MIXGSUR: a computer program for mixed-effects grouped-time survival analysis

Donald Hedeker
School of Public Health and Prevention Research Center
University of Illinois at Chicago

Correspondence to Dr. Donald Hedeker, Division of Epidemiology & Biostatistics (M/C 922), School of Public Health, University of Illinois at Chicago, 2121 West Taylor Street, Room 510, Chicago, IL, 60612-7260. Phone: (312) 996-4896. E-MAIL: HEDEKER@UIC.EDU

Abstract

MIXGSUR provides maximum marginal likelihood estimates for mixed-effects regression analysis of grouped-time survival data. Both proportional hazards (complementary log-log link) and proportional odds (logistic link) versions of the model are available, as well as partial proportional hazards and odds generalizations. The probit and log-log response functions are also available. Right-censoring is allowed for the (grouped) timing of the event. These models can be used for analysis of correlated grouped-time survival data, for example, data arising from a clustered design. For clustered data, the mixed-effects model assumes that data within clusters are dependent. The degree of dependency is jointly estimated with the usual model parameters, thus adjusting for dependence resulting from clustering of the data. MIXGSUR uses marginal maximum likelihood estimation, utilizing a Fisher-scoring solution. For the scoring solution, the Cholesky factor of the random-effects variance-covariance matrix is estimated, along with the effects of model covariates. Examples illustrating usage and features of MIXGSUR are provided.

Keywords: censored observations, proportional hazards model; discrete survival data; frailty; heterogeneity; clustering; multilevel data

1 Introduction

Models for grouped-time survival data are useful for analysis of failure time data when subjects are measured repeatedly at fixed intervals in terms of the occurence of some event, or when determination of the exact time of the event is only known within grouped intervals of time. Additionally, it is often the case that subjects are observed nested within clusters (*i.e.*, schools, firms, clinics), or are repeatedly measured in terms of recurrent events. In this case, use of grouped-time models that assume independence of observations [1, 2, 3] is problematic since observations from the same cluster or subject are usually correlated.

For data that are clustered and/or repeated, models including random effects provide a convenient way of accounting for association in correlated survival data. In terms of continuous-time survival data, several authors [4, 5, 6, 7, 8, 9, 10] have developed survival analysis models including random effects that are usually assumed to be distributed as a gamma distribution. These models are often termed frailty models or survival models including heterogeneity, and recent review articles describe many of these models [11, 12]. In terms of discrete-time survival models, Han and Hausman [13] have proposed a parametric proportional hazards model incorporating a gamma distribution specification of heterogeneity. Specifically, these authors utilize an ordinal logistic regression model for the discrete-time data and a gamma distribution for heterogeneity to arrive at a closed-form solution.

An alternative approach for dealing with correlated data uses the generalized estimating equations (GEE) method described by Liang and Zeger [14] to estimate model parameters under various "working" correlational structures. In this regard, Wei, Lin, and Weissfeld [15] have developed a continuous-time survival model for correlated data, while Guo and Lin [16] have proposed a multivariate model for grouped-time survival data.

Several authors have noted the relationship between ordinal regression models (using complementary log-log and logistic link functions) and survival analysis models for grouped and discrete time [17, 13, 18]. Recently, Hedeker, Siddiqui, and Hu [19] described a generalization of an ordinal random-effects regression model [20] to handle correlated grouped-time survival data. This model accomodates multivariate normally-distributed random effects, and additionally, allows for a general form for model covariates. Assuming either a proportional or partial proportional hazards or odds model, a maximum marginal likelihood solution is implemented using multi-dimensional quadrature to numerically integrate over the distribution of random-effects. An iterative Fisher scoring solution provides relatively quick convergence and standard errors for all model parameters. This paper describes the FORTRAN program MIXGSUR (mixed-effects grouped-time survival model) for the analysis of correlated grouped-time survival data. Examples will illustrate features of MIXGSUR.

2 Computational Methods

Hedeker, Siddiqui, and Hu [19] describe the statistical development of the mixed-effects grouped-time survival model; here we will present the key computational features. Using the terminology of multilevel analysis [21] let i denote the level-2 units (clusters) and let j denote the level-1 units (nested observations). Assume that there are $i = 1, \ldots N$ level-2 units and $j = 1, \ldots, n_i$ level-1 units nested within each level-2 unit. Suppose that there is a continuous random variable for the uncensored time of event occurence (which may not be observed), however assume that time (of assessment) can take on only discrete positive values $t = 1, 2, \ldots, m$. For each level-1 unit, observation continues until time t_{ij} at which point either an event occurs or the observation is censored. In discrete time, censoring means that the level-1 unit is observed at t_{ij} but not at $t_{ij} + 1$.

Define P_{ijt} to be the probability of failure, up to and including time interval t, that is,

$$P_{ijt} = \Pr\,[t_{ij} \leq t] \tag{1}$$

and so the probability of survival beyond time interval t is simply $1 - P_{ijt}$. Noting that, in discrete time, $1 - P_{ijt}$ represents the survivor function, McCullagh [17] proposed the following grouped-time version of the continuous-time proportional hazards model:

$$\log[-\log(1 - P_{ijt})] = \alpha_{0t} + \boldsymbol{x}'_{ij}\boldsymbol{\beta}\ . \tag{2}$$

The response function for P_{ijt} is the so-called complementary log-log function, which can be re-expressed in terms of the cumulative failure probability P_{ijt} as

$$P_{ijt} = 1 - \exp(-\exp(\alpha_{0t} + \boldsymbol{x}'_{ij}\boldsymbol{\beta}))\ . \tag{3}$$

Here, $\boldsymbol{x}_{ij}$ is a $p \times 1$ vector of covariates that vary either at level 1 or 2. The covariates $\boldsymbol{x}_{ij}$ do not vary with time, however, they may represent the average of a variable across time, or the value of the covariate at the time of the event. Since the integrated hazard function equals $-\log(1 - P_{ijt})$, the covariate effects ($\boldsymbol{\beta}$) are identical to those in the grouped-time version of the proportional hazards model described by Prentice and Gloeckler [2]. As such, the $\boldsymbol{\beta}$ coefficients are also identical to the coefficients in the underlying continuous-time proportional hazards model. Augmenting the coefficients $\boldsymbol{\beta}$, the intercept terms α_{0t} are a set of m constants that represent the logarithm of the integrated baseline hazard (*i.e.*, when $\boldsymbol{x} = \boldsymbol{0}$). As such, these terms represent cutpoints on the integrated baseline hazard function. While the above model is the same as that described in McCullagh [17], it is written so that the covariate effects are of the same sign as the Cox proportional hazards model [22]. A positive coefficient for a regressor then reflects increasing hazard with greater values of the regressor.

To add random effects into this model, simply write

$$P_{ijt} = 1 - \exp(-\exp z_{ijt}) \tag{4}$$

where $z_{ijt} = \alpha_{0t} + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{w}'_{ij}\boldsymbol{v}_i$, and $\boldsymbol{v}_i$ is the $r \times 1$ vector of unknown random effects for the level-2 unit i. In the case of only a random-intercept, the model simplifies to $z_{ijt} = \alpha_{0t} + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + v_i$, however, in what follows we will maintain the more general specification of r random effects.

The distribution of the r random effects $\boldsymbol{v}_i$ is assumed to be multivariate normal with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. Assuming the mean vector to be equal to $\mathbf{0}$, leads to the random effects being specified as deviations from the model. Typically, then, the variables specified in $\boldsymbol{w}$ will also be included in $\boldsymbol{x}$, so that the mean of the random effects is also estimated. Since the level-2 subscript i is present for the $\boldsymbol{w}$ vector, not all level-2 units are assumed to have the same number of level-1 observations nested within. Thus, there is no assumption of equal sample sizes within clusters for clustered data, or the same number of repeated observations per subject for repeated data.

2.1 Non-proportional Hazards

The parameters α_{0t} represent a common set of points on the integrated baseline hazard function for all observations. This assumption leads to a proportional hazards assumption for the effects of all model covariates. Alternatively, in some cases, it may be of interest to allow non-proportional hazards for a subset of covariates. Models of this type can be called partial proportional hazards models, since they are analogous to the partial proportional odds models described by Peterson and Harrell [23] in the context of a fixed-effects ordinal regression model.

To allow for differential baseline hazard functions, let us partition the $p \times 1$ covariate vector $\boldsymbol{x}$ into an $h \times 1$ vector $\boldsymbol{u}^*$ and a $(p - h) \times 1$ vector $\boldsymbol{x}^*$. The proportional hazards assumption is assumed to hold for the covariates in $\boldsymbol{x}^*$, but not for the covariates in $\boldsymbol{u}^*$. The partial proportional hazards model can then be written as:

$$z_{ijt} = \alpha_{0t} + (\boldsymbol{u}^*_{ij})'\boldsymbol{\alpha}^*_t + (\boldsymbol{x}^*_{ij})'\boldsymbol{\beta}^* + \boldsymbol{w}'_{ij}\boldsymbol{v}_i \tag{5}$$

or absorbing α_{0t} and $\boldsymbol{\alpha}^*_t$ into $\boldsymbol{\alpha}_t$,

$$z_{ijt} = \boldsymbol{u}'_{ij}\boldsymbol{\alpha}_t + (\boldsymbol{x}^*_{ij})'\boldsymbol{\beta}^* + \boldsymbol{w}'_{ij}\boldsymbol{v}_i \tag{6}$$

where, $\boldsymbol{u}_{ij}$ is a $(h+1) \times 1$ vector containing (in addition to a 1 for α_{0t}) the values of observation ij on the subset of h covariates for which interactions with the cutpoints of the integrated baseline hazard are desired, and $\boldsymbol{\alpha}_t$ is a $(h + 1) \times 1$ vector of regression coefficients associated with the h variables (and the intercept) in $\boldsymbol{u}_{ij}$. This extension then allows estimation of partial proportional hazards models, where the effects of a subset of covariates are allowed to vary across the cutpoints of the integrated baseline hazard. In the equation above, both vectors $\boldsymbol{x}$ and $\boldsymbol{\beta}$ have the * superscript to distinguish them as subsets from the original $\boldsymbol{x}$ and $\boldsymbol{\beta}$ vectors.

For simplicity of notation, in what follows, the distinction that these superscripts represent will not always be maintained.

3 Maximum Marginal Likelihood Estimation

To accomodate right censoring, define the status variable $\delta_{ij} = 0$ if level-1 unit ij is a censored observation and equal to 1 if the event occurs (fails) for that level-1 unit ij. The time variable t_{ij} now equals the value of t when either the ijth unit failed or was censored, and all observations either fail or are censored at discrete times $t = 1, \ldots, m$. Again, censoring at time t means that the level-1 unit ij is observed and has not failed at t_{ij} but is not observed at $t_{ij}+1$. With the above random-effects regression model, and response function (4), the probability of a failure at time t for a given level-2 unit i, conditional on $\boldsymbol{\upsilon}, \boldsymbol{\alpha}_t$, and $\boldsymbol{\beta}$ is given by:

$$\Pr(t_j = t \cap \delta_j = 1 \mid \boldsymbol{\upsilon}, \boldsymbol{\alpha}_t, \boldsymbol{\beta}) = P_{j\,t} - P_{j\,t-1} \tag{7}$$

where $P_{j\,0} = 0$ and $P_{j\,m+1} = 1$. Alternatively, the probability of being right censored at time t for a given level-2 unit i, conditional on $\boldsymbol{\upsilon}, \boldsymbol{\alpha}_t$, and $\boldsymbol{\beta}$, is equal to the cumulative probability of not failing at the time the observation is censored, and is given by:

$$\Pr(t_j = t \cap \delta_j = 0 \mid \boldsymbol{\upsilon}, \boldsymbol{\alpha}_t, \boldsymbol{\beta}) = 1 - P_{j\,t}\ . \tag{8}$$

Letting $\boldsymbol{t}_i$ denote the vector pattern of failure times from level-2 unit i for the n_i level-1 units nested within, the probability of any pattern $\boldsymbol{t}_i$, given $\boldsymbol{\upsilon}, \boldsymbol{\alpha}_t$, and $\boldsymbol{\beta}$, is equal to the product of the probabilities of the level-1 responses:

$$\ell(\mathbf{y}_{\boldsymbol{i}} \mid \boldsymbol{\beta}, \boldsymbol{\alpha}) = \prod_{j=1}^{n_i} \prod_{t=1}^{m} \left[(P_{ij\,t} - P_{ij\,t-1})^{\delta_{ij}} (1 - P_{ij\,t})^{1-\delta_{ij}} \right]^{d_{ijt}} \tag{9}$$

$$\text{where } d_{ijt} = \begin{cases} 1 & \text{if } t_{ij} = t \\ 0 & \text{if } t_{ij} \neq t \end{cases} .$$

The marginal density of $\boldsymbol{t}_i$ in the population is expressed as the following integral of the likelihood, $\ell(\cdot)$, weighted by the prior density $g(\cdot)$:

$$h(\boldsymbol{t}_i) = \int_{\boldsymbol{\upsilon}} \ell(\mathbf{y}_{\boldsymbol{i}} \mid \boldsymbol{\beta}, \boldsymbol{\alpha})\ g(\boldsymbol{\upsilon})\ d\boldsymbol{\upsilon} \tag{10}$$

where $g(\boldsymbol{\upsilon})$ represents the multivariate normal distribution of the $\boldsymbol{\upsilon}$ vector in the population. It is convenient to orthogonally transform the response model to use the maximum marginal likelihood (MML) estimation procedure discussed by Bock and Aitkin [24] in the context of a dichotomous factor analysis model. Specifically, let $\boldsymbol{\upsilon} = \boldsymbol{S\theta}$, where $\boldsymbol{SS'} = \boldsymbol{\Sigma}$ is the Cholesky decomposition of $\boldsymbol{\Sigma}$. The model for z_{ijt} is then written as:

$$z_{ijt} = \boldsymbol{u}'_{ij}\boldsymbol{\alpha}_t + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{w}'_{ij}\boldsymbol{S}\boldsymbol{\theta}_i \tag{11}$$

The marginal density then becomes

$$h(\boldsymbol{t}_i) = \int_{\boldsymbol{\theta}} \ell(\mathbf{y}_{\boldsymbol{i}} \mid \boldsymbol{\theta}, \boldsymbol{\alpha})\ g(\boldsymbol{\theta})\ d\boldsymbol{\theta} \tag{12}$$

where $g(\boldsymbol{\theta})$ represents the multivariate standard normal density. In addition to transforming the response model to the multivariate standard normal distribution, a further consequence of transforming from $\boldsymbol{\upsilon}$ to $\boldsymbol{\theta}$ is that the Cholesky factor $\boldsymbol{S}$, which is a lower triangular matrix, is estimated instead of the covariance matrix $\boldsymbol{\Sigma}$. As the Cholesky factor is essentially the square-root of the covariance matrix, this then allows more stable estimation of near-zero variance terms.

For the estimation of the $(p-h)$ covariate coefficients $\boldsymbol{\beta}^*$, the population parameters $\boldsymbol{S}$ (with $r \times (r+1)/2$ elements), the $(h+1) \times m$ (integrated) baseline hazard values and interaction terms in $\boldsymbol{\alpha}_t$ $(t = 1, \ldots, m)$, the marginal log-likelihood for the patterns from the N level-2 units,

$$\log L = \sum_{i}^{N} \log h(\boldsymbol{t}_i)$$

is differentiated with respect to each parameter (see [19]). Fisher's method of scoring can then be used to provide the solution to these likelihood equations. For this, provisional estimates for the vector of parameters $\boldsymbol{\Theta}$, on iteration ι are improved by

$$\boldsymbol{\Theta}_{\iota+1} = \boldsymbol{\Theta}_{\iota} - \mathcal{E}\left[\frac{\partial^2 \log L}{\partial\boldsymbol{\Theta}_{\iota}\, \partial\boldsymbol{\Theta}'_{\iota}}\right]^{-1} \frac{\partial \log L}{\partial \boldsymbol{\Theta}_{\iota}} \tag{13}$$

where the information matrix, or expectation of the matrix of second derivatives, is given by

$$\mathcal{E}\left[\frac{\partial^2 \log L}{\partial\boldsymbol{\Theta}_{\iota}\, \partial\boldsymbol{\Theta}'_{\iota}}\right] = -\sum_{i=1}^{N} h^{-2}(\boldsymbol{t}_i)\ \frac{\partial h(\boldsymbol{t}_i)}{\partial\boldsymbol{\Theta}_{\iota}} \left(\frac{\partial h(\boldsymbol{t}_i)}{\partial\boldsymbol{\Theta}_{\iota}}\right)'.$$

At convergence, the large-sample variance covariance matrix of the parameter estimates is then obtained as the inverse of the information matrix.

3.1 Numerical Quadrature

In order to solve the above likelihood equations, numerical integration on the transformed $\boldsymbol{\theta}$ space must be performed. For this, Gauss-Hermite quadrature can be used to approximate the above integrals to any practical degree of accuracy. In Gauss-Hermite quadrature, the integration is approximated by a summation on a specified number of quadrature points Q for each dimension of the integration; thus, for the transformed $\boldsymbol{\theta}$ space, the summation goes over Q^r points. As the number of random effects r is increased, the terms in the summation (Q^r) increases exponentially in the quadrature solution. Fortunately, as is noted by Bock, Gibbons and Muraki [25] in the context of a dichotomous factor analysis model, the number of points in each dimension can be reduced as the dimensionality is increased without impairing the accuracy of the approximations; they indicated that for a five-dimensional solution as few as three points per dimension were sufficient to obtain adequate accuracy.

3.2 Frequency Weighted Data

The above solution can be modified to accommodate frequency weighted data, for example, when the same $n_i \times 1$ response pattern $\boldsymbol{t}_i$, $n_i \times r$ random-effects design matrix $\boldsymbol{W}_i$, and $n_i \times p$ covariate matrix $\boldsymbol{X}_i$ are observed for a number of level-2 units. In this context, the data can be represented in tabular form, each entry in the table representing a frequency or count of level-2 observations with the same values of $\boldsymbol{t}_i$, $\boldsymbol{W}_i$ and $\boldsymbol{X}_i$. This can occur, for example, in the longitudinal context if subjects are measured on the same number of repeated trials, n, and the frequency of each of the possible $(2m)^n$ response patterns is given for, say, a discrete number of treatment groups. The number of possible response patterns equals $(2m)^n$ since at each of the n trials, there are m possible censored codings, plus m possible event codings.

3.3 Other Response Functions

The description and development above is in terms of the complementary log-log response function, which leads to the proportional hazards model. Alternatively, the proportional odds model has also been proposed [17, 26] especially when the hazards of groups of subjects are thought to converge with time. For the proportional odds model, the logistic response function replaces the complementary log-log function in the derivation. MIXGSUR provides for both response functions, as well as two other less common choices: the probit and log-log response functions.

3.4 Model Parameterization in MIXGSUR

The model, as presented, is not the precise parameterization used in MIXGSUR. Starting with a model with proportional hazards (or odds), instead of the model given by

$$z_{ijt} = \alpha_{0t} + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{w}'_{ij}\boldsymbol{\upsilon}_i \tag{14}$$

MIXGSUR estimates the parameters of the following model:

$$z_{ijt} = \gamma_t + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{w}'_{ij}\boldsymbol{\eta}_i \tag{15}$$

where the random effects $\boldsymbol{\eta}_i$ are normally distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Note that the mean vector is not assumed to be zero, as was the case in the parameterization using $\boldsymbol{\upsilon}_i$ in (14). Also note that the baseline hazard parameters α_{0t} are replaced with the "threshold" parameters γ_t, with the restriction that $\gamma_1 = 0$. The intercept of the model α_{01} in the first parameterization is an element of $\boldsymbol{\mu}$ (if the intercept is assumed to vary between clusters, *i.e.*, the intercept is a random effect) or an element of $\boldsymbol{\beta}$ (if the intercept is not specified as a random effect, but instead specified as a fixed effect). of the latter parameterization.

For example, with a random-intercepts model ($r = 1$ and $w_{ij} = 1$),

$$z_{ijt} = \gamma_t + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \eta_i \tag{16}$$

where η_i has mean μ and variance σ. The model intercept α_{01} in the original formulation (14) is given by μ in this latter formulation of the model. More generally, the cutpoints of the baseline hazard function (α_{0t}) are then given as: $\alpha_{01} = \mu, \alpha_{02} = \mu + \gamma_2, \alpha_{03} = \mu + \gamma_3, \ldots, \alpha_{0m} = \mu + \gamma_m$. Note that if right-censoring is not present at the last timepoint (*i.e.*, when $t = m$), then the program has no information to estimate either α_{0m} or γ_m. Thus, MIXGSUR estimates $m - 1$ thresholds if right-censoring is present at the last timepoint (since $\gamma_1 = 0$), and $m - 2$ thresholds if there is no right-censoring at the last timepoint (since $\gamma_1 = 0$ and no information is available to estimate γ_m) . Estimates of these $m-1$ (or $m-2$) thresholds are listed sequentially in the output.

Now consider the partial proportional hazards (or odds) model given in (5). In its place, MIXGSUR estimates the model given by:

$$z_{ijt} = \gamma_t + (\boldsymbol{u}^*_{ij})'\boldsymbol{\zeta}_t + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{w}'_{ij}\boldsymbol{\eta}_i \tag{17}$$

under the restrictions that $\gamma_1 = 0$ and $\boldsymbol{\zeta}_1 = \mathbf{0}$. Suppose the first h covariates in the $p \times 1$ covariate vector $\boldsymbol{x}$ are considered to have non-proportional hazards (*i.e.*, they comprise the $h \times 1$ covariate vector $\boldsymbol{u}^*$). Partition the $p \times 1$ parameter vector $\boldsymbol{\beta}$ in (17) into $\boldsymbol{\beta}_{(h)}$ (the first h elements) and $\boldsymbol{\beta}_{(p-h)}$ (the remaining $p-h$ elements). The effects at the first timepoint for these h covariates are then denoted as $\boldsymbol{\alpha}^*_1$ in (5) and as $\boldsymbol{\beta}_{(h)}$ in (17). After the first timepoint, the effects of these h covariates are $\boldsymbol{\alpha}^*_t$ in (5) and as $\boldsymbol{\beta}_{(h)} + \boldsymbol{\zeta}_t$ in (17). Thus, the parameter vector $\boldsymbol{\zeta}_t$ represents differences (or deviations) in the effect of the h covariates at timepoint t, relative to the first timepoint. Since each $\boldsymbol{\zeta}_t$ is of size $h \times 1$, with the restriction that $\boldsymbol{\zeta}_1 = 0$, there are a total of $h \times (m-1)$ parameters to be estimated. In listing the estimates of these $h \times (m-1)$ parameters, following the listing of the thresholds, MIXGSUR lists sequentially the $m - 1$ (deviational) effects for covariate 1, the $m - 1$ (deviational) effects for covariate 2, ... , the $m - 1$ (deviational) effects for covariate h. Again, it should be noted that in the absence of right-censoring at the last timepoint, only $h \times (m - 2)$ parameters can be estimated.

Since the parameters estimated by MIXGSUR are linear expressions of the parameters in the model given by (5), it is relatively easy to express the estimated parameters in either representation. To facilitate this, MIXGSUR can calculate linear combinations of the estimated parameters, printing out the re-expressed parameter estimates and corresponding standard error estimates. This feature will be illustrated in the examples below.

4 Program description and usage

MIXGSUR is currently available in executable form for both MS-DOS and MACINTOSH computers. The MIXGSUR instructions *must* be stored in the file MIXGSUR.DEF (described below), and the user begins program execution using the MIXGSUR.EXE file (issuing the command MIXGSUR in DOS, or double-clicking on the MIXGSUR icon on the MACINTOSH). Here, we will detail the procedure for running the program. MIXGSUR makes use of the following files:

- input data file
- MIXGSUR.DEF - main definition file for analysis options and settings
- main output file

4.1 Structure of the input data file

This file contains all data (*i.e.*, level-2 identifier, grouped survival time, censoring variable, and covariates) to be read in by the program. It is read in free format and must be a standard text (ASCII) file with no hidden characters or word processing format codes. Variable fields must be separated by one or more blanks. The data are assumed to consist of multiple level 1 observations within a higher-order (2nd level) unit. There must be a level-2 ID variable for each record and the data must be sorted by this level-2 ID variable. The nested measurements (level 1) of a cluster (level 2) take up as many records in this file as there are level 1 units within that cluster. Thus, some clusters can have, for example, 40 records while others may have 20 to 50 records.

The fields of variables that are read in, separated by one or more blanks, on a line (or lines) are as follows (the order of the variables does not matter):

$$ID \;\; EventTime \;\; EventStatus \;\; Xvector \;\; Wvector$$

where, ID refers to an the level-2 ID number which does not change across level-1 units, $EventTime$ is the value of (grouped) time at which the observation was either censored or experienced the event, $EventStatus$ equals 0 or 1 if the observation was censored or experienced the event, respectively, $Xvector$ is the covariate vector for the observation, and $Wvector$ is the part of the design matrix for the random effects. All variables are read as REAL*8 with the exception of the level-2 IDs which are read as INTEGER. All missing data must have a *numeric* missing value code, in particular, *missing values left as blank fields will definitely cause problems.*

4.2 Analysis options and settings - MIXGSUR.DEF

This file contains the information to determine which statistical model should be fit to the data in the input data file. Although a word processor can be used to create this file, it must be saved as a standard text (ASCII) file with no hidden characters or word processing format codes. The analysis options and settings that comprise this file are described in Tables 1a, 1b, 1c, and 1d.

Insert Tables 1a, 1b, 1c, and 1d about here

Except where noted, this file is read in free format. This file is created by the user directly *before* typing the command MIXGSUR (or on the MACINTOSH, double-clicking on the MIXGSUR file). This filename and extension (MIXGSUR.DEF) *must* be used and should be in the same directory as the program MIXGSUR.EXE or accessible via appropriate PATH statements.

4.3 Main output file

This file contains descriptive information about the variables read in to MIXGSUR, as well as the main results of the specified analysis. The examples of the output file provided below illustrate the contents of this file. In terms of numbers of observations, the number of level-2 units, the total number of level-1 units, and the number of level-1 units for each level-2 unit are listed. For each variable (except the ID variable) read in to the program, the following descriptive statistics are provided: minimum, maximum, mean, and standard deviation. These descriptive statistics are based on the total number of level-1 observations. For the grouped-time variable *EventTime*, a frequency count is provided which lists for each category the number (and proportion) of level-1 observations. It should be noted that this frequency count includes both observations that are censored and observations experiencing the event. An optional listing of the frequencies and proportions of *EventTime* by the levels of one of the model covariates may be obtained. Starting values, either user-defined or program-generated, are listed for all model parameters. Finally, MIXGSUR indicates the number (and percentage) of level-2 units with non varying level-1 responses on *EventTime*.

In terms of program results, the number of iterations required to achieve convergence is listed, followed by the number of quadrature points requested, and the value of the log-likelihood at convergence. As mentioned, the log-likelihood value can be used to perform likelihood-ratio tests. Following the log-likelihood value is a listing of the ridge value. The ridge is an incremental adjustment which is made to the diagonal elements of the information matrix if the program encounters a non-increasing likelihood or some other indication of numerical difficulty during the iterations. This adjustment often improves the chances of convergence. At present, the ridge starts at zero, and is increased by 0.1 each time that difficulties are encountered. At convergence, the ridge is set back to zero in order to obtain the correct standard errors for the model parameters, however the listing of the ridge value indicates its value prior to being reset to zero. As such, the listed ridge value is indicative of the degree of computational difficulty that the program encountered.

For each parameter of the model, maximum marginal likelihood estimates, standard errors, z-values, and p-values are then provided. These p-values are two-tailed, except for the variance and threshold parameters where one-tailed p-values are given. This use of the standard errors to perform hypothesis tests for the variance and threshold parameters is controversial (see Bryk and Raudenbush [27] page 55). Also, it is important to realize that it is the Cholesky factor of the random-effects variance-covariance matrix that is estimated, and not the variance-covariance matrix itself. If only one random effect is requested in the model, the Cholesky factor is simply the square root of the variance, that is, the standard deviation. Analogously, with multiple random effects, the Cholesky factor represents the matrix square root.

Following the parameter estimates (and associated statistics), MIXGSUR lists a correlation matrix associated with the estimates of all model parameters. This correlation matrix does not contain correlations of the variables themselves, but correlations of the estimated model parameters. This matrix may be helpful in determining the degree to which collinearity is present in terms of the model parameters.

4.4 Some Common MIXGSUR Errors

There are a few errors which can prevent MIXGSUR from running correctly, or even running at all. First, as mentioned, missing values that are not given a specified numeric missing value code, but instead are left as blank fields, may cause the program to fail or to estimate a model which is incorrect from the user's perspective. To see if this is occurring, the user can check the correctness of each variable's descriptive statistics (minimum, maximum, mean, and standard deviation) listed in the output file. If these descriptive statistics are incorrect, the data are not being read into the program correctly and a common reason is that missing values are being left as blank fields in the data file. Second, the CATYX option (described in Tables 1b and 1d) is fairly unforgiving. The values listed by the user for the levels of the crosstabulation variable *must* be exactly the same as the values that are found in the data file. If a strange error prevents MIXGSUR from running and this option is selected, the user can set CATYX=0 to avoid this option. Third, the NPR option (described in Table 1b), which is used to list data to the screen, can cause MIXGSUR to stop in certain cases (essentially, when the number of digits to be listed for a variable exceeds the format specification of the program). If the program stops after indicating (on the screen) the number of random and fixed effects in the model, but prior to listing any iterative results to the screen, the user can set NPR=0 and re-run the program. Fourth, problems can develop if the user tries to fit a model with a single random effect, and that random effect is not the intercept. In this case, the procedure used to generate starting values for the program is poor. Instead, the user can choose the START option (described in Table 1b) and specify "naive" starting values of 0 for the mean of the random effect and for the covariate effects, some fraction of the assumed residual variance for the random-effect variance term (*e.g.*, .5 or 1), and some positive increasing values for the thresholds (*e.g.*, from .2 to 1). Finally, if the program "blows up," it may be that the model that is specified is not estimable. In this case, the user should try fitting a less complicated model by specifying fewer random effects, or fewer covariates, or collapsing some of the ordered grouped-time categories if these are very sparse. If the number of random effects is 1, and problems still exist, it may be that the random-effect variance cannot be reliably estimated as being different from zero. In this case, a model without random effects may be warranted.

5 Examples of Mixed-effects Grouped-Time Survival Analysis

MIXGSUR can estimate a variety of models for correlated grouped-time survival data. An analysis of a clustered dataset where students are clustered within classrooms is presented first. Students' time to smoking initiation is analyzed, including one random term in the model to account for the clustering of students within classrooms. This random classroom effect describes the way in which students from the same classroom respond similarly, relative to the sample as a whole. A second example will illustrate use of mixed-effects regression analysis for repeated, or within-subjects, data. This analysis will focus on two repeated observations clustered within students, and will include a subject-varying random effect in the model to account for differences attributable to the student. These two examples will serve to highlight some of the results that are obtained from mixed-effects analysis, and will be accompanied by listings of specific file setups that are used to run MIXGSUR.

These two examples are termed two-level models in the multilevel literature [21]. The first considers students (level-1) nested within classrooms (level-2), while the second considers repeated observations (level-1) nested within students (level-2). At present, MIXGSUR does not allow a three-level analysis which would consider, concurrently, students nested within classrooms *and* repeatedly observed.

5.1 Dataset

The Television School and Family Smoking Prevention and Cessation Project (TVSFP) study [28] was designed to test independent and combined effects of a school-based social-resistance curriculum and a television-based program in terms of tobacco use prevention and cessation. The initial study sample consisted of seventh-grade students who were pretested in January, 1986. Students who took the pretest completed an immediate post-intervention questionnaire in April, 1986, a one-year follow-up questionnaire (April, 1987), and a second year follow-up (April, 1988). The study involved students of schools from Los Angeles and San Diego. Randomization to various design conditions was at the school level, while much of the intervention was delivered to students within classrooms. For the illustrations given below, a subset of the TVSFP data was used. We concentrated on students from Los Angeles schools, where the schools were randomized to one of four study conditions: (a) a social-resistance classroom curriculum (CC); (b) a media (television) intervention (TV), (c) a combination of CC and TV conditions; and (d) a no-treatment control group. These conditions form a 2 x 2 design of CC (yes or no) by TV (yes or no).

Two outcomes of interest from the study include onset of cigarette experimentation and onset of alcohol experimentation. At each of the four timepoints, students answered the questions: "have you ever smoked a cigarette?" and "how many alcoholic drinks have you ever had in your whole life?" (dichotomized between "none" and "only sips"). Since the intervention was implemented following the pretest, the analyses presented focus on data from the three post-intervention timepoints, and include only those students who had not answered "yes" at pretest. The first example, focusing on smoking initiation, includes 1556 students who indicated at pretest that they had never smoked a cigarette. The second example, focusing on smoking and alcohol initiation jointly, includes 1002 students who indicated at pretest that they had never smoked a cigarette or had an alcoholic drink.

5.2 A Clustered Example

For the first illustration of the random-effects grouped-time survival analysis model, we focus on the nesting of students within classrooms. For smoking initiation, there were 1556 students from 134 classrooms, with between one to 23 students per classroom. Since MIXGSUR allows for both the logistic and complementary log-log response functions, both proportional odds and proportional hazards grouped-time models can be fit to these data. This example will use the proportional hazards model.

A partial list of these data for smoking initiation is given in Table 2. The variables are, in order, classroom ID, grouped-time of event (1 = post-intervention, 2 = 1-year followup, and 3 = 2-year followup), event status (0 = censored, 1 = event), a column of ones for the intercept, sex (0 = female, 1 = male), CC (0 = no, 1 = yes), TV (0 = no, 1 = yes), and the product of CC and TV.

Insert Table 2 about here

For these data, in addition to performing a mixed-effects analysis (to include and estimate the effect of clustering on the students' outcomes), MIXGSUR can be used to perform a more basic analysis ignoring data clustering. Tables 3a and 3b list the DEF files for the models ignoring and including data clustering, respectively.

Insert Tables 3a and 3b about here

In both, the grouped-time to smoking initiation (SmkOnset) is modeled in terms of sex, CC, TV, and CC by TV interaction. The titles for these two DEF files are slightly different to reflect the different analyses. In the file designations on lines 3 to 5, the same input data file is specified in both DEF files, however, different OUT files are indicated so that the results of both analyses are saved to different files.

Line 6 in the DEF files contains the primary specifications for these two models. For both, NPR=1, so the data from the first subject will be listed to the screen prior to the analysis. Eight fields of variables are to be read from the input data file, with the first model having 0 and 5 random and fixed effects, respectively, while the second specifies 1 random and 4 fixed effects. As a result of specifying zero random effects (in Table 3a), blank lines are given (below line 6) for the field and label specifications of the random effects. A convergence criterion of .0001 is chosen, and three ordered categories are indicated for the grouped-time variable (with actual values of 1, 2, and 3 specified in both DEF files). The missing value (MISS) option is requested, and missing value codes of 9 are given subsequently for all variables. The number of quadrature points is set to 10 for both models, though the model without random effects does not use quadrature for the solution (thus, for this model only, this specification is irrelevant though it must be given). In terms of the response function, the complementary log-log response function is selected for the analysis. The right-censoring option is indicated, and a censoring variable will later be identified. The covariates and mean of the random effects are to be added to the effects of the thresholds. For survival data, this ensures that the effects are of the same sign as the Cox proportional hazards model, specifically, a positive coefficient is indicative of an increased hazard with increasing values of the covariate. Finally, on line 6, two linear combinations of the estimated parameters are requested (and will be specified later in the DEF file).

In both DEF files, the classroom ID (field 1) is identified as the cluster ID (though again, for the model without random effects this is an irrelevant, though necessary, specification). Also on line 7, the grouped-time variable is identified as being in the second field. The next three lines in the DEF files indicate the fields for the random effects, covariates, and right-censoring variable. The codings for the grouped-time outcome variable are then indicated, as are the missing value codes for all model variables. Labels are then provided for the grouped-time outcome, random effect, and and covariates.

The last two lines of the DEF files correspond to the two linear combinations of the estimated parameters that are requested. The parameter order is as follows: mean of the random effects (R parameters), covariates (P parameters), unique elements of the random-effect variance-covariance matrix (R $\times$ (R-1) / 2 parameters), thresholds (MAXJ-2 or MAXJ-1 parameters, if right-censoring is not indicated, CEN=0, or is indicated, CEN=1, respectively), and interactions with the thresholds (PARTIAL $\times$ (MAXJ-2) or PARTIAL $\times$ (MAXJ-1), depending on the censoring specification). Note that if PARTIAL>1 then the ordering for these interactions with the thresholds is as follows: the first MAXJ-1 (or MAXJ-2) parameters are for the first of the PARTIAL variables, the next MAXJ-1 (or MAXJ-2) parameters are for the second of the PARTIAL variables, etc. In both DEF files, the first linear combination is simply the sum of the intercept and first threshold, while the second linear combination is the sum of the intercept and second threshold. These re-expressions of the estimated parameters then provide estimates of α_{02} and α_{03} (note: the estimate of the intercept equals α_{01}).

Table 4 lists the results for the model specified by the DEF file given in Table 3b, that is, from the analysis treating students nested within classrooms.

Insert Table 4 about here

As can be seen from Table 4, the nesting of the 1556 students within the 134 classrooms is associated with classroom sizes between 1 and 23. Descriptive statistics are listed for all variables considered in the analysis, including a listing of the frequencies in each of the grouped-time outcome categories (again, this includes censored and non-censored observations). Following the descriptive statistics, the program lists the number and percentage of classrooms (*i.e.*, the level-2 units) with non-varying responses across students; 3 classrooms (2.24%) had identical SmkOnset values from tudents. In terms of results, significant effects are not observed for sex or the classroom condition variables. The intracluster correlation (in this case, the intra*classroom* correlation) from this analysis equals 0.021. Comparing the fit of this model to a model that does not include a random classroom effect (*i.e.*, the model specified by the DEF file in Table 3a), yields a likelihood-ratio $\chi^2 = -2[-1593.233 - (-1592.354)] = 1.76$ on 1 degree of freedom ($p < .18$), indicating no clear evidence of a non-zero intra-classroom correlation. Finally, the estimates of the (integrated) baseline hazard at the three timepoints equal -1.70706 (the intercept), -0.99051 (the first linear transform), and -0.47416 (the second linear transform). These values indicate an overall increase in smoking initiation across time.

5.3 A Within-Subjects Example

In this example, the same dataset will be used, but the analysis will jointly examine smoking an alcohol initiation. Specifically, for illustration of the model to repeated survival data, the smoking and alcohol onset data is analyzed together, ignoring the clustering of students within classrooms. The analysis will focus on 1002 students who had never tried either smoking or alcohol at pretest. Since the intervention effect was non-significant in the previous analysis, for simplicity, only the effect of gender is considered, in addition to differences between cigarette and alcohol onset. Terms for gender, substance, and gender by substance interaction will be included in the model, in addition to a random subject effect. Table 5 lists the data from the first 8 subjects.

Insert Table 5 about here

The 7 variables are, in order, subject ID, grouped-time of event (1 = post-intervention, 2 = 1-year followup, and 3 = 2-year followup), event status (0 = censored, 1 =

event), a column of ones for the intercept, sex (0 = female, 1 = male), substance (0=cigarette, 1=alcohol), and gender by substance interaction

It should be noted, that for a given dataset, not all of the variables in the file need to be included in a given analysis. MIXGSUR allows the user to choose among the variables in an input data file for a particular model. In this data file, each subject's data consist of two lines, and missing value codes (9) are present for some subjects at specific timepoints. For inclusion into the analysis, a subject's data (both the grouped-time outcome variable and all model covariates being used in a particular analysis) at a specific timepoint must be complete. The number of repeated observations per subject then depends on the number of timepoints for which there are non-missing data for that subject. The use of missing value codes is not the only way of dealing with missing data. An alternative way of handling the missing data is for the user to physically remove the records with missing data, so that each subject would have a varying number of records in the file.

Tables 6a and 6b list the MIXGSUR.DEF files for two mixed-effects grouped-time survival analyses of these data.

Insert Tables 6a and 6b about here

For both models, grouped-time onset is modeled in terms of a dummy-coded substance effect (AlcOnset: cigarette = 0 and alcohol = 1), a sex effect (SexMale: female = 0 and male = 1), and a substance by sex interaction (AlcSex). In terms of the random effects, both DEF files specify a random-intercepts model, that is, a random subject effect. The first model allows an interaction of grouped-time by substance, while the second model additionally allows an interaction of grouped-time by sex. In so doing, both of these models are partial proportional hazards models, similar in nature to the partial proportional odds models described by Peterson and Harrell [23] in the context of a fixed-effects model. Again, the titles for these two DEF files reflect the different analyses, and different OUT files are indicated so that the results of both analyses are saved to different files.

Line 6 in the DEF files specifies models with 1 random and 3 fixed effects, and three ordered categories for the grouped-time variable (with actual values of 1, 2, and 3 given below in both DEF files). The missing value (MISS) option is requested, and missing value codes of 9 are given subsequently for all variables. The number of quadrature points is set to 10 and the complementary log-log response function is selected for both models. The right-censoring option is indicated, and a censoring variable will later be identified. The first DEF file (Table 6a) indicates that an interaction is to be formed with the first covariate (AlcOnset) and the grouped-time cutpoints. The second DEF file (Table 6b) further indicates interactions for the first two covariates (AlcOnset and SexMale). As in the previous example, the model terms will be added to the threshold effects (*i.e.*, the cutpoints of the integrated baseline hazard). Finally, on line 6, four and six linear combinations of the estimated parameters are requested, respectively.

In both DEF files, the subject ID (field 1) is identified as the cluster ID, and the grouped-time outcome variable is identified as being in the second field. The next three lines in the DEF files indicate the fields for the random effects, covariates, and right-censoring variable. The codings for the grouped-time outcome variable are then indicated, as are the missing value codes for all model variables. Labels are provided for the grouped-time outcome, random effect, and covariates.

In terms of the missing value specification, even though missing values are coded only for the grouped-time outcome variable in the input data file, numeric missing value codes must be specified in the MIXGSUR.DEF file for all model terms (if MISS=1). In this case, the value 9 was specified for all variables since for the Onset variable this value is the correct missing value code, while for all other model terms (intercpt, AlcOnset, SexMale, and AlcSex) this value was never observed.

The last two lines of the DEF files list the coefficients for the two linear combinations of the estimated parameters that are requested. Noting the order of parameters that the program assumes (discussed above), in both DEF files, the first linear combination is simply the sum of the intercept and first threshold, while the second linear combination is the sum of the intercept and the second threshold. In both DEF files, the next two linear combinations indicate the sum of the AlcOnset effect with each of the two AlcOnset by grouped-time interaction terms. These two re-expressions provide estimates of the AlcOnset effect at each of the latter two timepoints, respectively, while the AlcOnset effect relects the difference in substance onset at the first timepoint. The last two linear combinations included in the second DEF file provides similar estimates for the SexMale effect.

Table 7 lists the results from the model allowing for varying effects across timepoints for substance and sex (the model specified by the DEF file in Table 6b).

Insert Table 7 about here

Following the descriptive information provided by MIXGSUR, the results indicate a marginally significant substance by sex interaction, and significant "main effects" of AlcOnset and SexMale. Due to the coding of variables and the interactions present in this model, the AlcOnset effect represents the increased hazard for alcohol, relative to cigarettes, at the first timepoint for females. Since the AlcSex term is negative and marginally significant, there is some suggestion that this increased hazard for alcohol relative to cigarettes is not as pronounced for males. The significant SexMale effect in this model represents the increased hazard in terms of cigarette onset for males, relative to females, at the first timepoint. Again, due to the marginal significance of the AlcSex term, there is some suggestion that this difference between males and females is more pronounced for cigarettes than for alcohol.

Of particular interest in this analysis are the interactions with the thresholds for AlcOnset and SexMale. For both, the second of these interaction terms are significant, whereas the first are not. These terms represent differences from the first timepoint, and so suggest that the effects of AlcOnset and SexMale differ significantly between the first and last timepoint. Based on the sign of these interaction terms, it appears that the increased hazard for alcohol relative to cigarettes is more pronounced at the last timepoint than the first. Interestingly, the SexMale interactions are negative, indicating that the significant difference between males and females at the first timepoint largely dissipates by the last timepoint. This latter observation is confirmed by the non-significance of the last linear transform, which is an estimate of the SexMale effect at the last timepoint.

6 Hardware and software specifications

MIXGSUR is written in standard FORTRAN-77 with double arithmetic precision. It was originally developed for MS-DOS personal computers and later ported over to the MACINTOSH environment. As a result, it's use on MACINTOSH personal computers does not take advantage of the system's menu-orientated interface. For use in either the MS-DOS or MACINTOSH environment, MIXGSUR requires a math coprocessor. The program stores all necessary matrices and vectors in a single one-dimensional array. Thus there are no fixed limitations on the numbers of level-2 units, level-1 units, or model variables. MIXGSUR utilizes some MATCAL subroutines [29] for performing the matrix algebra operations.

7 Availability

The MIXGSUR program is available at no charge. Those interested in obtaining a copy of the program should contact the first author by electronic mail at HEDEKER@UIC.EDU.

8 Acknowledgements

Thanks are in order to Dr. R. Darrell Bock, Dr. Robert D. Gibbons, Dr. Richard Campbell, Dr. Ohid Siddiqui, Dr. Frank Hu, and Dr. Brian Flay for many valuable discussions, comments, and suggestions that contributed in the preparation of this program. In particular, the author is grateful to Dr. Richard Campbell for suggesting the development of the random-effects model within the grouped-time survival data context, to Dr. Bock for providing the code for the MATCAL [29] matrix algebra subroutines, and to Dr. Brian Flay for providing the clustered data-set for the example. The development of MIXGSUR was supported by the National Institutes of Mental Health Grants MH44826-01A2 and MH56146-01.

References

[1] W.A. Thompson, On the treatment of grouped observations in life studies, Biometrics 33, (1977) 463-470.

[2] R.L. Prentice and L.A. Gloeckler, Regression analysis of grouped survival data with application to breast cancer data, Biometrics 34, (1978) 57-67.

[3] P. Allison, Discrete-time methods for the analysis of event histories, in: Sociological Methodolgy, Ed. S. Leinhardt, pp. 61-98 (Jossey-Bass, San Francisco, 1982).

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

[5] T. Lancaster, Econometric methods for the duration of unemployment Econometrica 47, (1979) 939-956.

[6] D. Clayton and J. Cuzick, Multivariate generalizations of the proportional hazards model (with discussion), Journal of the Royal Statistical Society, Series A, 148, (1985) 82-117.

[7] S.G. Self and R.L. Prentice, Incorporating random effects into multivariate relative risk regression models, in: Modern Statistical Methods in Chronic Disease Epidemiology, Eds. S.H. Moolgavkar and R.L. Prentice, pp. 167-178 (Wiley, New York, 1986).

[8] G. Guo and G. Rodriquez, 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, (1992) 969-976.

[9] M.C. Paik, W.-Y. Tsai, and R. Ottman, Multivariate survival analysis using piecewise gamma frailty, Biometrics 50, (1994) 975-988.

[10] J.H. Shih and T. Louis, Assessing gamma frailty models for clustered failure time data, Lifetime Data Analysis 1, (1995) 205-220.

[11] A. Pickles and R. Crouchley, A comparison of frailty models for multivariate survival data, Statistics in Medicine 14, (1995) 1447-1461.

[12] P. Hougaard, Frailty models for survival data, Lifetime Data Analysis 1, (1995) 255-273.

[13] A. Han and J. Hausman, Flexible parametric estimation of duration and competing risk models, Journal of Applied Econometrics 5, (1990) 1-28.

[14] K.-Y. Liang and S.L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73, (1986) 13-22.

[15] L.J. Wei, D.Y. Lin, and L. Weissfeld, Regression analysis of multivariate incomplete failure time data by modeling marginal distributions, Journal of the American Statistical Association 84, (1989) 1065-1073.

[16] S.W. Guo and D.Y. Lin, Regression analysis of multivariate grouped survival data, Biometrics 50, (1994) 632-639.

[17] P. McCullagh, Regression models for ordinal data (with discussion), Journal of the Royal Statistical Society Series B 42, (1980) 109-142.

[18] J.D. Teachman, V.R.A. Call, and K.P. Carver, Marital status and the duration of joblessness among white men, Journal of Marriage and the Family 56, (1994) 415-428.

[19] D. Hedeker O. Siddiqui, and B.F. Hu, Random-effects regression analysis of correlated grouped-time survival data, Biometrics (under review).

[20] D. Hedeker and R.D. Gibbons, A random-effects ordinal regression model for multilevel analysis, Biometrics 50, (1994) 933-944.

[21] H. Goldstein, Multilevel Statistical Models (Halsted Press, New York, 1995).

[22] D.R. Cox, Regression models and life tables (with discussion), Journal of the Royal Statistical Society, Series B, 74, (1972) 187-220.

[23] B. Peterson and F.E. Harrell, Jr., Partial proportional odds models for ordinal response variables, Applied Statistics 39, (1990) 205-217.

[24] R.D. Bock and M. Aitkin, Marginal maximum likelihood estimation of item parameters: an application of the EM algorithm, Psychometrika 46, (1981) 443-459.

[25] R.D. Bock, R.D. Gibbons, and E. Muraki, Full-information item factor analysis, Applied Psychological Measurement 12, (1988) 261-280.

[26] S. Bennett, Analysis of survival data by the proportional odds model, Statistics in Medicine 2, (1983) 273-277.

[27] A.S. Bryk and S.W. Raudenbush, Hierarchical Linear Models: Applications and Data Analysis Methods (Sage Publications, Inc., Newbury Park CA, 1992).

[28] B.R. Flay, B.R. Brannon, C.A. Johnson, *et al.* The television, school and family smoking cessation and prevention project: I. theoretical basis and program development, Preventative Medicine 17, (1989) 585-607.

[29] R.D. Bock and B.H. Repp, MATCAL: Double Precision Matrix Operations Subroutines, (National Educational Resources, Chicago, 1974).

Table 1a
Analysis options and settings specified in MIXGSUR.DEF: lines 1-5

Line 1 - A title of 60 characters

Line 2 - A subtitle of 60 characters

Line 3 - name of input data file. Any legal filename of 80 characters or less can be specified.

Line 4 - name of main output file. Any legal filename of 80 characters or less can be specified.

Line 5 - name of definition file to be saved or retrieved. Any legal filename of 80 characters or less can be specified. Note that a name for this file *must* be specified even in batch processing, although in batch processing nothing is done to this file.

Table 1b
Analysis options and settings specified in MIXGSUR.DEF: line 6

Line 6 - NPR NF R P CONV MAXJ MISS START WT CATYX NQUAD FUNC CEN PARTIAL ADD LINFN

NPR = number of level-2 units whose data will be listed on the screen (*usually set to 1*).

NF = number of fields of data to read from the input data file.

R = number of random effects.

P = number of fixed effects (not including the mean of the random effects).

CONV = convergence criterion (*usually set to .001 or .0001*).

MAXJ = number of ordered grouped-time categories.

MISS = 0 if no missing values are present in the data, or 1 if missing values are present (codes will later be defined).

START = 0 if automatic starting values are to be used, or 1 if user-defined starting values are to be used.

WT = 0 if each 2nd level unit is weighted equally, or 1 for differential weighting.

CATYX = 0 if a crosstab of any variable by the grouped-time variable is not to be done, and 1 if such a crosstab is to be done.

NQUAD = number of quadrature points (per random-effect dimension) to use in the numerical integration (*usually set between 10 and 20 for models with one random effect, and between 5 and 10 for models with multiple random effects*).

FUNC = 0 for the probit, 1 for the logistic, 2 for the complementary log-log, or 3 for the log-log response function.

CEN = 0 for no right-censoring, or 1 to include right-censoring.

PARTIAL = number of P fixed effects to interact with baseline hazard function parameters (these are assumed to be the first PARTIAL variables of the P fixed effects)

ADD = 1 to add the the P fixed effects to the cutpoints (*i.e.*, the baseline hazard function parameters), or -1 to subtract the P fixed effects from the cutpoints (*usually set to 1 for survival analysis and -1 for ordinal regression*).

LINFN = number of linear transforms of the estimated parameters to estimate.

Table 1c
Analysis options and settings specified in MIXGSUR.DEF: lines 7-9

Line 7 - two parameters are to be read on this line: the field of the input data file which contains the (level-2) IDs, followed by the field of the input data file which contains the grouped-time variable.

Line 8 - R parameters are to be read on this line: the field(s) of the input data file which contain(s) the R random effects.

Line 9 - P parameters are to be read on this line: the field(s) of the input data file which contain(s) the P fixed effects.

Table 1d
Analysis options and settings specified in MIXGSUR.DEF: lines after line 9

next line - (*if WT = 1*) - the field of the input data file which contains the weight to be assigned to each level-2 unit.

next line - (*if CEN = 1*) - the field of the input data file which contains the CENSOR variable (coded 0=censor and 1=event).

next line - the MAXJ values of the ordinal grouped-time variable.

next line - (*if CATYX = 1*) - two parameters and a list of values: the field of the input data file which contains the variable that is to be crosstabulated with the grouped-time variable, followed by the number of levels of this variable, and a list of the values for all of these levels.

next line *(if MISS = 1)* - missing value code for the grouped-time variable.

next line *(if MISS = 1)* - R missing value codes for the random-effect variables.

next line *(if MISS = 1)* - P missing value codes for the fixed effects.

next line - an 8 character label for the grouped-time variable.

next line - R labels for the random effects in 8 character width fields.

next line *(if START = 1)* - R starting values for the means of the random effects.

next line - P labels for the covariates in 8 character width fields (a maximum of 10 labels per line).

next line *(if START = 1)* - P starting values for the covariate effects.

next line *(if START = 1)* - ((R $\times$ (R+1)) / 2) starting values for the variance and covariance terms of the random effects given in "packed" form, *e.g.*, for a 2 x 2 covariance matrix, the order of the starting values should be: variance(1), covariance(1,2) and variance(2).

next line - (*if START = 1*) - MAXJ-1 starting values for cutpoints (if CEN=1) or MAXJ-2 starting values for cutpoints (if CEN=0).

final lines - (*if LINFN* $>$ 0) - LINFN by NPAR coefficients for the linear re-expressions of the estimated parameters (note: NPAR = P + (R $\times$ (R+1)/2) + NCUT $\times$ (PARTIAL+1), where NCUT = MAXJ-2 (if CEN=0) or MAXJ-1 (if CEN=1). Each of these LINFN sets of coefficients are multiplied by the "original" parameter estimates according to the following order: the P fixed effects, the (R $\times$ (R+1)/2) unique elements of the random-effect variance-covariance matrix (in packed form), the NCUT cutpoints, and the PARTIAL $\times$ NCUT interactions with the cutpoints (NCUT coefficients for the first of PARTIAL variables, NCUT coefficients for the second of PARTIAL variables, etc.)

Table 2 - Data from example 5.2: first 55 students from 4 classrooms

193101	3	0	1	1	0	0	0
193101	3	0	1	0	0	0	0
193101	1	1	1	1	0	0	0
193101	1	1	1	1	0	0	0
193101	3	0	1	1	0	0	0
193101	3	0	1	0	0	0	0
193101	3	1	1	0	0	0	0
193101	3	0	1	0	0	0	0
193101	3	0	1	1	0	0	0
193101	3	0	1	1	0	0	0
193101	3	0	1	1	0	0	0
193101	1	1	1	0	0	0	0
193101	3	0	1	0	0	0	0
193101	1	0	1	0	0	0	0
193101	3	0	1	0	0	0	0
193101	3	0	1	0	0	0	0
193101	1	1	1	1	0	0	0
193101	3	0	1	0	0	0	0
193101	3	0	1	1	0	0	0
193101	3	0	1	0	0	0	0
193101	1	0	1	0	0	0	0
194101	2	0	1	1	0	0	0
194101	2	0	1	0	0	0	0
194101	3	1	1	0	0	0	0
194101	1	0	1	0	0	0	0
194101	1	1	1	0	0	0	0
194101	3	0	1	0	0	0	0
194101	2	1	1	1	0	0	0
194101	3	0	1	1	0	0	0
194101	2	1	1	0	0	0	0
194101	2	1	1	1	0	0	0
194101	1	1	1	1	0	0	0
194101	2	1	1	0	0	0	0
194102	3	0	1	1	0	0	0
194102	2	1	1	0	0	0	0
194102	1	0	1	0	0	0	0
194102	2	1	1	1	0	0	0
194102	3	0	1	0	0	0	0
194102	3	0	1	0	0	0	0
194102	2	1	1	0	0	0	0
194102	3	0	1	1	0	0	0
194103	3	1	1	1	0	0	0
194103	1	0	1	0	0	0	0
194103	2	0	1	0	0	0	0
194103	1	1	1	1	0	0	0
194103	1	0	1	0	0	0	0
194103	3	1	1	0	0	0	0
194103	3	0	1	1	0	0	0
194103	3	1	1	1	0	0	0
194103	3	1	1	1	0	0	0
194103	2	0	1	0	0	0	0
194103	3	1	1	0	0	0	0
194103	1	0	1	1	0	0	0
194103	2	1	1	1	0	0	0
194103	3	1	1	0	0	0	0

Table 3a
MIXGSUR.DEF file for example 5.2: student-level analysis ignoring clustering

```
--------------------------------------------------------------------------------
TVSFP Onset of SMOKING (Waves B through D) Survival Analysis
0 random and 5 fixed effects
C:\DATA\DRUGONS\SMKEX.RRM
SMCEX0.OUT
SMCEX0.def
1 8  0 5  0.00010 3 1 0 0 0  10  2  1  0  1   2
1 2

4 5 6 7 8
3
1 2 3
9

9 9 9 9 9
SmkOnset

InterceptSexMale CC       TV       CCTV
1 0 0 0 0   1 0
1 0 0 0 0   0 1
--------------------------------------------------------------------------------
```

Table 3b
MIXGSUR.DEF file for example 5.2: students-within-classrooms analysis

```
--------------------------------------------------------------------------------
TVSFP Onset of SMOKING (Waves B through D) Survival Analysis
1 random and 4 fixed effects
c:\data\drugons\smkEX.rrm
smcEX1.out
smcEX1.DEF
1 8  1 4  0.00010 3 1 0 0 0  10  2  1  0  1   2
1 2
4
5 6 7 8
3
1 2 3
9
9
9 9 9 9
SmkOnset
Intercpt
SexMale CC       TV       CCTV
1 0 0 0 0  0  1 0
1 0 0 0 0  0  0 1
--------------------------------------------------------------------------------
```

Table 4

Output file for example 5.2: students-within-classrooms analysis

---

MIXGSUR - The program for mixed-effects grouped-time survival analysis

TVSFP Onset of SMOKING (Waves B through D) Survival Analysis
1 random and 4 fixed effects

Response function: complementary log-log

Covariate(s) and random-effect(s) mean added to thresholds
==> positive coefficient = negative association between regressor
and ordinal outcome

Numbers of observations

Level 2 observations = 134
Level 1 observations = 1556

The number of level 1 observations per level 2 unit are:

```
  21  12   8  14  10  15  15  16   8   2   6   5  13   8   8  14  11  19  13
   4   9   8  10   9   8  18   7  14  11   2   4   8   7  13  12   6  10   6
  15   9  12  17  10  12   9   7   5   6  10  13  13   9  16  13  13  15   4
  14  13   2   6  12   8  13   8  11  20  10  12  15  19  17  13   7  10  11
   6  15   5   5   2   5   1   7  19  23  23  20  17  20  20  13  15  11  12
  13   7   8  19  11  18  13  15  13  10   7  11  21  10   9  14  20  18  22
  16  11  16  14   7  13  17  23  10  16   7  13  13   7  13   5   8  14  13
   9
```

Descriptive statistics for all variables

Variable	Minimum	Maximum	Mean	Stand. Dev.
SmkOnset	1.00000	3.00000	2.07326	0.82316
Intercpt	1.00000	1.00000	1.00000	0.00000
SexMale	0.00000	1.00000	0.47686	0.49963
CC	0.00000	1.00000	0.47879	0.49971
TV	0.00000	1.00000	0.49036	0.50007
CCTV	0.00000	1.00000	0.22686	0.41894
Event	0.00000	1.00000	0.40746	0.49152

Categories of the response variable SmkOnset

Category	Frequency	Proportion
1.00	474.00	0.30463
2.00	494.00	0.31748
3.00	588.00	0.37789

```
Starting values
---------------

mean        -1.016
covariates   0.008 -0.148  0.020  0.238
var. terms   0.406
thresholds   0.985  1.523


==> The number of level 2 observations with non-varying responses
    =      3 (   2.24 percent )



-------------------------------------------------------
* Final Results - Maximum Marginal Likelihood Estimates *
-------------------------------------------------------

Total Iterations =  18
Quad Pts per Dim =  10
Log Likelihood   =    -1592.354
Ridge            =        0.100


Variable        Estimate     Stand. Error            Z          p-value
--------      ------------   ------------   ------------   ------------
Intercpt        -1.70706         0.10619      -16.07503      0.00000  (2)
SexMale          0.05948         0.08381        0.70972      0.47788  (2)
CC               0.12936         0.11224        1.15251      0.24911  (2)
TV               0.09144         0.13642        0.67025      0.50270  (2)
CCTV            -0.16049         0.18569       -0.86427      0.38744  (2)

Random effect variance term (standard deviation)
Intercpt         0.18657         0.09666        1.93013      0.02680  (1)

Thresholds
     1           0.71655         0.04678       15.31830      0.00000  (1)
     2           1.23289         0.05513       22.36445      0.00000  (1)


note: (1) = 1-tailed p-value
      (2) = 2-tailed p-value



Calculation of the intracluster correlation
-------------------------------------------
residual variance = pi*pi / 6 (assumed)
cluster  variance = (0.187 * 0.187) = 0.035

intracluster correlation = 0.035 / (0.035 + (pi*pi/6)) = 0.021
-



Correlation of the Maximum Marginal Likelihood Estimates

                      1         2         3         4         5         6
```

		Intercpt	SexMale	CC	TV	CCTV	VarCov1
1	Intercpt	1.0000					
2	SexMale	-0.4927	1.0000				
3	CC	-0.5651	-0.0002	1.0000			
4	TV	-0.4293	0.0125	0.3786	1.0000		
5	CCTV	0.2691	0.0678	-0.5936	-0.7364	1.0000	
6	VarCov1	-0.0849	0.0576	-0.0271	-0.0605	0.1178	1.0000
7	Thresh1	-0.4911	0.0844	0.2270	0.0663	-0.0636	0.1127
8	Thresh2	-0.5294	0.0583	0.1612	0.0797	-0.0174	0.1493

		7 Thresh1	8 Thresh2
7	Thresh1	1.0000	
8	Thresh2	0.7237	1.0000

----
* Transforms of parameter estimates *

Transpose of the Transform Matrix (parameters by transforms)

		1	2
1	Intercpt	1.0000	1.0000
2	SexMale	0.0000	0.0000
3	CC	0.0000	0.0000
4	TV	0.0000	0.0000
5	CCTV	0.0000	0.0000
6	VarCov1	0.0000	0.0000
7	Thresh1	1.0000	0.0000
8	Thresh2	0.0000	1.0000

Transform	Estimate	Stand. Error	Z	p-value
1	-0.99051	0.09266	-10.68925	0.00000
2	-0.47416	0.09010	-5.26280	0.00000

note: p-values are 2-tailed

Correlation of the MML Transformed Estimates

	1	2
1	1.0000	
2	0.9109	1.0000

Table 5
Data from example 5.3: first 15 subjects (with two observations each)

193101121	3	0	1	1	1	1
193101121	3	0	1	0	1	0
193101210	3	1	1	1	1	1
193101210	3	0	1	0	1	0
193101212	3	0	1	1	1	1
193101212	3	0	1	0	1	0
193101214	3	0	1	1	1	1
193101214	3	0	1	0	1	0
193101225	1	0	1	1	0	0
193101225	1	0	1	0	0	0
193101302	1	1	1	1	0	0
193101302	3	0	1	0	0	0
193101306	2	1	1	1	0	0
193101306	3	0	1	0	0	0
193101315	1	1	1	1	1	1
193101315	1	1	1	0	1	0
193101319	3	0	1	1	0	0
193101319	3	0	1	0	0	0
193101322	3	0	1	1	1	1
193101322	3	0	1	0	1	0
193101326	1	1	1	1	0	0
193101326	3	0	1	0	0	0
193101327	1	1	1	1	0	0
193101327	1	0	1	0	0	0
194101112	2	1	1	1	0	0
194101112	2	0	1	0	0	0
194101205	1	0	1	1	0	0
194101205	1	0	1	0	0	0
194101208	1	0	1	1	0	0
194101208	1	1	1	0	0	0

Table 6a
MIXGSUR.DEF file for example 5.3: random-intercepts model
(Integrated) baseline hazard varies by substance

```
------------------------------------------------------------------------------
TVSFP Onset of Smoking & Etoh (Waves B-D) Survival Analysis
1 random & 3 fixed - Varying Baseline Hazard for Substances
C:\DATA\DRUGONS\SMALEX.RRM
SACEX1.OUT
SACEX1.DEF
1 7  1 3 0.00010 3 1 0 0 0  10  2  1   1  1  4
1 2
4
5 6 7
3
1 2 3
9
9
9 9 9
Onset
Intercpt
AlcOnsetSexMale AlcSex
1  0 0 0   0   1 0  0 0
1  0 0 0   0   0 1  0 0
0  1 0 0   0   0 0  1 0
0  1 0 0   0   0 0  0 1
------------------------------------------------------------------------------
```

Table 6b
MIXGSUR.DEF file for example 5.3: random-intercepts model
(Integrated) baseline hazard varies by substance and sex

```
------------------------------------------------------------------------------
TVSFP Onset of Smoking & Etoh (Waves B-D) Survival Analysis
1 random & 3 fixed - Varying Baseline Hazard for Subs & Sex
C:\DATA\DRUGONS\SMALEX.RRM
SACEX2.OUT
SACEX2.DEF
1 7  1 3 0.00010 3 1 0 0 0  10  2  1   2  1  6
1 2
4
5 6 7
3
1 2 3
9
9
9 9 9
Onset
Intercpt
AlcOnsetSexMale AlcSex
1  0 0 0   0   1 0  0 0  0 0
1  0 0 0   0   0 1  0 0  0 0
0  1 0 0   0   0 0  1 0  0 0
0  1 0 0   0   0 0  0 1  0 0
0  0 1 0   0   0 0  0 0  1 0
0  0 1 0   0   0 0  0 0  0 1
------------------------------------------------------------------------------
```

Table 7
Output file for example 5.3: random-intercepts model
(Integrated) baseline hazard varies by substance and sex

```
--------------------------------------------------------------------------------
 MIXGSUR - The program for mixed-effects grouped-time survival analysis

 TVSFP Onset of Smoking & Etoh (Waves B-D) Survival Analysis
 1 random & 3 fixed - Varying Baseline Hazard for Subs & Sex

 Response function: complementary log-log

 Covariate(s) and random-effect(s) mean added to thresholds
 ==> positive coefficient = negative association between regressor
     and ordinal outcome


 Numbers of observations
 -----------------------

 Level 2 observations =   1002
 Level 1 observations =   2004


 The number of level 1 observations per level 2 unit  are:

    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2    2
```

```
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2   2
    2   2   2   2   2   2   2   2   2   2   2   2   2   2

Descriptive statistics for all variables
----------------------------------------

Variable          Minimum          Maximum            Mean       Stand. Dev.

Onset             1.00000          3.00000          2.06587          0.81220
Intercpt          1.00000          1.00000          1.00000          0.00000
AlcOnset          0.00000          1.00000          0.50000          0.50012
SexMale           0.00000          1.00000          0.44511          0.49710
AlcSex            0.00000          1.00000          0.22255          0.41607
Event             0.00000          1.00000          0.41118          0.49217

Categories of the response variable Onset
--------------------------------------------

Category        Frequency       Proportion

    1.00           599.00          0.29890
    2.00           674.00          0.33633
    3.00           731.00          0.36477

Starting values
---------------

mean        -0.913
covariates   0.239  0.049 -0.085
var. terms   0.406
thresholds   1.044  1.567  0.000  0.000  0.000  0.000

==> The number of level 2 observations with non-varying responses
    =    744 (  74.25 percent )
```

```
-----------------------------------------------------
* Final Results - Maximum Marginal Likelihood Estimates *
-----------------------------------------------------

Total Iterations =   9
Quad Pts per Dim =  10
Log Likelihood   =    -1932.599
Ridge            =        0.000


Variable        Estimate     Stand. Error          Z          p-value
--------      ------------   ------------   ------------   ------------
Intercpt        -2.87732       0.16039      -17.93960      0.00000  (2)
AlcOnset         0.75070       0.14601        5.14133      0.00000  (2)
SexMale          0.40969       0.18410        2.22535      0.02606  (2)
AlcSex          -0.31281       0.16111       -1.94156      0.05219  (2)

Random effect variance term (standard deviation)
Intercpt         1.32152       0.09026       14.64052      0.00000  (1)

Thresholds
     1           1.15491       0.11096       10.40838      0.00000  (1)
     2           2.02279       0.13782       14.67735      0.00000  (1)

Interaction of Thresholds by  AlcOnset
     1           0.13212       0.12445        1.06166      0.28839  (2)
     2           0.31463       0.15336        2.05153      0.04021  (2)

Interaction of Thresholds by  SexMale
     1          -0.17031       0.11358       -1.49944      0.13376  (2)
     2          -0.35856       0.14373       -2.49468      0.01261  (2)


note: (1) = 1-tailed p-value
      (2) = 2-tailed p-value



Calculation of the intracluster correlation
-------------------------------------------
residual variance = pi*pi / 6 (assumed)
cluster  variance = (1.322 * 1.322) = 1.746

intracluster correlation = 1.746 / (1.746 + (pi*pi/6)) = 0.515



Correlation of the Maximum Marginal Likelihood Estimates

                      1        2        3        4        5        6
                   Intercpt AlcOnset SexMale  AlcSex   VarCov1  Thresh1
    1   Intercpt    1.0000
    2   AlcOnset   -0.6649   1.0000
    3   SexMale    -0.6081   0.3539   1.0000
    4   AlcSex      0.3533  -0.5082  -0.5790   1.0000
    5   VarCov1    -0.4464   0.1385   0.0183  -0.0533   1.0000
    6   Thresh1    -0.6126   0.4612   0.3188  -0.0583   0.1464   1.0000
    7   Thresh2    -0.7071   0.5483   0.3582  -0.1311   0.2818   0.7329
```

8	ThrInt1	0.2856	-0.5566	-0.0407	-0.0167	0.0951	-0.6310
9	ThrInt2	0.2763	-0.6003	-0.0495	0.0018	0.1422	-0.4376
10	ThrInt3	0.3370	-0.0971	-0.4991	0.0873	0.0147	-0.5781
11	ThrInt4	0.3818	-0.1476	-0.5661	0.1761	-0.0320	-0.4034

		7	8	9	10	11
		Thresh2	ThrInt1	ThrInt2	ThrInt3	ThrInt4
7	Thresh2	1.0000				
8	ThrInt1	-0.4259	1.0000			
9	ThrInt2	-0.5529	0.6952	1.0000		
10	ThrInt3	-0.4082	0.0542	0.0535	1.0000	
11	ThrInt4	-0.5653	0.0363	0.0275	0.7047	1.0000

----
* Transforms of parameter estimates *

Transpose of the Transform Matrix (parameters by transforms)

		1	2	3	4	5	6
1	Intercpt	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
2	AlcOnset	0.0000	0.0000	1.0000	1.0000	0.0000	0.0000
3	SexMale	0.0000	0.0000	0.0000	0.0000	1.0000	1.0000
4	AlcSex	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	VarCov1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	Thresh1	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
7	Thresh2	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000
8	ThrInt1	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000
9	ThrInt2	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000
10	ThrInt3	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
11	ThrInt4	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Transform	Estimate	Stand. Error	Z	p-value
1	-1.72242	0.12741	-13.51864	0.00000
2	-0.85453	0.11601	-7.36604	0.00000
3	0.88283	0.12875	6.85666	0.00000
4	1.06533	0.13400	7.95046	0.00000
5	0.23939	0.16100	1.48684	0.13706
6	0.05113	0.15683	0.32604	0.74440

note: p-values are 2-tailed

Correlation of the MML Transformed Estimates

	1	2	3	4	5	6
1	1.0000					
2	0.7037	1.0000				
3	-0.6774	-0.4112	1.0000			
4	-0.5125	-0.6065	0.6394	1.0000		
5	-0.6137	-0.4881	0.3732	0.3447	1.0000	
6	-0.4541	-0.6191	0.3038	0.2677	0.7914	1.0000
