

**A General Multilevel Multistate Competing Risks Model for Event History Data, with  
an Application to a Study of Contraceptive Use Dynamics**

*Fiona Steele, Harvey Goldstein\**

Centre for Multilevel Modelling

Institute of Education

University of London

20 Bedford Way

London WC1H 0AL

*William Browne\**

Mathematical Sciences

University of Nottingham

University Park

Nottingham NG7 2RD

Contact author: Fiona Steele

Tel: 020 7612 6657

Fax: 020 7612 6658

Email: [F.Steele@ioe.ac.uk](mailto:F.Steele@ioe.ac.uk)

\*Fiona Steele is Research Lecturer in Statistics, Harvey Goldstein is Professor of Statistical Methods, and William Browne is Lecturer in Statistics.

## **Abstract**

We propose a general discrete-time model for multilevel event history data. The model is developed for the analysis of longitudinal repeated episodes within individuals where there are multiple origin states and multiple transitions from a state (competing risks). Transitions from each origin state are modelled jointly to allow for correlation across states in the unobserved individual characteristics that influence transitions. Implementation of the method in *MLwiN* is described. The model is applied in an analysis of contraceptive use dynamics in Indonesia where transitions from two origin states, contraceptive use and non-use, are of interest. A distinction is made between two ways in which an episode of contraceptive use may end: a transition to non-use or a switch to another method. After adjusting for a range of background characteristics, we find evidence of a positive residual correlation between the risk of discontinuation and the risk of moving from non-use to use; this suggests that women who have short (long) episodes of contraceptive use tend also to have short (long) episodes of non-use.

*Keywords:* Event history analysis, competing risks, multilevel model, multistate model, contraceptive use

## 1. Introduction

Event history data are collected in many surveys, providing a longitudinal record of events such as births, deaths, and changes in employment and marital status. These data are often highly complex, with common features including repeated events, multiple origin states and multiple types of transition from each state (competing risks). While there are methods for handling repeated events combined with either multiple origin states or competing risks, existing methodology does not allow all three features to be handled simultaneously. In this paper, we propose a general event history model for the analysis of repeated durations where there may be multiple origin states and multiple transitions from those states.

The methodological development is motivated by a study of contraceptive use dynamics. Event history data on episodes of contraceptive use and non-use are now collected in a number of developing countries, as part of the Demographic and Health Survey (DHS) programme. These surveys collect monthly data on contraceptive use, non-use and pregnancy for a period of 5-6 years before the survey date. The information recorded includes the methods of contraception used and the reason for discontinuation when an episode of use ends. Previous studies of contraceptive use dynamics using these data have focused on contraceptive discontinuation, allowing for repeated episodes of use and different reasons for discontinuation in a competing risks framework (e.g. Steele et al. 1996b). Episodes of non-use are ignored. However, the transition from non-use to use is also important for family planning programme evaluation since women who do not quickly resume contraceptive use after a birth, or after discontinuing use of a method, may be at risk of having an unintended pregnancy. In this paper, we consider episodes of both contraceptive use and non-use and model transitions between use and non-use simultaneously. Use and non-use of contraception

are examples of multiple origin states. By jointly modelling transitions from different origin states, it is possible to test explicitly for state-dependent covariate effects. For example, the effects of background characteristics such as age might differ for transitions from use and non-use. Joint modelling of transitions also allows for residual correlation in individual transition rates across states, which might arise because of unobserved factors that affect transitions from each state.

In the model for contraceptive discontinuation, an episode of contraceptive use is defined as a continuous period of using the same method. We distinguish between two types of transition from use of a given method: a transition to non-use or a switch to a different method (a transition within the ‘use’ state). These two types of event are examples of competing risks. An episode of non-use is defined as a continuous period of non-use (excluding months of pregnancy). The only possible type of transition from non-use is to use. We therefore have a situation where the number and type of transitions that can occur depend on the origin state. The model developed in this paper can handle state-dependent competing risks.

The model we propose is a generalised multilevel discrete-time event history model.

A multilevel model is used to allow for the hierarchical structure that arises from having repeated episodes (of use or non-use of contraception) nested within individuals. The model includes individual random effects for each origin state and for each type of transition from a given state; these random effects may be correlated across origin states and transitions to allow for shared unobserved individual factors. One advantage of choosing a discrete-time formulation is that it allows the model to be cast as a multilevel model for multinomial response data, which may be fitted using existing software.

The remaining sections of the paper are organised as follows. In Section 2, we give a brief outline of previous work on event history analysis for repeated events, competing risks and multiple states. We then describe a general multilevel multistate competing risks model that allows all three of these common features of event history data to be incorporated simultaneously. The application of this model to a study of contraceptive use dynamics in Indonesia is presented in Section 3. Finally, in Section 4, further extensions to the proposed model are discussed.

## **2. Methodology**

### **2.1 Previous work on repeated events, multiple origin states and competing risks**

When an event may occur more than once over an individual's lifetime, the durations between events may be correlated due to the presence of unobserved individual-level factors. Repeated events are usually handled by including individual-specific random effects in an event history model, leading to a multilevel model. The random effect represents individual 'frailty', and the random effect variance measures unobserved heterogeneity among individuals. Vaupel et al. (1979) describe how, in the presence of unobserved heterogeneity, the population hazard rate may be observed to decline over time, even if the hazard rates of individuals in the population are constant throughout the observation period, due to high risk individuals experiencing the event early and leaving the least susceptible individuals in the 'at risk' sample. Repeated observations on individuals allow unobserved heterogeneity to be better identified. Multilevel event history models have been developed for the analysis of hierarchical duration data, where the hierarchical structure results from repeated events

within individuals or clustering of individuals within some higher-level grouping such as geographical area. Multilevel extensions of continuous-time proportional hazards models include Clayton and Cuzick (1985), Goldstein (2003, Chapter 10), Guo and Rodríguez (1992) and Sastry (1997), while discrete-time approaches include Davies et al. (1992) and Steele et al. (1996a).

Another extension of the basic event history model allows for the possibility of multiple states. There may be several transient origin states between which individuals move, perhaps more than once. Several previous studies have considered models for repeated transitions between multiple states. Enberg et al. (1990) consider transitions between welfare and work using a random effects model but, since they assume individual random effects are uncorrelated across states, their approach amounts to fitting a separate model for each origin state. In many applications, this assumption of independence may be invalid since there may be unobserved factors which influence transitions from more than one origin state. Goldstein et al. (2002) also use a discrete-time random effects model, but model jointly transitions from two origin states, allowing for correlation between the state-specific random effects. An alternative approach is to use a fixed effects model such as the proportional hazards model proposed by Lindeboom and Kerkhofs (2001) to analyse movements between sickness and work spells, clustered by workplace. While existing methods allow for multiple origin states and repeated events, it is assumed that only one type of transition can occur from each state.

Competing risks are another common feature of event history data. In many situations, there are several competing destinations from a given state, or an event may be experienced for one of several reasons. To allow for unobserved individual heterogeneity in the risks of competing events, various competing risks event history models have been proposed. These

models typically include individual-level random effects for each alternative destination. Enberg et al. (1990) consider a discrete-time competing risks model with individual- and destination-specific random effects, which is essentially a multilevel multinomial logit model. However, their model assumes that the random effects are uncorrelated across competing risks, an assumption which is likely to be unrealistic since there may be common unobservables affecting more than one type of transition. Hill et al. (1993) propose a nested logit model which relaxes this independence assumption. For alternative destinations which may be regarded as similar with respect to unmeasured risk factors, the error terms are decomposed into a component which is common to similar alternatives and a component which is destination-specific. A different approach to relaxing the independence assumption is adopted by Steele et al. (1996b). They propose a discrete-time competing risks model, formulated as a multilevel multinomial model, which includes individual- and destination-specific random effects that may be correlated across destinations. Theirs is a more general model than that of Hill et al. (1993) and can be extended to several hierarchical levels where the effects of duration and covariates may vary across higher-level units. However, neither approach allows for the possibility of multiple origin states.

## **2.2 Multilevel discrete-time competing risks model**

In this section we describe the discrete-time competing risks model proposed by Steele et al. (1996b). The more general model proposed in the present paper is an extension of this model that allows for multiple origin states.

We focus on discrete-time models for several reasons. First, in our application durations of use and non-use of contraception are measured in discrete-time units as the data were

collected monthly. It is very common for durations to be measured in discrete time, particularly in studies of human populations in which event times are often collected retrospectively. When durations are recorded in reasonably broad intervals, such as months, there will be multiple ties. While ties present no problem in the estimation of discrete-time models, some adjustment is required if a continuous-time model is used. For example the widely used Cox proportional hazards model, estimated via partial likelihood, requires some modification (see, e.g., Kalbfleish and Prentice, 1980). A second reason for favouring discrete-time event history models is that they are essentially discrete response models. This allows the use of existing methodology for multilevel discrete response data when there are repeated events. Other benefits of the discrete-time approach include straightforward inclusion of time-varying covariates and the possibility to allow for non-proportional hazards. Non-proportional hazards are handled by including interactions between the duration variable(s) (treated as explanatory variable(s) in a discrete-time model) and covariates.

One disadvantage of a discrete-time approach, however, is the need to expand the dataset so that there is an observation for each time unit. If the width of the discrete time intervals is short relative to the observation period this may lead to a very large dataset, but with increasing computational power and storage this is becoming a less severe problem. One strategy to reduce the size of the expanded dataset is to group discrete-time intervals; for example, quarterly rather than monthly observations might be created. While grouping intervals leads to a loss of information, in our experience there is often little impact on parameter estimates and standard errors. In the application to contraceptive use dynamics, for example, results were robust to using six-month rather than monthly intervals.

An episode is defined as a continuous period of time spent in the same state until an event occurs. Suppose that for each time interval  $t$  in episode  $j$  for individual  $k$ , we observe a multinomial variable  $y_{tjk}$  which denotes whether an event has occurred and the type of event. Suppose there are  $R$  end events. Denote the multinomial response by  $y_{tjk}$  where  $y_{tjk} = r$  if an event of type  $r$  has occurred in time interval  $t$ ,  $r = 1, \dots, R$ , and  $y_{tjk} = 0$  if no event has occurred. The hazard of an event of type  $r$  in interval  $t$ , denoted by  $h_{tjk}^{(r)}$ , is the probability that an event of type  $r$  occurs in interval  $t$ , given that no event of any type has occurred before interval  $t$ .

The log-odds of an event of type  $r$  versus no event may be modelled as a function of episode duration and covariates, using methods for unordered multinomial response data. Using a logit link, the multilevel discrete-time competing risks model may be written

$$\log\left(\frac{h_{tjk}^{(r)}}{h_{tjk}^{(0)}}\right) = \boldsymbol{\alpha}^{(r)T} \mathbf{z}_t^{(r)} + \boldsymbol{\beta}^{(r)T} \mathbf{x}_{tjk}^{(r)} + u_k^{(r)}, \quad r = 1, \dots, R. \quad (1)$$

The effect of duration is represented by  $\boldsymbol{\alpha}^{(r)T} \mathbf{z}_t^{(r)}$  which can take a number of forms, including a polynomial function or a step (piecewise constant) function of time. The covariates, represented by  $\mathbf{x}_{tjk}^{(r)}$ , may be defined at the level of the discrete time unit (time-dependent), or at the episode or individual level. Equation (1) defines a proportional hazards model where the effects of covariates are assumed to be constant across time. Non-proportional effects may be accommodated simply by adding interactions between  $\mathbf{z}_t^{(r)}$  and  $\mathbf{x}_{tjk}^{(r)}$ .

In a competing risks model, the effects of duration and covariates may differ for each event type, as indicated by the  $r$  superscript for  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$ . It is also possible that the form of  $\mathbf{z}_t$  and

the set of covariates  $\mathbf{x}_{ijk}$  may vary across event types. Unobserved individual-specific factors may differ for each type of event; these are represented by  $R$  random effects  $u_k^{(r)}$ . The random effects are assumed to follow a multivariate normal distribution, with covariance matrix  $\mathbf{\Omega}_u$ ; non-zero correlation between random effects allows for shared or correlated unobserved risk factors across competing risks. The model may be extended further to allow coefficients of  $\mathbf{z}_i^{(r)}$  and  $\mathbf{x}_{ijk}^{(r)}$  to vary randomly across individuals.

Model (1) may be estimated as a multilevel multinomial model (Goldstein, 2003, Chapter 4). Several software packages may be used, including MLwiN (Rasbash et al., 2000), PROC NLMIXED in SAS (SAS Institute, 1999) and WinBUGS (Spiegelhalter et al., 2000). Further details of the multinomial model for competing risks are given in Steele et al. (1996b).

### 2.3 A multilevel discrete-time model for competing risks and multiple states

The model we propose is an extension of (1) to handle situations where there are both competing risks and multiple origin states. The approaches of Steele et al. (1996b) and Goldstein et al. (2002) are combined in a general framework. In this general model, the number and type of transitions may differ for each state. It is also possible that the end of an episode does not necessarily lead to a change in state. For example, in our application an episode of contraceptive use may end in a transition to the non-use state or a transition within the use state (a method switch).

Suppose that there are  $R_i$  ways in which an episode in state  $i$  ( $i = 1, \dots, s$ ) can end. Denote by  $h_{ijk}^{(r_i)}$  the hazard of making a transition of type  $r_i$  ( $r_i = 1, \dots, R_i$ ) from origin state  $i$  in time

interval  $t$  of episode  $j$  for individual  $k$ . The hazard of no transition is denoted by  $h_{ijk}^{(0)}$ . A

multilevel model for competing risks and multiple states may be written

$$\log\left(\frac{h_{ijk}^{(r_i)}}{h_{ijk}^{(0)}}\right) = \boldsymbol{\alpha}_i^{(r_i)T} \mathbf{z}_{ii}^{(r_i)} + \boldsymbol{\beta}_i^{(r_i)T} \mathbf{x}_{ijk}^{(r_i)} + u_{ik}^{(r_i)}, \quad r_i = 1, \dots, R_i; \quad i = 1, \dots, s. \quad (2)$$

In (2) duration and covariate effects may depend both on the origin state  $i$  and on the type of

transition  $r_i$ . Unobserved individual-level factors, represented by  $u_{ik}^{(r_i)}$ , may also vary

according to state and transition. The  $\sum_{i=1}^s r_i$  random effects are assumed to follow a

multivariate normal distribution.

## 2.4 Data preparation and estimation

In order to estimate a discrete-time event history model, the data must first be restructured to what is often called a person-period format. This involves expanding the data so that there is a record for each time interval in each episode. For example, an episode which ended during the third time interval would be expanded to obtain three records, for  $t = 0$ ,  $t = 1$  and  $t = 2$ .

Suppose there are competing risks and the episode ended for reason  $r = 2$ , then the multinomial response variable for the three intervals would be  $(y_{0,jk}, y_{1,jk}, y_{2,jk}) = (0, 0, 2)$ . If the individual had been right-censored during the third time interval, their sequence of responses would be  $(0, 0, 0)$ . After this data expansion, models (1) or (2) may be estimated using any software that can handle multilevel multinomial response data. Further details of the data structure required are given in Appendix A.

In the analysis that follows we have used a hybrid Gibbs-Metropolis sampling algorithm.

Gibbs sampling is used to update the random effects variance matrix, while single-site

random walk Metropolis sampling is used for all the other parameters. As we have no prior information on likely parameter values we have incorporated suitable ‘diffuse’ prior distributions in the model. Details of the MCMC estimation algorithm and the chosen prior distributions are given in Appendix B. This method has been implemented in *MLwiN*. Details of *MLwiN*'s MCMC estimation engine are given in Browne (2002).

### **3. Contraceptive use dynamics in Indonesia**

We consider an application of the multilevel multistate competing risks model in an analysis of changes in contraceptive use over time. Two origin states are considered: contraceptive use and non-use. An episode of non-use always ends in a transition to use, while for an episode of contraceptive use there are two competing risks: a woman may discontinue use of all contraception and become a non-user, or she may switch to a different method.

#### **3.1 Data and sample definition**

The data are from the 1997 Indonesia Demographic and Health Survey (IDHS), a nationally representative survey of ever-married women age 15-49 (Central Bureau of Statistics, 1998). Contraceptive histories were collected retrospectively using a calendar for a six-year period before the survey. The calendar has a tabular format with a row for each month of the observation period, a column containing information on pregnancies, births and contraceptive use, and another column recording the main reason for discontinuation for each episode of contraceptive use. The analysis is based on episodes of contraceptive use and non-use for 14677 women who were married throughout the observation period and who had previously used contraception.

An episode is defined as a continuous period of non-use or use of the same contraceptive method. Periods of non-use that are interrupted by pregnancy are treated as two separate episodes, one ending when the woman becomes pregnant, and the other starting after the birth. Periods of non-use while a woman is pregnant are excluded. The period of non-use after pregnancy is considered as a new episode since interest is focused on non-use while a woman is at risk of conception. Episodes of male or female sterilisation are excluded from the analysis since no transition is possible from these permanent methods of contraception. This results in the loss of a very small number of episodes since sterilisation is relatively unpopular in Indonesia, and few women in the sample were sterilised after the start of the six-year observation period. The sample is further restricted to women who had previously used contraceptives, and to episodes of use or non-use which began after the start of the observation period. Episodes that were in progress at the start of the calendar period, i.e. left-truncated episodes, were necessarily excluded since the start date was not asked for these episodes. The final analysis sample contains 17 843 episodes of use and 21 285 episodes of non-use.

The IDHS also collected complete birth histories and a large amount of demographic and socio-economic information from each woman and her household. A number of covariates were used in the analysis: current age (treated as time-dependent), education level, type of region of residence, an indicator of socio-economic status based on household possessions, contraceptive method (for episodes of contraceptive use) and an indicator of whether the episode followed a live birth (for episodes of non-use). The socio-economic status indicator has been used in previous studies (Curtis and Blanc, 1997; Steele and Curtis, 2003) and is based on a simple household possessions score. Households receive one point for having each

of the following: piped or bottled drinking water, flush toilet, vehicle, radio, and a floor that is not dirt. The total score ranges from 0 to 5 and is categorised as low (0-1), medium (2-3), or high (4-5). Contraceptive method is classified as 1) pills or injectables (short-term hormonal methods), 2) Norplant® or intra-uterine device (IUD) (longer-term clinical methods), 3) other modern reversible methods (mainly condoms), and 4) traditional methods. Descriptive statistics for all covariates are given in Table 1.

### 3.2 Modeling strategy

The multilevel multistate competing risks model in (2) is applied in the analysis of transitions from  $s=2$  states, contraceptive use and non-use. From the ‘use’ state ( $i=1$ ) there are  $R_1=2$  possible transitions, while from the ‘non-use’ state ( $i=2$ ) there is only  $R_2=1$  transition.

In order to fit a discrete-time event history model, the data first must be expanded so that there is a response for each time interval in an episode. The expanded dataset using one-month intervals has 543 737 observations. To reduce computational time, the length of discrete-time intervals is increased to six months which reduces the size of the dataset to 109 666 observations. Comparison of single-level models using one- and six-month intervals reveal that increasing the length of discrete-time intervals to six months has little effect on the parameter estimates or standard errors (results not shown). In aggregating time intervals, the number of episodes does not change. If there is more than one episode within a six-month interval, all such episodes are retained in the reduced dataset, with a duration of one six-month interval recorded for each.

Duration effects are modelled in different ways for use and non-use states. For transitions from contraceptive use, a piecewise constant formulation is found to be a good fit to the observed logit-hazard. A step function is fitted for duration intervals of 0-5 months, 6-11 months, 12-23 months, 24-35 months, and 36 or more months. For transitions from non-use to use, a polynomial function of the cumulative duration of non-use is used.

### **3.3 Results**

Cumulative transition probabilities were calculated using separate life tables for each origin state. Based on a multiple-decrement life table, within the first 12 months of use 13% of women have become non-users and 13% have switched to a different method of contraception. After 24 months, 23% have discontinued while 18% have switched methods. The probability of moving from non-use to use increases rapidly with duration of non-use. Within 12 months of the start of an episode of non-use, 57% of women have started to use contraception, while 70% start within 24 months. These high rates are due largely to women resuming contraceptive use after a brief period of non-use following a birth.

#### **3.3.1 Random effects**

We begin by fitting a model including duration effects only, before adding the covariates listed in Table 1. The estimated random effects covariance matrix from both models is shown in Table 2. There is evidence of unobserved heterogeneity between women in the hazards of all types of transition, but particularly for transitions from contraceptive use. From the upper panel of Table 2, it can be seen that before including covariates there is a strong negative residual correlation (estimated as -0.71) between the logit-hazards for the transition from use to non-use and from non-use to use. The negative correlation implies that women with a high

(low) hazard of moving from non-use to use tend to have a low (high) hazard of discontinuation. In other words, women with short (long) periods of use before a discontinuation generally have long (short) periods of non-use. On further examination of the data, we find that the shortest periods of non-use follow a live birth. These short postnatal episodes of non-use are usually followed by a long period of using the same method of contraception, in order to space or limit subsequent births. After controlling for covariates, in particular the indicator of whether a period of non-use immediately followed a live birth, we find that the residual correlation becomes moderate and positive (see the estimate of 0.28 in the lower panel of Table 2). A positive correlation implies that women with short periods of contraceptive use tend also to have short periods of non-use, and those who use contraceptives for longer periods tend to have longer breaks in use. The correlations between the random effects for the other pairs of transitions are both small and neither is significant at the 5% level.

### **3.3.2 Fixed effects**

The estimated coefficients and standard errors corresponding to the fixed part of the full model are shown in Table 3. For all types of transition, the effects of current age, education level, type of region of residence, and household socio-economic status are considered. For transitions from use, the type of contraceptive method used is treated as a time-dependent covariate. The indicator of whether an episode of non-use follows a live birth is included only in the model for the transition from non-use.

We begin by examining the effect of duration of use and covariates on transitions from contraceptive use to non-use (‘discontinuation’) or to use of another method (‘switching’). The risk of discontinuation is fairly constant over the first three years of use, but greater for

longer durations, while the risk of switching is highest in the first six months of use, then decreases. Age has a negative effect on both discontinuation and switching; older women are more likely than young women to continue use of the same method. Education has a positive effect on both discontinuation and switching, but the effect on the risk of switching is stronger. Urban women are more likely than rural women to discontinue, but type of region has no effect on the rate of switching. Socio-economic status has different effects on discontinuation and switching; a high level of socio-economic status is associated with low discontinuation rates, but higher switching rates, possibly reflecting access to a wider choice of methods for better-off women. Norplant®/IUD users are less likely than users of any other method to become non-users or to change to a different method. Users of traditional methods are also relatively unlikely to switch methods. In contrast, condom users (the main constituent of the ‘other modern’ group) are the most likely to abandon contraceptive use or to change to another method.

We now turn to the factors associated with transitions from non-use to use. The probability that a non-user becomes a user decreases sharply with the duration of non-use. Older women are less likely than young women to become a contraceptive user. Educated women, those living in urban areas, and women of higher socio-economic status are more likely than uneducated, rural, or poorer women to make the transition from non-use. Finally, if the episode of non-use follows a birth rather than an episode of contraceptive use, a woman is considerably more likely to adopt contraception and the negative effect of duration of non-use is stronger. This effect distinguishes between short breaks in contraceptive use after a birth and longer-term non-use, possibly following a problem with contraception such as side-effects.

#### 4. Discussion

We have shown how to specify and fit general discrete-time event history models with multiple origin states and multiple transitions from those states. We have illustrated this for repeated episodes within individuals but our models can be extended readily to further levels of nesting. For example, in the application presented here, community-specific random effects may be added to allow for clustering of contraceptive behaviour within neighbourhoods or villages.

We have assumed that random effects follow a multivariate normal distribution. This leads to an extremely flexible model in which there may be several correlated random effects. As with any statistical analysis, however, it is important to carry out diagnostic checks for departures from normality and other model assumptions. Langford and Lewis (1998) propose a range of procedures for multilevel data exploration, including methods for detecting and adjusting for outliers. It may also be possible to protect against non-normality using ‘sandwich’ or robust standard errors (Goldstein, 2003, p.80-81). Another approach is to assume a non-normal random effects distribution, for example a multivariate t-distribution, which could be implemented in WinBUGS (Spiegelhalter et al. 2000).

We have ignored the possibility of within-individual between-episode random variation in durations. In principle we can fit this using episode-specific random effects, but in this case the within-individual variation in episode durations is not significant, possibly due to a relatively low proportion of women who experience more than one transition of each type. Furthermore, in general it would seem preferable to model episode heterogeneity using random coefficients associated with individual level covariates. Thus, for example, the age

relationship within individuals may vary across individuals and this can be modelled by including random coefficients for the age group category coefficients.

Where the transition states form an ordered categorisation we can use corresponding ordered category models, for example by modelling cumulative log-odds (Goldstein, 2003, Chapter 4). This could arise, for example, in the modelling of illness duration where patients make transitions between clinical states which are ordered by severity. For such models we can also assume an underlying propensity with a probit link and this can be fitted via MCMC.

Our models can also be extended readily to the multivariate case where, for each individual, we wish to study more than one type of episode at a time; for example, durations of contraceptive use episodes and intervals between births. For each episode type we form the same set of discrete time intervals and, for each time interval, the response is multivariate with dimension  $p$ , where  $p$  is the number of episode types. A dummy variable is created to indicate each episode type, and these are interacted with covariates to allow covariate effects to vary across the different types of episode. For ordered models and for binary response models, using a probit link, we can directly incorporate correlations between the underlying normal distributions at the episode level and at higher levels. This then provides covariance matrix estimates for the episode types at all levels of the data hierarchy.

## References

- Browne, WJ and Draper D. (2000) Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. *Computational Statistics* **15**, 391-420.
- Browne, W. J. (2002). *MCMC estimation in MLwiN*. London, Institute of Education.
- Central Bureau of Statistics (CBS) [Indonesia], State Ministry of Population/National Family Planning Coordination Board (NFPCB), Ministry of Health (MOH), and Macro International Inc. (MI) (1998) *Indonesia Demographic and Health Survey 1997*. Calverton, Maryland: CBS and MI.
- Clayton, D.G. and Cuzick, J. (1985) Multivariate generalizations of the proportional hazards model (with discussion). *Journal of the Royal Statistical Society, Series A*, **148**, 82-117.
- Curtis, S.L. and Blanc, A. (1997) *Determinants of contraceptive failure, switching, and discontinuation: An analysis of DHS contraceptive histories*. DHS Analytical Reports No. 6, Calverton, Maryland: Macro International Inc.
- Davies, R.B., Elias, P. and Penn, R. (1992) The relationship between a husband's unemployment and his wife's participation in the labour-force. *Oxford Bulletin of Economics and Statistics*, **54**, 145-171.
- Enberg, J. Gottschalk, P. and Wolf, D. (1990) A random-effects logit model of work-welfare transitions. *Journal of Econometrics*, **43**, 63-75.
- Goldstein, H. (2003) *Multilevel Statistical Models*. 3<sup>rd</sup> edition. London: Arnold.
- Goldstein, H., Pan, H. and Bynner, J. (2002) A note on methodology for analysing longitudinal event histories using repeated partnership data from the National Child Development Study (NCDS). Working paper, Institute of Education, London. Downloadable from <http://k1.ioe.ac.uk/hgpersonal/>.

- Guo, G. and Rodríguez, G. (1992) Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala. *Journal of the American Statistical Association*, **87**, 969-976.
- Hill, D.H., Axinn, W.G. and Thornton, A. (1993) Competing hazards with shared unmeasured risk factors. *Sociological Methodology*, **23**, 245-277.
- Kalbfleish, J.D. and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data*. New York: Wiley.
- Langford, I.H. and Lewis, T. (1998) Outliers in multilevel data (with discussion). *Journal of the Royal Statistical Society, Series A*, **161**, 121-160.
- Lindeboom, M. and Kerkhofs, M. (2000) Multistate models for clustered duration data – an application to workplace effects on individual sickness absenteeism. *The Review of Economics and Statistics*, **82**, 668-684.
- Rasbash, J., Browne, W.J., Goldstein, H., Yang, M., et al. (2000). *A User's Guide to MLwiN (Second Edition)*. London: Institute of Education.
- SAS Institute Inc. (1999) SAS/STAT<sup>®</sup> *User's Guide, Version 8*. Cary, NC: SAS Institute inc.
- Sastry, N. (1997) A nested frailty model for survival data, with an application to the study of child survival in Northeast Brazil. *Journal of the American Statistical Association*, **92**, 426-435.
- Spiegelhalter, D.J., Thomas, A. and Best, N.G. (2000) *WinBUGS Version 1.3 User Manual*. Cambridge: Medical Research Council Biostatistics Unit.
- Steele, F. and Curtis, S.L. (2003) Appropriate methods for analysing the effect of method choice on contraceptive discontinuation. *Demography*, **40**, 1-22.
- Steele, F., Diamond, I. and Amin, S. (1996a) Immunization uptake in rural Bangladesh: a multilevel analysis. *Journal of the Royal Statistical Society, Series A*, **159**, 289-299.

Steele, F., Diamond, I. and Wang, D. (1996b) The determinants of the duration of contraceptive use in China: A multilevel multinomial discrete hazards modelling approach. *Demography*, **33**,12-33.

Vaupel, J.W., Manton, K.G. and Stallard, E. (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16**, 439-454.

**Table 1.** Distribution of women/episodes by covariates, Indonesia 1997

	Number of women	% of women
<i>Woman-level variables</i>		
Education		
None	1257	8.6
Primary	7643	52.1
Secondary +	5777	39.4
Type of region of residence		
Rural	10393	70.8
Urban	4284	29.2
Socio-economic status		
0-1 (low)	2418	16.5
2-3 (medium)	7261	49.5
4-5 (high)	4998	34.0
	Number of episodes	% of episodes
<i>Episode-level variables: non-use</i>		
Episode follows a live birth		
No	7677	36.1
Yes	13608	63.9
Age (years)		
<25	7748	36.4
25-34	10323	48.5
35-49	3214	15.1
<i>Episode-level variables: use</i>		
Contraceptive method		
Pill/injectable	13736	77.0
Norplant®/IUD	2740	15.4
Other modern	323	1.8
Traditional	1044	5.9
Age (years)		
<25	6774	38.0
25-34	8468	47.5
35-49	2601	14.6

**Table 2.** Random effects covariance matrix from models of transitions from contraceptive use and non-use, Indonesia 1992-97

	Use → non-use (Discontinuation)		Use → other method (Method switch)		Non-use → use	
	Est. †	(95% interval estimate)	Est.	(95% interval estimate)	Est. †	(95% interval estimate)
<i>Duration effects only</i>						
Use → non-use	0.524	(0.398,0.659)				
Use → other method	0.055	(-0.032,0.154)	0.748	(0.663,0.865)		
	<b>0.092<sup>a</sup></b>					
Non-use → use	-0.155	(-0.204,-0.112)	0.002	(-0.045,0.042)	0.089	(0.071,0.116)
	<b>-0.710<sup>a</sup></b>		<b>0.001<sup>a</sup></b>			
<i>Duration + covariates</i>						
Use → non-use	0.223	(0.149,0.309)				
Use → other method	0.021	(-0.065,0.114)	0.702	(0.616,0.818)		
	<b>0.058<sup>a</sup></b>					
Non-use → use	0.068	(0.015,0.109)	0.016	(-0.090,0.093)	0.231	(0.172,0.305)
	<b>0.278<sup>a</sup></b>		<b>0.016<sup>a</sup></b>			

†Coefficients are the modal estimates from 50 000 chains.

<sup>a</sup>Correlation between random effects.

**Table 3.** Estimated coefficients and standard errors from model of transitions from contraceptive use and non-use, Indonesia 1992-97

	Use → non-use (Discontinuation)		Use → another method (Method switch)		Non-use → use	
	Est.†	(SE)	Est.†	(SE)	Est.†	(SE)
Constant	-2.306	(0.085)	-3.424	(0.123)	-1.906	(0.106)
Duration (months)						
0-5	0	-	0	-	-	-
6-11	-0.031	(0.048)	-0.345	(0.052)	-	-
12-23	-0.064	(0.046)	-0.586	(0.054)	-	-
24-35	0.076	(0.057)	-0.568	(0.070)	-	-
36+	0.160	(0.066)	-0.440	(0.082)	-	-
Duration	-	-	-	-	-0.752	(0.071)
Duration <sup>2</sup>	-	-	-	-	0.051	(0.008)
Age (years)						
<25	0	-	0	-	0	-
25-34	-0.408	(0.038)	-0.374	(0.047)	-0.319	(0.031)
35-49	-0.844	(0.057)	-0.593	(0.068)	-0.710	(0.045)
Education						
None	0	-	0	-	0	-
Primary	0.014	(0.074)	0.424	(0.107)	0.210	(0.053)
Secondary+	0.234	(0.077)	0.831	(0.110)	0.508	(0.057)
Region of residence						
Rural	0	-	0	-	0	-
Urban	0.135	(0.040)	0.068	(0.049)	0.138	(0.035)
Socio-economic status						
0-1 (low)	0	-	0	-	0	-
2-3 (medium)	-0.107	(0.050)	0.354	(0.072)	0.187	(0.040)
4-5 (high)	-0.167	(0.057)	0.307	(0.078)	0.384	(0.047)
Contraceptive method						
Pill/injectable	0	-	0	-	-	-
Norplant®/IUD	-1.244	(0.061)	-1.106	(0.071)	-	-
Other modern	0.521	(0.120)	0.508	(0.137)	-	-
Traditional	0.024	(0.068)	-0.623	(0.102)	-	-
Episode follows live birth					3.030	(0.098)
Episode after live birth*Duration	-	-	-	-	-0.312	(0.077)
Episode after live birth*Duration <sup>2</sup>	-	-	-	-	0.024	(0.009)

†Estimates are the modal estimates from 50 000 chains.

## Appendix A: Data Preparation

Suppose that a woman uses contraception for 3 time intervals, then discontinues and does not use contraception for 2 intervals, then uses contraception again for 4 intervals, before switching to another method:

Individual ( $k$ )	Episode ( $j$ )	State ( $i$ )	Duration (in 6-month intervals)	Transition	Censor
1	1	1	3	Discontinue	0
1	2	2	2	Start to use	0
1	3	1	4	Switch	0

*Censor* indicates whether the episode is right-censored; here, the duration of each episode is completely observed.

The first step in restructuring the data for a multinomial discrete-time model is to create a multinomial response for each time interval (six months intervals, here). The multinomial response  $y_{tijk}$  has  $R_i$  categories for state  $i$ , where  $R_i = 2$  for  $i = 1$  and  $R_i = 1$  for  $i = 2$ . The multinomial response is coded as follows:

$$\begin{aligned}
 y_{tijk} = & \quad 0 & \quad \text{if no event has occurred} \\
 & \quad 1 & \quad \text{if individual discontinues contraceptive use } (i = 1), \text{ or} \\
 & & \quad \text{if individual starts to use contraception } (i = 2) \\
 & \quad 2 & \quad \text{if individual switches to another method } (i = 1)
 \end{aligned}$$

In addition, two indicator variables,  $I_1$  and  $I_2$ , denoting the origin state are created. These are interacted with  $t$  and covariates. The restructured dataset is as follows:

$k$	$j$	$i$	$t$	$y_{tijk}$	$I_1$	$I_2$	$I_1*t$	$I_2*t$
1	1	1	0	0	1	0	0	0
1	1	1	1	0	1	0	1	0
1	1	1	2	0	1	0	2	0
1	1	1	3	1	1	0	3	0
1	2	2	0	0	0	1	0	0
1	2	2	1	0	0	1	0	1
1	2	2	2	1	0	1	0	2
1	3	1	0	0	1	0	0	0
1	3	1	1	0	1	0	1	0
1	3	1	2	0	1	0	2	0
1	3	1	3	0	1	0	3	0
1	3	1	4	2	1	0	4	0

To fit a multilevel multinomial model in  $MLwiN$ , the data must be further restructured to obtain a set of binary responses for each multinomial response. This reconstruction is required only for episodes that originate in state 1, from which two types of exit are considered; since there is only one type of exit from non-use the indicator of event occurrence for episodes originating in state 2 is binary. For  $i=1$ , the multinomial response  $y_{t1jk}$  for each time interval is converted to two binary responses  $y_{t1jk}^{(r)}$ , where  $y_{t1jk}^{(r)}=1$  if  $y_{t1jk} = r$  and 0 otherwise ( $r = 1, 2$ ). For each time interval, the two binary responses are stacked. Thus, for the first episode in the example above the final data structure is as follows:

$t$	$r$	$\mathcal{Y}_{t1jk}^{(r)}$	$I_1^{(1)}$	$I_1^{(2)}$	$I_1^{(1)} * t$	$I_1^{(2)} * t$
0	1	0	1	0	0	0
0	2	0	0	1	0	0
1	1	0	1	0	1	0
1	2	0	0	1	0	1
2	1	0	1	0	2	0
2	2	0	0	1	0	2
3	1	1	1	0	3	0
3	2	0	0	1	0	3

The indicator for state 1,  $I_1$ , is replaced by indicators for  $r$ ,  $I_1^{(1)}$  and  $I_1^{(2)}$ . These are multiplied with duration and the covariates to allow duration and covariate effects to vary according to the type of transition from contraceptive use. The destination-specific individual random effects for state 1,  $u_{1k}^{(1)}$  and  $u_{1k}^{(2)}$ , are fitted by allowing the coefficients of  $I_1^{(1)}$  and  $I_1^{(2)}$  to vary randomly across individuals. In addition the random effect for state 2,  $u_{2k}$ , is obtained by allowing the coefficient of  $I_2$  to vary across individuals.

## Appendix B : Estimation of a Multilevel Multistate Competing Risks Model in MLwiN

All the MCMC results in this paper were obtained using a modified version of *MLwiN* (2.0) that will be made publicly available in the future. Here, we describe an MCMC algorithm for estimation of the multilevel multistate competing risks model of equation (2) in Section 2.3.

The algorithm is described in the context of the application to contraceptive use and non-use in Indonesia. There are  $s=2$  states, with  $R_1=2$  possible transitions from state  $i=1$ , and  $R_2=1$  transition from state  $i=2$ . There are six sets of fixed effects, which have been split into duration effects ( $\boldsymbol{\alpha}_1^{(1)}$ ,  $\boldsymbol{\alpha}_1^{(2)}$  and  $\boldsymbol{\alpha}_2$ ) and covariate effects ( $\boldsymbol{\beta}_1^{(1)}$ ,  $\boldsymbol{\beta}_1^{(2)}$  and  $\boldsymbol{\beta}_2$ ) and three sets of random effects ( $u_{1k}^{(1)}$ ,  $u_{1k}^{(2)}$  and  $u_{2k}$ ). All of these parameters are updated using single-site random walk Metropolis updating steps. We also have a 3\*3 variance matrix,  $\Omega_u$ , for the correlated sets of random effects and for this we use a Gibbs sampling step. For prior distributions we use ‘improper’ uniform priors for all of the fixed effects and a diffuse inverse-Wishart prior with parameters 3 and  $S_3 = 3*I$  (the identity matrix) for  $\Omega_u$ .

We make the following substitutions in (2) to simplify writing down the conditional posterior distributions:

$$\text{Let } \mu_{t1jk}^{(r)} = \exp(\boldsymbol{\alpha}_1^{(r)T} \mathbf{z}_{t1} + \boldsymbol{\beta}_1^{(r)T} \mathbf{x}_{t1jk} + u_{1k}^{(r)}), \quad r = 1, 2, \text{ and } \mu_{t2jk} = \exp(\boldsymbol{\alpha}_2^T \mathbf{z}_{t2} + \boldsymbol{\beta}_2^T \mathbf{x}_{t2jk} + u_{2k}).$$

The joint posterior distribution is proportional to

$$p(\Theta | y) \propto \prod_{jk} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1} \left(\mu_{t1jk}^{(1)} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1}\right)^{y_{t1jk}=0} \left(\mu_{t1jk}^{(2)} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1}\right)^{y_{t1jk}=1} \left(\mu_{t1jk}^{(2)} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1}\right)^{y_{t1jk}=2} \\ \times \prod_{jk} \left(1 + \mu_{t2jk}\right)^{-1} \left(\mu_{t2jk} \left(1 + \mu_{t2jk}\right)^{-1}\right)^{y_{t2jk}=0} \prod_k |\Omega_u|^{-1/2} \exp\left(-\frac{1}{2} \mathbf{u}_k^T \Omega_u^{-1} \mathbf{u}_k\right) \times p(\Omega_u)$$

where  $\mathbf{u}_k = (u_{1k}^{(1)}, u_{1k}^{(2)}, u_{2k})$  and  $\Theta$  is the set of all unknown parameters. When we come to calculate the conditional posterior distributions for the unknown parameters they generally do not have standard forms and consist of all the terms in the above joint posterior that contain the parameter of interest. For example the posterior distribution for  $\alpha_1^{(1)}$  has the form:

$$p(\alpha_1^{(2)} | y, \Phi) \propto \prod_{jk} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1} \left(\mu_{t1jk}^{(1)} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1}\right)^{y_{t1jk}=0} \left(\mu_{t1jk}^{(2)} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1}\right)^{y_{t1jk}=1} \left(\mu_{t1jk}^{(2)} \left(1 + \sum_{r=1}^2 \mu_{t1jk}^{(r)}\right)^{-1}\right)^{y_{t1jk}=2}$$

which is the first term in the joint posterior. Here  $\Phi = \Theta \setminus \{\alpha_1^{(1)}\}$ .

The MCMC algorithm works by updating each of the unknown parameters in turn by making a random draw from their conditional posterior distributions. The variance matrix,  $\Omega_u$ , is updated by Gibbs sampling and has an inverse Wishart conditional distribution:

$$p(\Omega_u^{-1} | y, \Theta / \{\Omega_u\}) \sim W_3 \left[ n_w + 3, \left( \sum_{k=1}^{n_w} \mathbf{u}_k \mathbf{u}_k^T + S_3 \right)^{-1} \right]$$

where  $n_w$  is the number of women in the dataset.

All other parameters are updated by random-walk Metropolis sampling which we will illustrate via the step for  $\alpha_1^{(1)}$ . At iteration  $m$  generate a proposed new value  $\alpha_1^{(1)*}$  from the random walk proposal distribution  $\alpha_1^{(1)*} \sim N(\alpha_1^{(1)}(m-1), \sigma_p^2)$  where  $\sigma_p^2$  is the proposal

distribution variance which will be tuned via the adaptive method originally used in Browne and Draper (2000).

The updating step is then:

$$\boldsymbol{\alpha}_1^{(1)}(m) = \boldsymbol{\alpha}_1^{(1)*} \text{ with probability } \min[1, p(\boldsymbol{\alpha}_1^{(1)*} | y, \Phi) / p(\boldsymbol{\alpha}_1^{(1)}(m-1) | y, \Phi)],$$

$$\boldsymbol{\alpha}_1^{(1)}(m) = \boldsymbol{\alpha}_1^{(1)}(m-1) \text{ otherwise.}$$

Similar steps are performed for each of the other unknown parameters. The procedure of updating all the unknown parameters is then repeated many times to generate a large sample of estimates for each parameter. We used a burn-in of 5 000 iterations to allow the chains of parameter estimates to converge and then sampled 50 000 iterations.