1		2.2314	0.2030
λ_2		0.1497	0.1540
β_{11}	AP	0.6875	0.9117
β_{12}	BMI	0.9530	0.8710
β_{21}	AP	0.4468	1.0854
β_{22}	BMI	0.6918	1.0287

Goodness-of-fit $\chi^2 = 6.05055$, d.f. $= 10$, $p = 0.8110$.

Table V. Estimated results for piecewise exponential models with two covariates, breast cancer screening data from TAMCAS.

Parameter	Covariate	Estimate	SE
λ_1 ($\leqslant 50$ years)		0.0008	0.0010
λ_1' (>50 years)		0.0015	0.0018
λ_2		0.1075	0.1258
β_{11} ($\leqslant 50$ years)	AP	0.8616	1.0787
β_{12} ($\leqslant 50$ years)	BMI	1.4208	0.8565
β_{11}' (>50 years)	AP	1.4833	1.1185
β_{12}' (>50 years)	BMI	0.4939	0.9133
β_{21}	AP	0.8935	1.0950
β_{22}	BMI	0.7270	0.8710

Goodness-of-fit $\chi^2 = 10.6983$, d.f. $= 23$, $p = 0.9861$.

4.2. *Piecewise exponential models*

For the piecewise exponential models described in Section 3.2, estimates of preclinical incidence without covariates before 50 years old is 0.0037 (SE = 0.0014), which is approximate to that from the Weibull model ($\hat{q}_{01} = 0.0035$). Preclinical incidence after 50 years old is 0.0056 (SE = 0.0022), which is also close to the estimate from the Weibull model, 0.0057.

For covariate effects, we allow regression coefficients with respect to preclinical incidence, q_{01}, to be estimated by different age groups. The regression coefficients in respect of the transition rate from the PCDP to clinical phase, q_{12}, is assumed to remain constant across all ages because 75 per cent of women are under 50 years of age. Thus: for those younger than 50 years old, $q_{01} = \lambda_1 \times \exp(\beta_{11}X_1 + \beta_{12}X_2)$; for those older than 50, $q_{01}' = \lambda_1' \times \exp(\beta_{11}'X_1 + \beta_{12}'X_2)$. The results are listed in Table V. The preclinical incidence in the lowest risk group (AP negative and BMI negative) for an old woman is double that for a young woman. Estimated regression coefficients show AP plays a more important role in the onset of breast cancer for old women than for young women, but BMI seems rather significant for young women.

4.3. *Model diagnostics*

In order to check whether the fitted model is adequate we compared the observed number with the expected number by the use of Pearson χ^2 and deviance statistics. There is a lack of

significant difference of the comparison between the observed and expected for the piecewise exponential model ($\chi^2_{(23)} = 10.70,\ p = 0.99$). A good fit is also corroborated by the deviance statistics ($\chi^2_{(23)} = 12.34,\ p = 0.96$).

5. DISCUSSION

There are several contributions from this study. From the practical aspect, this is the first use, to the best of our knowledge, of non-homogeneous Markov models with the incorporation of Weibull distribution or the piecewise exponential model to accommodate non-constant transition rate and of proportional hazard regression form to take relevant covariates into account. The development of non-homogeneous Markov models enable fitting disease progression models in the presence of increased or decreased risk with time. The consideration of relevant covariates in multi-state transition models has implications for exploring the natural history of certain chronic diseases. For example, biological markers may have different roles at different stages. Elucidation of respective contributions from relevant factors is valuable for the prevention of chronic disease.

For the example of selective breast cancer screening illustrated in this paper, results found that the Weibull or piecewise exponential model may be appropriate for this data set. The incidence rate of preclinical breast cancer among female relatives of breast cancer index cases increases with age. The incidence rate (per 1000) for age groups under 40, 40–49, 50–59, 60–69 and 70–79 are estimated as 3.33, 4.63, 6.07 and 7.62. AP and BMI may not only significantly account for the occurrence of breast cancer but play an important role for the progression from the PCDP to the clinical phase. Taking AP and BMI into account, the incidence rates (per 1000) of the lowest risk group (AP negative and BMI negative), two intermediate risk groups (AP negative and BMI positive, AP positive and BMI negative), and the highest risk group for aged under 50 years are estimated as 0.80, 3.31, 1.89 and 7.84, respectively (Table V). The corresponding figures for aged over 50 years are 1.5, 6.61, 2.46 and 10.83, respectively (Table V). This means that the highest risk group (AP positive and BMI positive), AP negative and BMI positive, AP positive and BMI negative for aged under 50 have ten-fold, four-fold and two-fold risk, respectively, for the onset of breast cancer as compared to the lowest risk group (AP negative and BMI negative). A similar finding is observed for aged over 50 years.

Annual transition rates from the PCDP to clinical phase from Table V were estimated as 0.11, 0.27, 0.23 and 0.56 for the lowest risk group (AP negative and BMI negative), two intermediate risk groups (AP negative and BMI positive, AP positive and BMI negative), and the highest risk group, respectively. The inverse of these estimates give mean sojourn times for the four groups as 9.1 years, 3.7 years, 4.3 years and 1.8 years. This suggests that different screening strategies may take this information into account.

From a technical viewpoint, a computer program for the non-homogeneous Markov model using SAS PROC IML in this study can be used with relative ease and has great flexibility and potential to accommodate practical problems usually encountered in disease progression modelling. For example, in addition to the Weibull distribution, this model can be amended to use other distributions like gamma, log-normal, log-logistic regression, and so on, to accommodate different shapes of the time with respect to the corresponding transition.

Table VI. Estimated results for the Weibull model addressing the effect of covariate AP on each type of transition, adjusted for the effect of covariate BMI.

Parameter	Covariate	Estimate	SE
(*i*) *Only consider the effect of covariate AP on state 0 to state 1*			
λ_1		0.0000061	0.00000048
k_1		2.2606	0.2081
λ_2		0.1757	0.1905
β_{11}	AP	0.3351	0.2955
β_{12}	BMI	1.0863	0.9936
β_{22}	BMI	0.8642	1.2053
			-2 log -likelihood $= 557.6551$
(*ii*) *Only consider the effect of covariate AP on state 1 to state 2*			
λ_1		0.0000079	0.00000163
k_1		2.2114	0.4964
λ_2		0.1897	0.5181
β_{12}	BMI	1.2763	2.2062
β_{21}	AP	-0.4289	0.3894
β_{22}	BMI	1.0922	2.8795
			-2 log -likelihood $= 558.2998$

There are several ways of assessing the effect of a certain covariate on transition from state 0 to 1 and 1 to 2 using our non-homogeneous exponential regression Markov models. The effect of each covariate on both types of transitions can be assessed together by comparing the model including both transitions with the model excluding both transitions on the basis of likelihood ratio test with 2 d.f. as above. However, one of the strengths of our model is that the effect on each type of transition can be estimated mutually adjusted by two separate likelihood ratio tests with 1 d.f. (the comparison between the model including both transitions and the model only including the transition from state 0 to 1 or the transition from 1 to 2). Table VI shows the results for the model addressing the effect of AP on each type of transition adjusted for the covariate of BMI. However, no significant results were found for the effect on respective transitions, state 0 to state 1 ($\chi^2_{(1)} = 0.18,\ p = 0.68$) and state 1 to state 2 ($\chi^2_{(1)} = 0.82,\ p = 0.37$) although AP may play a more important role in the transition from state 0 to state 1 than that from state 1 to state 2. The similar findings were found for BMI with respect to the transitions state 0 to state 1 ($\chi^2_{(1)} = 0.52,\ p = 0.47$) and state 1 to state 2 ($\chi^2_{(1)} = 1.36,\ p = 0.24$). It is noted that the -2 log-likelihood of respective transitions, state 0 to state 1 or state 1 to state 2 for AP and BMI in Table III, are not significantly different from the 557.48 for both AP and BMI for both types of transitions. The results are consistent with the assessment of the effect of BMI or AP on state 0 to state 1 and state 1 to state 2 separately using the Wald test in Table IV. This suggests that it is the covariance between betas which make the effect of AP significant. This is proved by large covariance between β_{11} and β_{21} (0.94) using the model in Table IV.

Since no interval cases can be obtained, estimation of parameters in this study is only based on left-censored (prevalent cases) and interval-censored cases (incident cases) without interval cases (uncensored cases). Our model can be also applied to data with interval cases. The likelihood functions for the first screen and the second screen are identical to the above.

As regards interval cases, since time to interval cases is exactly known, the instantaneous rate was adopted to model interval cases. A detailed likelihood function is illustrated in Appendix C.

Estimation of parameters using expression (A4) was performed on the basis of a simulated data set that were spawned from 5000 women who received two rounds of screens with a three-year interscreening interval. For simplicity, the progression from normal to PCDP follows a Weibull distribution with true values of λ_1 and k_1 in expression (3), assigned as 0.0042 and 1, respectively, that is an exponential distribution, a special case of Weibull distribution. The progression from the PCDP to the clinical phase follows an exponential distribution with true value of λ_2, 0.3512. Simulated data based on these parameters yields 57 prevalent cases (first screen), 40 incident cases (second screen) and 25 interval cases. Estimation of parameters through expression (A4) yields the estimates of λ_1, k_1 and λ_2, 0.0036 (95 per cent CI: $-0.0070-0.0142$), 1.04 (95 per cent CI: 0.43–1.65) and 0.36 (95 per cent CI: 0.27–0.47), respectively, that are very close to true values.

There are several limitations in this study. First, owing to only eight cases being found in the second screen, this leads to imprecision of MLEs based on the asymptotic theory. Small samples may result in the violation of the assumption of asymptotic theory in the likelihood ratio test. This may affect the results for the significant effects of AP and BMI on transition from state 0 to state 1 on the basis of the likelihood ratio test.

Secondly, estimation of parameters in this study is only based on left-censored (prevalent cases) and interval-censored cases (incident cases) without interval cases (uncensored cases). Although previous research [9] has demonstrated that the estimates are consistent with those from data which includes interval cases, low precision of estimates will be encountered. Thirdly, we assume 100 per cent sensitivity in this study. It could be argued that no allowance made for sensitivity may lead to biased estimates of transition parameters. However, previous research [9] has also shown that simultaneous estimation of sojourn time and sensitivity yields 95 per cent sensitivity for the combinations of the three screening tools. Accordingly, 100 per cent sensitivity assumption is not unreasonable.

In conclusion, exponential regression Markov models with the incorporation of the Weibull distribution and the proportional hazard regression form were developed to model the effects of covariates on the natural history of chronic disease dispensing with the constant hazard assumption and with necessity for data on interval cases. In addition to breast cancer, our non-homogeneous Markov model can be easily applied to data on screening for other chronic diseases with or without interval cases.

APPENDIX A: DEFINITION OF TRANSITION INTENSITIES

Following Cox and Miller [12], a three-state Markov process is defined as follows. Let $Y(t)$, the disease state occupied by an individual at age t, be a random variable with state space $\Omega = \{0, 1, 2\}$, where

$$Y(t) = \begin{cases} 0 & \text{Normal} \\ 1 & \text{PCDP} \\ 2 & \text{Clinical phase} \end{cases}$$

The transition intensities for the three-state Markov process in Section 3 are defined as following:

$$q_{ij}(t) = \lim_{\mathrm{d}t \to 0} \frac{P\{Y(t+\mathrm{d}t) = j \mid Y(t) = i\}}{\mathrm{d}t} \quad \text{for } i, j = 0, 1, 2 \text{ and } i \neq j$$

$$q_{ii}(t) = -\sum_{i \neq j} q_{ij} \quad \text{for } i = 0, 1, 2$$

The transition probability $p_{ij}(t_1, t_2)$ is defined under the Markov property

$$\begin{aligned} p_{ij}(t_1, t_2) &= \Pr(Y(t_2) = j \mid Y(t_1) = i,\ Y(t),\ 0 \leqslant t < t_1) \\ &= \Pr(Y(t_2) = j \mid Y(t_1) = i) \end{aligned}$$

for $i, j = 0, 1, 2$ and $t_1 < t_2$.

APPENDIX B: DERIVATION OF TRANSITION PROBABILITIES

To derive the transition probabilities, we first define cumulative hazard functions as $H_j(t_1, t_2) = \int_{t_1}^{t_2} q_{ij}(t)\,\mathrm{d}t$, $i = 0, 1$, $j = 1, 2$, and $i < j$; then

$$p_{00}(t_1, t_2) = \exp[-H_1(t_1, t_2)] \tag{A1}$$

The probability of observing one in state 0 at age t_1 and 1 at age t_2 is obtained [13] by integrating the product function of the probability of being state 0 from age t_1 to t, the intensity of transition to state 1 at age t and the probability of remaining in state 1 from age t to age t_2. Thus

$$p_{01}(t_1, t_2) = \int_{t_1}^{t_2} \exp[-H_1(t_1, t)] q_{01}(t) \exp[-H_2(t, t_2)]\,\mathrm{d}t \tag{A2}$$

$$p_{02}(t_1, t_2) = 1 - p_{00}(t_1, t_2) - p_{01}(t_1, t_2) \tag{A3}$$

The integration in expression (A2) does not have a closed form and is rather intractable, the numerical integration [14, 15] is therefore used to estimate the parameters.

It should be noted that an exponential distribution is a special case of the Weibull distribution whereby the shape parameter k equals 1. Therefore, if any of the above two kinds of transition time follows an exponential distribution, expressions (A1) to (A3) are still applicable by substituting 1 for k in either expression (3) or (4).

APPENDIX C: LIKELIHOOD FUNCTION FOR DATA WITH INTERVAL CASES

Since time to the occurrence of interval cases is exactly known, the likelihood function for interval cancers should be based on the density function with respect to $p_{02}(t)$

$$p^*_{02}(u) = \frac{\mathrm{d}\,p_{02}(u)}{\mathrm{d}u}$$

where u is the time to the occurrence of interval cases. The total likelihood for data including first, second screens and interval cases will be written down as follows:

$$L(\theta) = L_1(\theta) \times L_2(\theta) \times L_3(\theta) \tag{A4}$$

where $L_1(\theta)$ and $L_2(\theta)$ are the likelihoods for the first screen and the second screen (see expressions (5) and (6)), respectively, $L_3(\theta)$ is the likelihood for the interval cases

$$L_3(\theta) = \prod_{k=1}^{n_3} \{p_{02}^*(u_k)\}$$

where n_3 is the number of interval cases.

ACKNOWLEDGEMENTS

We thank Dr Stephen Duffy for helpful discussions and comments.

REFERENCES

1. Walter SD, Day NE. Estimation of the duration of a pre-clinical disease state using screening data. *American Journal of Epidemiology* 1983; **118**:865–886.
2. *Report of the Organizing Committee and Collaborators, Falun Meeting*. Breast-cancer screening with mammography in women aged 40–49 years. *International Journal of Cancer* 1996; **68**:693–699.
3. Etzioni R, Cha R, Feuer EJ, Davidov O. Asymptomatic incidence and duration of prostate cancer. *American Journal of Epidemiology* 1998; **148**:775–785.
4. Zelen M, Feinleib M. On the theory of screening for chronic disease. *Biometrika* 1969; **56**:601–613.
5. Prorok PC. The theory of periodic screening. II. Doubly bounded recurrence times and mean lead time and detection probability estimation. *Advances in Applied Probability* 1976; **8**:460–476.
6. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedure from mass screening programmes. *Biometrics* 1984; **40**:1–14.
7. Chen HH, Duffy SW. A Markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. *The Statistician* 1996; **45**(3):307–317.
8. Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov Chain Model of both entry to and exit from the preclinical detectable phase. *Statistics in Medicine* 1995; **14**:1531–1543.
9. Chen THH, Kuo HS, Yen MF, Lai MS, Tabar L, Duffy SW. Estimation of sojourn time in chronic disease screening without data on interval cases. *Biometrics* 2000; **56**:167–172.
10. Armitage P, Doll R. Stochastic models for carcinogenesis. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* 1961; **4**:9–38.
11. Lai MS, Yen MF, Kuo HS, Kong SL, Chen THH, Duffy SW. Efficacy of breast-cancer screening for female relatives of breast-cancer-index cases: Taiwan multicentre cancer screening (TAMCAS). *International Journal of Cancer* 1998; **78**:21–26.
12. Cox DR, Miller HD. *The Theory of Stochastic Processes*. Chapman & Hall: London, 1965.
13. Klein JP, Klotz JH and Grever MR. A biological marker model for predicting disease transitions. *Biometrics* 1984; **40**:927–936.
14. Stenger F. Integration formulae based on trapezoidal formula. *Journal of the Institute of Mathematics and its Application* 1973; **12**:103–114.
15. SAS/IML software, changes and enhancements through release 6.11. SAS Institute: 1995.

