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

Published in *Journal of Statistical Modelling*, **4**(2): 145-159.

*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

*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 states and multiple types of event (competing risks) which may vary across states. The different transitions are modelled jointly to allow for correlation across transitions in unobserved individual risk factors. Implementation of the methodology using existing multilevel models for discrete response data is described. The model is applied in an analysis of contraceptive use dynamics in Indonesia where transitions from two 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. Before adjusting for covariate effects, there is a strong negative residual correlation between the hazards of a transition from use to non-use and from non-use to use; this correlation is due to a tendency for short periods of non-use after a birth to be followed by long periods of using the same contraceptive method.

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 states and multiple types of event (competing risks). While there are methods for handling repeated events combined with either multiple 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 several states and multiple types of event. A key advantage of our approach is that it may be implemented using any software that can fit multilevel models to multinomial response data. Practical issues such as data preparation and model specification are also discussed in the paper, and in further documentation available from the *Statistical Modelling* website.

The methodological development is motivated by a study of contraceptive use dynamics. Monthly data on contraceptive use are now collected retrospectively in a number of developing countries, as part of the Demographic and Health Survey (DHS) programme. From these data it is possible to construct a contraceptive use history consisting of episodes of use and non-use. An episode of use is typically defined as a continuous period of using the same method of contraception. 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). Periods of non-use have been ignored. However, the transition from non-use to use is also important for evaluation of family planning programmes since women who do not quickly resume contraceptive use after a birth, or after discontinuing use of a method, are likely to 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 of different contraceptive methods and non-use simultaneously. Use and non-use of contraception are examples of non-absorbing states. By jointly modelling transitions from use and non-use, it is possible to test explicitly for state-dependent covariate effects. A joint model also allows for residual correlation in individual transition rates across states, which might arise because of unobserved factors that affect more than one type of transition.

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. These two types of event can be thought of as competing risks. 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 current 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 separate individual-specific random effects for transitions between use and non-use, which may be correlated to allow for shared unobserved individual-level 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 Terminology: episodes and states

In the general multistate model described below, an episode is defined as a continuous period of time spent in a state until an event occurs. The occurrence of an event does not necessarily lead to a change of state. Therefore there may be repeated episodes within the same state. For example, in the application to contraceptive use dynamics, an episode of use is defined as a continuous period of using the same contraceptive method. There are two types of event that may occur within the use state: a transition to non-use (referred to as a discontinuation), or a transition to a different method (a method switch). While the first type of event leads to a change in state, a method switch results in the start of a new episode in the use state (or a transition within the same state).

The above definition of an episode differs from that conventionally used in multistate models, where an episode is defined as a continuous period spent in the same state and, therefore, an event always leads to a transition to another state. It is possible to redefine episodes or states to follow this more usual definition. In our application, one option would be to redefine a use episode as a continuous period spent in the use state, i.e. using *any* contraceptive method.

However, this definition does not allow estimation of switching rates or method-specific discontinuation rates – both of which are commonly used for evaluation purposes. An alternative approach is to treat each contraceptive method as a separate state, so that a switch from one method to another leads to a change in state. This approach is problematic from a practical point of view. There are more than ten methods available in Indonesia, and some have few users. Thus having a state for each method would lead to a large transition matrix, with some extremely rare transitions which are not of substantive interest; of major interest for family planning programme evaluation are the broader types of transition (discontinuation, switching, and non-use to use). We therefore follow our original definition of episodes and states, but consider method-specific effects of duration and covariates on discontinuation and switching rates. Our definition of episodes and states may also be useful in other applications. For example in studies of employment, one could group different types of job or employer into a single employment state and treat a continuous period spent with the same employer as an episode.

2.2 Previous work on repeated events, multiple 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. 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 non-absorbing states between which individuals move. Several previous studies have considered models for repeated transitions between multiple states. Enberg et al. (1990) use 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 state. In many applications, this assumption of independence may be invalid since there may be unobserved factors which influence transitions from more than one state. Goldstein et al. (2004) also use a discrete-time random effects model, but model jointly transitions from two 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). While existing methods allow for multiple 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 possible ways in which an episode may end, or an event may occur for one of several reasons. To allow for unobserved individual heterogeneity in the risks of competing events, various 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. Their model, however, 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 states.

2.3 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 states. We focus on discrete-time models for several reasons. First, as is common in studies of human populations where event times are collected retrospectively, durations of use and non-use of contraception are measured in months. A second reason for favouring a discrete-time approach is that existing methodology for multilevel discrete response data may be used to handle repeated events. Other benefits of discrete-time models include straightforward inclusion of time-varying covariates and the possibility to allow for non-proportional hazards.

Suppose that durations of episodes are measured in discrete-time intervals indexed by t (t=0, 1, 2 etc). An event time is measured from the start of an episode; when an event occurs a new episode begins and the clock is reset to zero. For each discrete time interval t of episode j for individual k, we observe a multinomial variable y_{tjk} which denotes whether an event has occurred during the interval and the type of event. Suppose there are R end events. The multinomial response is coded so that $y_{tjk} = r$ if an event of type r has occurred in time interval t, r = 1, . . ., 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 the start of 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

$$\log\left(\frac{h_{tjk}^{(r)}}{h_{tjk}^{(0)}}\right) = \boldsymbol{\alpha}^{(r)^T}\mathbf{z}_t^{(r)} + \boldsymbol{\beta}^{(r)^T}\mathbf{x}_{tjk}^{(r)} + u_k^{(r)}, \qquad r = 1, \ldots, R. \tag{2.1}$$

The effect of duration is represented by $\boldsymbol{\alpha}^{(r)^T}\mathbf{z}_t^{(r)}$, where $\mathbf{z}_t^{(r)}$ is a vector of functions of t and $\boldsymbol{\alpha}^{(r)}$ is a parameter vector. For example if a quadratic function of duration is assumed for each type of event, $\mathbf{z}_t^{(r)} = (1, t, t^2)$. Alternatively if the duration effect is assumed piecewise constant, $\mathbf{z}_t^{(r)}$ will be a vector of dummy variables for (possibly grouped) time intervals. The covariates, represented by $\mathbf{x}_{tjk}^{(r)}$, may be defined at the level of the discrete time interval (time-dependent), or at the episode or individual level. Equation (2.1) 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 $\mathbf{x}_{tjk}$ 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 $\boldsymbol{\Omega}_u$; 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 $\mathbf{z}_t^{(r)}$ and $\mathbf{x}_{tjk}^{(r)}$ to vary randomly across individuals.

Model (2.1) may be estimated as a multilevel multinomial model (Goldstein, 2003, Chapter 4). Several software packages may be used, including MLwiN (Rasbash et al., 2003), 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).

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

The model we propose is an extension of (2.1) to handle situations where there are both competing risks and multiple states. The approaches of Steele et al. (1996b) and Goldstein et al. (2004) are combined in a general framework. In this general model, the number and type of events may differ for each state. As discussed in Section 2.1, an episode does not necessarily lead to a change in state, a situation we refer to as a transition within a state.

Suppose there are R_i ways in which an episode in state i (i = 1, . . ., s) can end. Denote by $h_{tijk}^{(r_i)}$ the hazard of making a transition of type r_i (r_i = 1, . . ., R_i) in state i in time interval t of episode j for individual k. The hazard of no transition is denoted by $h_{tijk}^{(0)}$. A multilevel model for competing risks and multiple states may be written

$$\log\left(\frac{h_{tijk}^{(r_i)}}{h_{tijk}^{(0)}}\right) = \boldsymbol{\alpha}_i^{(r_i)^T}\mathbf{z}_{ti}^{(r_i)} + \boldsymbol{\beta}_i^{(r_i)^T}\mathbf{x}_{tijk}^{(r_i)} + u_{ik}^{(r_i)}, \qquad r_i = 1, \ldots, R_i;\ i = 1, \ldots, s. \tag{2.2}$$

In (2.2) duration and covariate effects may depend both on the state i and the type of transition r_i. Unobserved individual-level factors, represented by random effects $u_{ik}^{(r_i)}$, 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.

2.5 Data preparation and estimation

In this section, we give a brief description of the data reconstruction required before fitting a discrete-time model for multiple states and competing risks. Further details are given in accompanying documentation available from the *Statistical Modelling* website.

In order to estimate a discrete-time event history model, the data must first be expanded so that there is a record for each discrete time interval of 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_{0jk}, y_{1jk}, y_{2jk}) = (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 (2.1) or (2.2) may be estimated using any software that can handle multilevel multinomial response data.

One potential disadvantage of a discrete-time approach is that the expanded dataset may be very large, particularly if the width of the discrete time intervals is short relative to the observation period. One strategy to reduce the number of records generated is to group discrete-time intervals. In our application, contraceptive use and non-use durations are grouped into six-month intervals. If the hazard function and covariate values are constant within each six month period, grouping will not lead to loss of information as long as the grouped intervals are weighted by exposure time; for each six-month interval, a weight is defined as the number of months during that interval for which the individual was exposed to the risk of an event. For example, suppose that the episode in the above example ends during month 14. If durations are grouped into six-monthly intervals, the episode will be expanded to three records (corresponding to intervals 0-5, 6-11, and 12-17 months). The woman was at risk of an event for the full six months of the first two six-month intervals, but only three months of the third interval. The response vector for this episode will be (0, 0, 2) as before, but the weight vector will be (6, 6, 3). These weights form the denominators for the multinomial responses. Provided weights are used, grouped time intervals need not be of equal width.

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 the Appendix. This method has been implemented in *MLwiN* v2.0. 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 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 for a six-year period before the survey. The analysis is based on episodes of contraceptive use and non-use for 14 677 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 use or non-use 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: age at the start of the episode, 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

As described in Section 2.5, before fitting 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 contained 543 737 observations. However, as for each transition the hazard is fairly constant within six-month intervals and none of the covariates are time-varying, the width of discrete-time intervals was increased to six months. This reduced the size of the dataset to 109 666 observations. In the analysis, observations were weighted by the number of months for which a woman was exposed to the risk of an event during that six-month interval (see Section 2.5). Note that aggregation of time intervals does not affect the number of episodes. If there was more than one episode within a six-month interval, all such episodes were retained in the reduced dataset, with a duration of one six-month interval (and appropriate weight) recorded for each.

Duration effects were modelled in different ways for use and non-use states. For transitions from contraceptive use, a piecewise constant formulation was found to be a good fit to the observed logit-hazard. A step function was fitted for duration intervals of 0-5, 6-11, 12-23, 24-35, and 36 or more months. For transitions from non-use to use, a polynomial function of the cumulative duration of non-use was used. Non-proportional effects of contraceptive method on discontinuation and switching rates were tested by including interactions between method and duration in the models for transitions from use. Interactions between method and background characteristics were also considered. Since none of these interactions were statistically significant, the selected model includes only main effects of method.

3.3 Results

Cumulative transition probabilities were calculated using separate life tables for each 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 began by fitting a model including only duration effects (method-specific for transitions from use), before adding the covariates listed in Table 1. The estimated random effects covariance matrices from both models are shown in Table 2. There is evidence of unobserved heterogeneity between women in the hazards of all types of transition. From the upper panel of Table 2, it can be seen that before including covariates there is a strong negative residual correlation (estimated as -0.78) 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 found 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 found that the residual correlation became small and non-significant (see the estimate of –0.05 in the lower panel of Table 2). The correlations between the random effects for the other pairs of transitions are also small and neither is significant at the 5% level.

3.3.2 Fixed effects

The estimated coefficients and standard errors for the fixed part of the full model are shown in Table 3. 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. 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. Turning to the effects of demographic and socio-economic characteristics, we see that 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.

Next we 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, but this negative duration effect is weaker for episodes of non-use following a birth. At all durations, a woman is considerably more likely to adopt contraception if the period of non-use follows a birth rather than an episode of contraceptive use. 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. Older, uneducated, rural or low socio-economic status non-users are less likely to become contraceptive users than those who are young, educated, urban or better-off.

4. Discussion

We have shown how to specify and fit general discrete-time event history models with multiple states and multiple types of transition. 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 could 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 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.

Acknowledgements

The authors would like to thank the Editor, Associate Editor and referees whose comments and suggestions have helped to improve the paper.

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*. 3rd edition. London: Arnold.

Goldstein, H., Pan, H. and Bynner, J. (2004) A flexible procedure for analysing longitudinal event histories using a multilevel model. *Understanding Statistics* (to appear).

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.

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., Steele, F. and Browne, W.J. (2003). *A User's Guide to MLwiN, Version 2.0.* London: Institute of Education.

SAS Institute Inc. (1999) SAS/STAT® *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.

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

	Number of women	% of women
Woman-level variables		
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
Episode-level variables: non-use		
Episode follows a live birth		
No	7677	36.1
Yes	13608	63.9
Age at start if episode (years)		
<25	8050	37.8
25-34	10171	47.8
35-49	3064	14.4
Episode-level variables: use		
Contraceptive method		
Pill/injectable	13736	77.0
Norplant®/IUD	2740	15.4
Other modern	323	1.8
Traditional	1044	5.9
Age at start of episode (years)		
<25	7165	40.2
25-34	8277	46.4
35-49	2401	13.5

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)
Duration effects only						
Use → non-use	0.625	(0.493,0.772)				
Use → other method	0.014 **0.020**[a]	(-0.078,0.103)	0.731	(0.639,0.829)		
Non-use → use	-0.369 **-0.783**[a]	(-0.436,-0.303)	0.084 **0.165**[a]	(0.013,0.150)	0.355	(0.288,0.422)
Duration + covariates						
Use → non-use	0.272	(0.189,0.380)				
Use → other method	0.005 **0.011**[a]	(-0.078,0.079)	0.674	(0.585,0.770)		
Non-use → use	-0.031 **-0.052**[a]	(-0.098,0.034)	0.089 **0.095**[a]	(-0.001,0.182)	1.325	(1.199,1.457)

†Coefficients are the modal estimates from 50 000 chains.

[a]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	-4.143	(0.085)	-5.275	(0.120)	-4.100	(0.113)
Duration (months)						
0-5	0	-	0	-	-	-
6-11	-0.057	(0.046)	-0.368	(0.050)	-	-
12-23	-0.076	(0.045)	-0.593	(0.052)	-	-
24-35	0.059	(0.055)	-0.579	(0.066)	-	-
36+	0.175	(0.064)	-0.423	(0.077)	-	-
Duration	-	-	-	-	-0.779	(0.065)
Duration^2	-	-	-	-	0.056	(0.008)
Age (years)						
<25	0	-	0	-	0	-
25-34	-0.398	(0.036)	-0.312	(0.044)	-0.209	(0.033)
35-49	-0.680	(0.061)	-0.505	(0.072)	-0.600	(0.052)
Education						
None	0	-	0	-	0	-
Primary	0.036	(0.073)	0.436	(0.112)	0.310	(0.060)
Secondary+	0.246	(0.075)	0.832	(0.114)	0.708	(0.066)
Region of residence						
Rural	0	-	0	-	0	-
Urban	0.126	(0.038)	0.055	(0.048)	0.261	(0.039)
Socio-economic status						
0-1 (low)	0	-	0	-	0	-
2-3 (medium)	-0.123	(0.048)	0.346	(0.071)	0.240	(0.046)
4-5 (high)	-0.197	(0.054)	0.288	(0.077)	0.451	(0.052)
Contraceptive method						
Pill/injectable	0	-	0	-	-	-
Norplant®/IUD	-1.226	(0.062)	-1.070	(0.070)	-	-
Other modern	0.525	(0.117)	0.553	(0.130)	-	-
Traditional	0.036	(0.066)	-0.596	(0.100)	-	-
Episode follows live birth					2.673	(0.087)
Episode after live birth*Duration	-	-	-	-	0.184	(0.069)
Episode after live birth*Duration^2	-	-	-	-	-0.014	(0.009)

†Estimates are the modal estimates from 50 000 chains.

Appendix: Estimation of a Multilevel Multistate Competing Risks Model

All the MCMC results in this paper were obtained using *MLwiN* v2.0. Here, we describe an MCMC algorithm for estimation of the multilevel multistate competing risks model of equation (2.2).

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 types of end event in state i=1, and R_2=1 end event in 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, Ω_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 Ω_u.

We make the following substitutions in (2.2) to simplify writing down the conditional posterior distributions:

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

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

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

where $\mathbf{u}_k=(u_{1k}^{(1)},u_{1k}^{(2)},u_{2k})$ and Θ 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 $\boldsymbol{\alpha}_1^{(1)}$ has the form:

$$p(\boldsymbol{\alpha}_1^{(2)} \mid y,\Phi) \propto \prod_{jk}\left((1+\sum_{r=1}^{2}\mu_{t1jk}^{(r)})^{-1}\right)^{y_{t1jk}=0}\left(\mu_{t1jk}^{(1)}(1+\sum_{r=1}^{2}\mu_{t1jk}^{(r)})^{-1}\right)^{y_{t1jk}=1}\left(\mu_{t1jk}^{(2)}(1+\sum_{r=1}^{2}\mu_{t1jk}^{(r)})^{-1}\right)^{y_{t1jk}=2}$$

which is the first term in the joint posterior. Here $\Phi = \Theta \setminus \{\boldsymbol{\alpha}_1^{(1)}\}$.

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, Ω_u, is updated by Gibbs sampling and has an inverse Wishart conditional distribution:

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

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 $\boldsymbol{\alpha}_1^{(1)}$. At iteration m generate a proposed new value $\boldsymbol{\alpha}_1^{(1)*}$ from the random walk proposal distribution $\boldsymbol{\alpha}_1^{(1)*}$~N($\boldsymbol{\alpha}_1^{(1)}$ (m-1),σ_p^2) where σ_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:

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

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