Some applications of generalized linear latent and mixed models in epidemiology:
Repeated measures, measurement error and multilevel modeling

Anders Skrondal[1] and Sophia Rabe-Hesketh[2]

[1] *Biostatistics Group, Division of Epidemiology, Norwegian Institute of Public Health*

[2] *Department of Biostatistics and Computing, Institute of Psychiatry, King's College London*

E-mail: anders.skrondal@fhi.no `http://www.gllamm.org`

ABSTRACT

We describe generalized linear latent and mixed models (GLLAMMs) and illustrate their potential in epidemiology. GLLAMMs include many types of multilevel random effect, factor and structural equation models. A wide range of response types are accommodated including continuous, dichotomous, ordinal and nominal responses as well as counts and survival times. Multivariate responses can furthermore be of mixed types. The utility of GLLAMMs is illustrated in three applications involving repeated measurements, measurement error and multilevel data.

INTRODUCTION

In this article we describe generalized linear latent and mixed models (GLLAMMs) and illustrate their potential in epidemiology.

We begin by briefly describing 'generalized linear models' (1) which encompass common epidemiological tools such as linear regression, dichotomous logistic regression and Poisson regression. Subsequently, we outline 'generalized linear mixed models' (2) where random effects are included in generalized linear models. These models are called mixed since both fixed and random effects are incorporated. Although generalized linear mixed models are undoubtedly useful, it turns out that this framework is too limited for some applications, including many in epidemiology.

This was our motivation (3,4) for extending generalized linear mixed models to the class of 'generalized linear latent and mixed models' (GLLAMM). GLLAMMs include many types of latent variables varying at different levels such as random effects, common factors and latent classes. Latent variables can be regressed on covariates and other latent variables. A wide range of response types are accommodated including continuous, dichotomous, ordinal and nominal responses as well as counts and survival times (discrete and continuous). Multivariate responses can furthermore be of mixed type, some responses may for instance be continuous and others dichotomous.

The utility of GLLAMMs is illustrated in three applications: logistic regression for repeated measurement data (course of illness in schizophrenia), logistic regression with covariate measurement error (diet and coronary heart disease), and multilevel modelling of nominal data (abuse of antibiotics).

GENERALIZED LINEAR MODELS

Let y_i be the response and $\mathbf{x}_i$ explanatory variables or covariates for unit i, and define the conditional expectation of the response given the covariates as μ_i, i.e. $\mu_i \equiv \mathrm{E}[y_i|\mathbf{x}_i]$.

Generalized linear models are specified as

$$g(\mu_i) \;=\; \mathbf{x}_i'\boldsymbol{\beta} = \nu_i, \qquad (1)$$

where the linear combination $\nu_i = \beta_0 + \beta_1 x_{i1} + \cdots = \mathbf{x}_i'\boldsymbol{\beta}$ is called the 'linear predictor' and $\boldsymbol{\beta}$ are fixed effects. g is a 'link function', linking the expected response μ_i to the linear predictor ν_i. The specification is completed by choosing a conditional distribution for the responses y_i given the covariates $\mathbf{x}_i$, $f(y_i|\mathbf{x}_i)$, from the exponential family of distributions.

Links and distributions

Some of the links and distributions that can be combined for generalized linear models are:

Links:
identity
reciprocal
logarithm
logit
probit
compl. log-log

Distributions:
normal
gamma
Poisson
binomial

Common special cases

Linear regression

Continuous outcomes are sometimes encountered in epidemiology, for instance birthweight in perinatal epidemiology. The conventional linear regression

model for continuous responses results from combining the identity link, $g(\mu_i) = \mu_i$, with the normal distribution:

$$\mu_i = \nu_i, \tag{2}$$

$$y_i|\mathbf{x}_i \sim \mathrm{N}(\nu_i, \sigma^2). \tag{3}$$

Logistic regression

Dichotomous responses are legion in epidemiology, the archetypical example being disease (1:present, 0:absent). The expectation of a dichotomous response y_i is just the probability that $y_i = 1$, i.e. $\mu_i = \Pr(y_i = 1|\mathbf{x}_i)$. The logistic regression model for dichotomous responses results from combining the bernoulli distribution (binomial with $N = 1$) with the logit link:

$$\ln\left(\frac{\mu_i}{1-\mu_i}\right) = \nu_i.$$

An equivalent way of formulating the logistic regression model is often useful. This approach, which appears to be unfamiliar to most epidemiologists, express logistic regression as a latent response model (4). Here, we consider a linear regression model for a continuous latent response or underlying variable y_i^* (called 'liability' in genetics),

$$y_i^* = \nu_i + \epsilon_i,$$

where the residual ϵ_i has a logistic distribution with zero mean and variance $\pi^2/3$. The dichotomous observed response y_i is simply related to the continuous latent response y_i^* via a threshold model

$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0 \\ 0 & \text{otherwise.} \end{cases}$$

We cannot resist pointing out that some epidemiologists tend to dichotomize any outcome variable, a practice that should be avoided. It should also be noted that other approaches than logistic regression, such as linear risk models with an identity link, have been advocated for dichotomous responses in epidemiology (5).

Poisson regression

Epidemiologists using cohort designs are often interested in studying the incidence rates of diseases or other outcomes. The conventional Poisson regression model for counts and rates results from combining the log link with the Poisson distribution:

$$\ln \mu_i = \nu_i,$$

$$\Pr(y_i \,|\, \mathbf{x}_i) = \frac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!}.$$

GENERALIZED LINEAR MIXED MODELS

A crucial assumption of generalized linear models is that the responses of different units i are independent given the covariates $\mathbf{x}_i$. Unfortunately, this assumption is often unrealistic since data are frequently of a multilevel nature with units i nested in clusters j. Examples of such two-level designs include repeated measurements (units) nested in subjects (clusters) or subjects (units) nested in families (clusters). There will often be *unobserved heterogeneity* at the cluster level representing confounders that are omitted either because they cannot be measured or because their existence is unknown. The unobserved heterogeneity induces dependence among the units, even after controlling for observed heterogeneity (covariates) at the unit and cluster levels.

The combined effect of all unobserved cluster-level covariates is modeled by including random effects $eta_{mj}^{(2)}$ in the linear predictor which take on the same value for all units in the same cluster

$$g(\mu_{ij}) = \nu_{ij} = \underbrace{\mathbf{x}_{ij}'\boldsymbol{\beta}}_{\text{Fixed part}} + \underbrace{\sum_{m=0}^{M-1} \eta_{mj}^{(2)} z_{mij}^{(2)}}_{\text{Random part}}. \tag{4}$$

Here, $\mu_{ij} \equiv \mathrm{E}[y_{ij}|\mathbf{x}_{ij}, \mathbf{z}_{ij}^{(2)}, \boldsymbol{\eta}_j^{(2)}]$ where $\boldsymbol{\eta}_j^{(2)} = (\eta_{0j}^{(2)}, \cdots, \eta_{M-1,j}^{(2)})'$ are random effects varying at level 2 and $\mathbf{z}_{ij}^{(2)}$ corresponding covariates. Specifically, $\eta_{mj}^{(2)}$ is a random effect of covariate $z_{mij}^{(2)}$ for cluster j. It is typically assumed that the random effects are multivariate normal,

$$\boldsymbol{\eta}_j^{(2)} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Psi}^{(2)}).$$

Common special cases

Random intercept model

The simplest generalized linear mixed model is the random intercept model where $M = 1$ and $z_{0ij}^{(2)} = 1$,

$$\nu_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + \eta_{0j}^{(2)}. \tag{5}$$

Here, $\eta_{0j}^{(2)}$ is a random intercept, allowing the overall level of the linear predictor to vary between clusters j over and above the variability explained by the covariates $\mathbf{x}_{ij}$.

Random coefficient model

Letting $M = 2$, $z_{0ij}^{(2)} = 1$ and $z_{1ij}^{(2)}$ be a unit-specific covariate, we obtain a model with a random coefficient in addition to a random intercept

$$\nu_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + \eta_{0j}^{(2)} + \eta_{1j}^{(2)} z_{1ij}^{(2)}. \tag{6}$$

The model could of course include several random coefficients. Usually $x_{1ij}^{(2)} = z_{1ij}^{(2)}$, implying that β_1 represents the mean effect and $\eta_{1j}^{(2)}$ cluster-specific random deviations from this mean effect. While random intercepts can be thought of as main effects of

omitted covariates, random coefficients represent interactions between omitted and included covariates. Here, $\eta_{1j}^{(2)}$ is a random coefficient, allowing the effects of $z_{1ij}^{(2)}$ to vary between clusters j. In repeated measurement modelling, $z_{1ij}^{(2)}$ would typically represent time, so that the rate of change of the response over time (the slope) can vary between subjects. The model can be further extended to include quadratic components $(z_{1ij}^{(2)})^2$ and higher order polynomials.

Truly multilevel models

Data are often 'truly' multilevel with more than two hierarchical levels. For instance, children i (level 1) may be nested in doctors j (level 2) who are nested in hospitals k (level 3), as in an application considered later. In the repeated measurement setting we could have measurement occasions i (level 1) nested in subjects j (level 2) who are nested in families k (level 3). We would expect dependence between different subjects in the same family, which can be modeled by introducing family-level random effects. However, the dependence should be even higher among different responses for the same subject, which can be modeled by using subject-level random effects (in addition to the family-level effects).

A generalized linear mixed model for three-level data can be written as

$$g(\mu_{ijk}) = \nu_{ijk} = \tag{7}$$
$$\underbrace{\mathbf{x}'_{ijk}\boldsymbol{\beta}}_{\text{Fixed part}} + \underbrace{\sum_{m_2=0}^{M_2-1}\eta_{m_2jk}^{(2)}z_{m_2ijk}^{(2)}}_{\text{Level-2 random part}} + \underbrace{\sum_{m_3=0}^{M_3-1}\eta_{m_3k}^{(3)}z_{m_3ijk}^{(3)}}_{\text{Level-3 random part}}.$$

Here, $\mu_{ijk} \equiv \mathrm{E}[y_{ijk}|\mathbf{x}_{ijk}, \mathbf{z}_{ijk}^{(2)}, \mathbf{z}_{ijk}^{(3)}, \boldsymbol{\eta}_{jk}^{(2)}, \boldsymbol{\eta}_{k}^{(3)}]$ where $\boldsymbol{\eta}_{jk}^{(2)} = (\eta_{0jk}^{(2)}, \cdots, \eta_{M_2-1,jk}^{(2)})'$ are level-2 random effects with corresponding covariates $\mathbf{z}_{ijk}^{(2)}$ and $\boldsymbol{\eta}_{k}^{(3)} = (\eta_{0k}^{(3)}, \cdots, \eta_{M_3-1,k}^{(3)})'$ are level-3 random effects with corresponding covariates $\mathbf{z}_{ijk}^{(3)}$.

The random effects at each level are multivariate normal,

$$\begin{aligned}\boldsymbol{\eta}_{jk}^{(2)} &\sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Psi}^{(2)}),\\ \boldsymbol{\eta}_{k}^{(3)} &\sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Psi}^{(3)}),\end{aligned}$$

and assumed to be independent across levels. If required, the model can easily be extended to four or more hierarchical levels.

GENERALIZED LINEAR LATENT AND MIXED MODELS (GLLAMMs)

Although generalized linear mixed models are very useful, the framework is too limited for many problems in epidemiology. GLLAMMs therefore provide five extensions to generalized linear mixed models (where we refer to $\boldsymbol{\eta}$ as latent variables, including random effects, factors, etc.):

1. Multilevel factor structures
2. Multilevel structural equations, regressing latent variables on other latent variables and covariates
3. Discrete latent variables
4. Additional response types
5. Responses of mixed types

Multilevel factor structures

In equation (4) each random effect multiplies a single covariate. GLLAMMs allow each random effect to multiply a linear combination of covariates. Consider a two-level model

$$g(\mu_{ij}) = \nu_{ij} = \underbrace{\mathbf{x}'_{ij}\boldsymbol{\beta}}_{\text{Fixed part}} + \underbrace{\sum_{m=0}^{M-1}\eta_{mj}^{(2)}\boldsymbol{\lambda}_m^{(2)\prime}\mathbf{z}_{mij}^{(2)}}_{\text{Random part}}, \tag{8}$$

where $\mathbf{z}_{mij}^{(2)}$ is a vector of covariates with corresponding vector of coefficients, called factor loadings, $\boldsymbol{\lambda}_m^{(2)}$. For identification, the first coefficient $\lambda_{m1}^{(2)}$, is set to one. Note that the model reduces to a generalized linear mixed model if $\mathbf{z}_{mij}^{(2)}$ is a scalar for all m.

This extension of the generalized linear mixed model allows factor models to be incorporated in multilevel models. Here (some of) the level-one units are the response variables of the factor model and the $\mathbf{z}_{mij}^{(2)}$ are dummy variables that assign factor loadings to the appropriate responses.

The basic idea of factor models is that one or more unobserved variables, latent traits or factors 'explain' the dependence between different observed measurements for a subject, in the sense that the measurements are conditionally independent given the factor(s). We now consider a number of subjects j providing yes/no responses to a set of measurements or 'items' i, for instance to ascertain asthma among children with symptoms such as wheezing, heavy breathing and coughing at night. By regarding the symptoms as nested in children, a logit factor model relating the probability of having a symptom to a single factor $\eta_{1j}^{(2)}$, interpretable as the childrens' unobserved severity of asthma, can be defined as

$$\begin{aligned}\mathrm{logit}(\mu_{ij}) &= \mathbf{x}'_{ij}\boldsymbol{\beta} + \eta_{1j}^{(2)}\boldsymbol{\lambda}_1^{(2)\prime}\mathbf{z}_{1ij}^{(2)}\\ &= \beta_i + \eta_{1j}^{(2)}\lambda_{1i}^{(2)},\end{aligned} \tag{9}$$

where $\mathbf{x}_{ij}$ and $\mathbf{z}_{1ij}^{(2)}$ are vectors whose ith element equals 1 and all other elements equal 0. The 'factor' $\eta_{1j}^{(2)}$ represents the 'true asthma' of the jth child, β_i reflects the relative prevalence of the ith symptom and $\lambda_{1i}^{(2)}$, the factor loading, reflects how well the ith symptom discriminates between children with different severities of asthma. This model is also

known as a two-parameter *item response model*, 'two-parameter' referring to β_i and $\lambda_{1i}^{(2)}$ for each item i. If there are several children nested in families, further latent variables could be added at the family level (level 3). Another example where factor models are useful is for repeated fallible measurements of nutrients, an application we will return to in a later section.

A more general 'two-level' (here three-level due to variables being considered level 1 units) factor model would allow different factor structures at the two levels. Such a model has been used for attitudes to abortion (3), where questionnaire items were nested in subjects who were nested in polling districts. Note that the practice of substituting different kinds of 'scores' for factors in multilevel modeling should be abandoned since this will in general preclude valid statistical inference (6).

Multilevel structural equations

As an example of the use of structural equations, consider the item response model in equation (9). If there are several children nested in families k, 'true asthma' may depend on child-specific covariates (e.g. age, w_{1jk}) and family-level covariates (e.g. presence of a pet, w_{2k}), and there may be residual heterogeneity between families represented by $\eta_{1k}^{(3)}$, for instance of a genetic nature. Using indices i, j, k for symptoms, children and families, respectively, we can write the 'measurement model' as

$$\text{logit}(\mu_{ijk}) = \mathbf{x}'_{ijk}\boldsymbol{\beta} + \eta_{1jk}^{(2)}\boldsymbol{\lambda}_1^{(2)\prime}\mathbf{z}_{1ijk}^{(2)}, \qquad (10)$$

and add the 'structural equation model'

$$\eta_{1jk}^{(2)} = \eta_{1k}^{(3)} + \gamma_{11}w_{1jk} + \gamma_{12}w_{2k} + \zeta_{1jk}^{(2)}. \qquad (11)$$

In general, GLLAMMs allow latent variables (random coefficients and/or factors) to be regressed on other latent variables (random coefficients and/or factors) and covariates (3,4). It is important to note that GLLAMMs extend conventional structural equation models (7) by permitting latent variables to be regressed on same or higher level latent variables.

Discrete latent variables

For two-level models, the latent variables can have a discrete distribution with non-zero probability at a finite number of points (of dimensionality equal to the number of random effects). This is useful if the level 2 units are believed to fall into a number of groups or 'latent classes' within which the latent variables do not vary (8). For instance, such a formulation seems natural for medical diagnosis of say myocardial infarction where it is presumed that the patient is either ill or not ill, but the physician must resort to several fallible indicators of disease. For the case of depression, however, it might be more natural to view the disease as having different degrees of severity, consistent with a continuous illness distribution.

If the number of latent classes, or masses, is chosen to maximize the likelihood, the nonparametric maximum likelihood estimator (NPMLE) can be achieved (9,10). Importantly, this approach enables us to relax the assumption of multivariate normal latent variables and thus makes our inferences more robust. We consider the use of NPMLE in our covariate measurement error application.

Additional response types

In addition to the response types typically considered in generalized linear mixed models, GLLAMMs accommodate several additional response types and links.

An ordinal response is one among a set of categories with a common a priori ordering. An example is severity of illness in terms of the categories 'absent', 'mild', 'moderate' and 'severe'. The following links can be used for ordinal responses:

	Ordinal responses:
family:	ordinal logit
	ordinal probit
	ordinal compl. log-log
	scaled ord. probit

Nominal responses are in terms of categories which do not have a common a priori ordering. A polytomous response is one among a set of unordered categories. For instance, the subject may be asked which contraceptive was used (if any) during the last sexual intercourse. Rankings involve ordering of categories, for instance a patient's preference ordering of different drugs for a specific condition. The following link is available for nominal responses:

Polytomous & Rankings:
multinomial logit

Responses of mixed types

Different links and families can be combined for multivariate responses of mixed type. One example would be the modelling of asthma, where some symptoms could be ascertained as dichotomous (present/absent), some as ordinal (severe/mild/absent) and others as counts (number of episodes). Handling mixed responses is also necessary for modeling joint processes, for instance joint modeling of survival and repeated measurements (11) or hospital delivery and child mortality (12). The possibility of specifying models with mixed responses also turns out to be required for our covariate measurement error application.

IMPLEMENTATION

All models in the GLLAMM framework can be estimated using the program gllamm (13). This program, written in Stata (14), implements maximum likelihood estimation and empirical Bayes prediction for GLLAMMs. Numerical integration by adaptive Gauss-Hermite quadrature (15) is used to integrate

out the latent variables and obtain the marginal log-likelihood. This log-likelihood is maximized by Newton-Raphson using numerical first and second derivatives. Empirical Bayes predictions are posterior means of the latent variables given the observed responses with the parameter estimates plugged in. These predictions have many uses, for instance in predicting 'true' exposure in covariate measurement error models, plotting growth trajectories for individual units in longitudinal data and in model diagnostics. Both posterior means and posterior standard deviations are obtained by numerical integration using adaptive quadrature.

Monte Carlo experiments (16) have been performed to investigate our maximum likelihood methodology. The performance of adaptive quadrature was good in all cases, larger numbers of quadrature points being required for more difficult situations (17). Comparing gllamm with other software (using e.g. IGLS, PQL, MCMC and quadrature) good agreement was found between parameter estimates, standard errors and log-likelihood values (13,15). The exceptions were cases where simulation or parametric bootstrapping demonstrated that PQL produced biased estimates in contrast to quadrature (13,18,19).

MISSING DATA

The GLLAMM framework treats the variables of a multivariate response as level-1 units, thereby automatically handling unbalanced designs, for instance due to missing data. Maximum likelihood produces valid inferences when data are missing at random (MAR), where the probability of a response being missing (given the covariates) may depend on other observed responses but not on missing responses (20). Importantly, the stronger assumption of data missing completely at random (MCAR), where the probability of missingness (given the covariates) neither depends on observed nor missing responses, need not be invoked. Furthermore, missing data that are not MAR can be handled in GLLAMM by explicitly modeling the missingness mechanism (21).

SOME APPLICATIONS IN EPIDEMIOLOGY

We describe some applications of GLLAMMs in order to illustrate the potential of this methodology for epidemiological research. However, we do not purport to exhaust the types of applications that are likely to be useful in epidemiology. Furthermore, it should be appreciated that our applications are simplified for didactic reasons.

Logistic regression for repeated measurements: Course of illness in schizophrenia

The Madras Longitudinal Schizophrenia Study (22) followed up 86 patients monthly after their first hospitalization for schizophrenia. An important question is whether the course of illness differs between men and women and between patients with early and late onset. Here we consider a subset of data previously analyzed (23), namely data on whether thought disorder was present or not at 0, 2, 6, 8 and 10 months after hospitalization.

Performing a complete-case or 'listwise' analysis would discard the information from 16 patients contributing a total of 45 responses (out of the scheduled 96). This will obviously lead to reduced power but more importantly produce biased estimates unless the measurements are missing completely at random (MCAR).

The variables are:

- [y]: dummy variable for thought disorder (1: present, 0: absent)
- [Month]: number of months since first hospitalization
- [Man]: dummy for patient being a man (1: man, 0: woman)
- [Early]: dummy for early onset (1: before age 20, 0: at age 20 or later)

Following the approach adopted in (23), we will estimate a dichotomous logistic regression model with thought disorder [y] as response and explanatory variables [Month], [Man], [Early] and the interactions [Man]×[Month] and [Early]×[Month]. This model allows us to investigate the linear trend (on the logit scale) of time as well as differences between genders and between times of onset, not just in the overall odds of thought disorder but also in the trend over time.

Let y_{ij} be the repeated measurement of thought disorder at occasion i for patient j. To model the dependence among the repeated measurements of thought disorder (given the covariates) we can include a patient-specific random intercept η_{0j}, giving the model

$$\ln\left(\frac{\Pr(y_{ij}=1|\mathbf{x}_{ij})}{1-\Pr(y_{ij}=1|\mathbf{x}_{ij})}\right) \;=\; \mathbf{x}'_{ij}\boldsymbol{\beta}+\eta_{0j},$$

where η_{0j} is normally distributed with zero mean. The inclusion of η_{0j} allows the overall logit of thought disorder to vary over patients, even after controlling for the covariates $\mathbf{x}_{ij}$ (due to omitted patient-specific covariates).

The random intercept logistic regression model can equivalently be expressed as a latent response model,

$$y^*_{ij} \;=\; \mathbf{x}'_{ij}\boldsymbol{\beta}+\eta_{0j}+\epsilon_{ij},$$

where the random intercept η_{0j} is now included. The dichotomous observed response is related to the latent response via a threshold model as previously specified for conventional logistic regression. The latent response formulation is useful for investigating

Table 1: Repeated measurements of thought disorder – Estimates for dichotomous logistic regressions with random intercept and with random intercept and random slope for [Month]

	Random intercept model		Random coefficient model	
	Est	(SE)	Est	(SE)
Fixed part				
β_0 [Cons]	0.71	(0.41)	1.00	(0.63)
β_1 [Month]	-0.37	(0.07)	-0.46	(0.12)
β_2 [Man]	-0.25	(0.57)	-0.23	(0.84)
β_3 [Man]×[Month]	-0.14	(0.11)	-0.19	(0.16)
β_4 [Early]	1.19	(0.62)	1.56	(0.91)
β_5 [Early]×[Month]	-0.19	(0.11)	-0.24	(0.17)
Random part				
$\mathrm{var}(\eta_{0j})$	2.60	(0.91)	9.47	(3.59)
$\mathrm{var}(\eta_{1j})$		–	0.18	(0.09)
$\mathrm{cov}(\eta_{1j}, \eta_{0j})$		–	-0.97	(0.52)
Log-likelihood	-217.32		-210.81	

the properties of logistic regression models with latent variables (24,4). For instance, the strength of the residual within-patient dependence can be expressed by the intra-class correlation for the repeated latent responses

$$\rho \;=\; \mathrm{cor}(y^*_{ij}, y^*_{i'j} \,|\, \mathbf{x}_{ij}, \mathbf{x}_{i'j}) = \frac{\mathrm{var}(\eta_{0j})}{\mathrm{var}(\eta_{0j}) + \pi^2/3}.$$

The estimates for the random intercept model are given in Table 1. The odds of thought disorder decrease over time in late-onset women with an estimated odds ratio of $\exp(-0.37) = 0.69$ per month. Early onset patients seem to have a higher odds of thought disorder at first hospitalization (odds ratio $\exp(1.19) = 3.28$). Interestingly, it also appears as if early onset patients have a greater decline in their odds of thought disorder over time. The intra-class correlation is estimated as $\widehat{\rho} = 0.44$, demonstrating that those having a higher than expected (lower) risk of thought disorder on one occasion tend to have a higher (lower) risk at other occasions, taking into account the covariates.

The random intercept model assumes that the logit of thought disorder declines at the same rate over time for all patients with the same covariate values. Since this may be unrealistic, we now allow these rates to vary randomly between patients by including a random slope η_{1j} of [Month] in the model, giving the random coefficient model

$$\ln\left(\frac{\Pr(y_{ij}=1|\mathbf{x}_{ij}, z_{1ij})}{1-\Pr(y_{ij}=1|\mathbf{x}_{ij}, z_{1ij})}\right) \;=\; \mathbf{x}'_{ij}\boldsymbol{\beta}+\eta_{0j}+\eta_{1j}z_{1ij},$$

where z_{1ij} represents [Month]. The random intercept η_{0j} and slope η_{1j} are bivariate normally distributed with zero means.

Estimates for the random coefficient model are reported in Table 1. The change in log-likelihood suggests that the random slope should be included in the logistic regression model. Overall, the fixed effects estimates are quite similar to those for the random intercept model. The random slope variance is estimated as 0.18 and the covariance between intercept and slope as -0.97, corresponding to a correlation of -0.78. Therefore those at higher risk of thought disorder at the time of hospitalization experience a greater reduction in their risk over time than those at lower risk. It is important to note that the random intercept variance and the correlation between the random intercept and coefficient are interpreted at [Month]=0. (Subtracting 5 months from [Month] yields an estimated correlation close to zero).

To gain more insight into the model, we have plotted the conditional or *subject-specific* probabilities of thought disorder given various values of the random intercept (± 3) and slope (± 0.4) for women with early onset. These are shown as dashed curves in Figure 1 where the dotted curve is the conditional probability for random intercept and slope both equal to their population means of zero, thus representing a 'typical' individual. Also shown as a solid bold curve

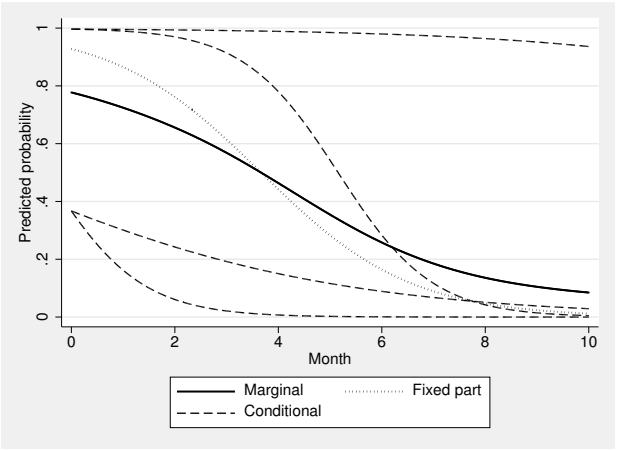

Figure 1: Conditional and marginal predicted probabilities of thought disorder for women with early onset. Dotted curve for conditional probability from random coefficient model when both random intercept and slope are zero.

is the *population average* or marginal probability of thought disorder obtained by integrating the conditional probability over the random effects distribution. Note that the population average curve is considerably flatter than that of a typical patient. Such attenuation of the effects of covariates in marginal models compared with conditional models is a well-known phenomenon for dichotomous responses (25).

Generalized estimating equations (GEE), see e.g. (26), are often used for estimating marginal relationships in longitudinal data. Using GEE with an 'exchangeable' correlation structure, the estimated effect of [Month] becomes -0.26, attenuated com-

pared with −0.37 and −0.46 for the random intercept and random coefficient models, respectively. Figure 2 shows the population average probability curves for GEE and the random intercept and random coefficient models. As might be expected, all

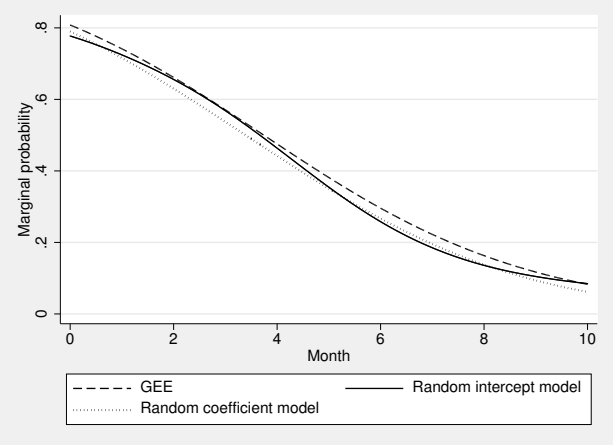

Figure 2: Marginal predicted probabilities of thought disorder for women with early onset from random effects models and GEE.

three marginal curves nearly coincide.

Finally, some comments on random effects modeling versus GEE are in order. GEE is an estimation method for clustered data where the dependence within clusters is treated as a nuisance. The merit of GEE is that valid inferences are produced for population average effects as long as the mean structure is correctly specified, even if the dependence structure is misspecified. It should be noted that random effects models can also be made more robust by using nonparametric maximum likelihood estimation (NPMLE), see the following section.

GEE has a number of severe limitations as compared to random effects modeling that are often not recognized. No insight is gained regarding individual trajectories of change in contrast to random effects modeling. In fact, longitudinal information is not exploited at all since estimation proceeds as if the data were repeated cross-sections (27). It is also evident that causal processes necessarily operate at the subject level so that subject-specific effects are required for etiological inference. In contrast, population average effects are merely descriptive and largely determined by the degree of heterogeneity in the population. Finally, GEE is not based on a statistical model which precludes likelihood based inference and can lead to logical inconsistencies.

Commands for estimating the random effects models in `gllamm` are given in the Appendix. More details on how to obtain the estimates, predictions and plots shown in this section are provided in (28).

Logistic regression with covariate measurement error: Diet and heart disease

We consider data from an investigation of the relationship between diet and coronary heart disease (30). At the time of recruitment, 337 middle-aged men weighed their food intake over a 7-day period, allowing food constituents to be derived. A subsample of 76 of the men repeated this 6 months later, and all the men were then followed up for heart disease. We will estimate the effect of dietary fibre intake on heart disease, controlling for occupation. The relevant variables are:

- [Chd]: dummy variable for coronary heart disease (1: present, 0: absent)
- [Lfibre1]: log of dietary fibre intake (grams/day) at first occasion
- [Lfibre2]: log of dietary fibre intake (grams/day) at second occasion
- [Bus]: dummy for man working for London Transport (1: London Transport, 0: bank staff)

Following Clayton (30), we view covariate measurement error models as composed of three submodels: [1] a disease model, [2] a measurement model and [3] an exposure model. Since we have different response types for unit j; continuous measurements of log dietary fibre intake y_{1j} and y_{2j} and dichotomous coronary heart disease y_{3j}, we need a model handling multivariate responses of mixed types.

As *disease model*, we consider a logistic regression model for [Chd] (y_{3j}) with a latent ('unobserved' or 'true') covariate η_j,

$$\ln\left(\frac{\Pr(y_{3j}=1|x_j)}{1-\Pr(y_{3j}=1|x_j)}\right) = \beta_0+\beta_1 x_j+\eta_j\lambda. \quad (12)$$

One of the covariates, [Bus] (x_j), is perfectly measured or 'observed' and has a *direct effect* β_1 on the logit. The other covariate, 'true log fibre intake' (η_j), is latent and measured with error. The factor loading λ represents the regression coefficient for true log fibre intake.

We specify a *classical measurement model* relating measured log fibre intake y_{ij} to true log fibre intake η_j:

$$y_{ij} = \eta_j+\epsilon_{ij}, \quad (13)$$

where the measurement error ϵ_{ij} has a normal distribution with mean zero. True log fibre intake was measured by y_{1j} for all men at the first occasion. At the second occasion measures y_{2j} were missing for many of the men. The repeated measurements for a man are assumed to have the same mean as the true covariate for that man. The measurement errors ϵ_{ij} are specified as independently normally distributed with zero mean and constant measurement error variance and are independent of the true covariate η_j. It follows that the measurements are conditionally independent of one another given the true

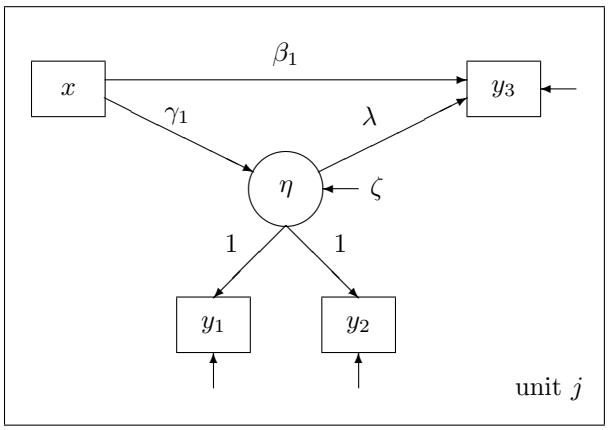

Figure 3: Path diagram with direct and indirect effects of [Bus] on [Chd] via true fibre intake.

Table 2: Diet and CHD – Estimates for dichotomous logistic regression with covariate measurement error based on normal and nonparametric exposure distributions

	Normality		NPMLE	
	Est	(SE)	Est	(SE)
Disease model				
β_0 [Cons]	3.64	(2.02)	3.54	(1.96)
β_1 [Bus]	-0.19	(0.34)	-0.18	(0.34)
λ	-1.96	(0.73)	-1.93	(0.70)
Measurement model				
var(ϵ_{ij})	0.02	(0.00)	0.02	(0.00)
Exposure model				
γ_0 [Cons]	2.86	(0.02)	2.86	(0.02)
γ_1 [Bus]	-0.12	(0.03)	-0.12	(0.03)
var(ζ_j)	0.07	(0.01)	0.07†	(–)
Log-likelihood	-186.93		-177.87	

†Empirical variance of estimated discrete distribution.

covariate. Furthermore, the measurements are conditionally independent of the outcome y_{3j} given the true covariate, a property known as *nondifferential measurement error*.

To complete specification of the covariate measurement error model, we need to define an *exposure model*

$$\eta_j = \gamma_0 + \gamma_1 x_j + \zeta_j, \tag{14}$$

where the covariate [Bus] affects true log fibre intake and the residual or 'disturbance' ζ_j has a normal distribution with mean zero.

The quality of measurements is often quantified in terms of the *reliability* of the measures. Here, we define the reliability as the proportion of the total variance that is due to variability between the units' true covariate values, given the observed covariate x_j:

$$R = \frac{\text{var}(\zeta_j)}{\text{var}(\zeta_j) + \text{var}(\epsilon_{ij})}. \tag{15}$$

The structure of the model is perhaps best conveyed in a path diagram as shown in Figure 3. Here circles represent latent variables and rectangles observed variables. Long arrows represent linear relations in the linear predictor and short arrows represent residual variability. For the exposure and measurement models, this residual variability is represented by an additive error term, but for the disease model it represents bernoulli variability.

Importantly, the covariate measurement error model specifies true log fibre intake as an intermediate variable in the causal pathway from occupation [Bus] to disease [Chd]. It follows that [Bus] may have an *indirect effect* $\gamma_1\lambda$ in addition to the direct effect β_1. The *total effect*, the sum of the direct and indirect effect, then becomes $\beta_1 + \gamma_1\lambda$. Note that true log fibre intake can be viewed as the latent variable in a factor model with mixed responses; continuous responses for the two measurements of log fibre intake and a dichotomous response for disease. The factor loading for [Chd] is λ, whereas the loadings are fixed at one for the fibre measures. Since the latent variable is regressed on a covariate, the specified covariate measurement error model represents a generalization of the conventional *Multiple-Indicator Multiple-Cause* (MIMIC) *model* (31) to include direct effects and non-continuous responses.

Estimates for the logistic regression model with covariate measurement error are presented in Table 2. Estimates for the model presented above are shown under 'Normality' in the table. Here, $\widehat{\lambda} = -1.96$ represents the estimated effect of true log-fibre on coronary heart disease with corresponding odds ratio $\exp(-1.96) = 0.14$. This extremely large estimated protective effect of log-fibre is probably due to omitting important confounding variables such as exercise which is protective of heart disease and increases food intake, including fibre. Note that the 'naive' estimate, treating the first log-fibre measurement as perfect and discarding the second measurement, is -1.63, which is attenuated as expected. True log-fibre has an estimated mean of $\widehat{\gamma}_0 = 2.86$ among bank staff and $\widehat{\gamma}_0 + \widehat{\gamma}_1 = 2.74$ for transport staff.

It is interesting to investigate the direct, indirect and total effects of occupation on CHD. The reduced fibre intake among transport staff results in an increased odds of CHD, the odds ratio for this indirect effect of occupation (not shown in the table) being estimated as 1.27 (95% CI from 1.02 to 1.57). The protective direct effect of being transport staff $\exp(-0.19) = 0.83$ (95% CI from 0.43 to 1.61) counteracts this, the odds ratio for the total effect of occupation (not shown in the table) being estimated as 1.05 (95% CI from 0.55 to 2.01).

The residual variance of true log fibre intake is estimated as $\widehat{\text{var}}(\zeta_j) = 0.07$ and the measurement error variance as $\widehat{\text{var}}(\epsilon_{ij}) = 0.02$. Using equation (15), we obtain an estimated reliability $\widehat{R} = 0.77$.

Instead of assuming a normal distribution for the true covariate η_j, we can make the analysis more robust by letting the distribution be unspecified. The nonparametric maximum likelihood estimator (NPMLE) of the distribution is discrete (32,33) with probability masses at a finite number of locations. The number of masses is determined to achieve largest possible likelihood. NPMLE for covariate measurement error models has been discussed in several papers, e.g. (34,35). NPMLE for the logistic regression model with covariate measurement error required 8 mass points, the resulting estimates are presented under NPMLE in Table 2. We note that the estimates are very similar to those assuming normality, indicating that the normality assumption is tenable for this application.

Although commonly used, the classical measurement error model in equation (13) has a number of limitations. It assumes that the fallible measures have the same mean (no relative bias) and measurement error variance, which is reasonable if the measures are essentially exchangeable replicates. However, if the measurements are separated in time there may be a 'drift' in the mean measurement (36). More importantly, if the fallible measures were obtained by different methods, we should allow the measures to have different means, scales, and measurement error variances. These limitations can be rectified by using a *congeneric measurement model* (37).

There are often no replicate measures available to identify and estimate the covariate measurement error models considered above. In this case it is useful to perform a sensitivity analysis, investigating how the regression estimate for a fallible covariate depends on the magnitude of its measurement error variance. Suppose that we only had measures of log-fibre intake at the first occasion. Figure 4 shows a plot of the estimated odds-ratio $\exp(\beta_u)$ for different assumed values of the measurement error variance $\mathrm{var}(\epsilon_{i1})$. We see that the difference between the

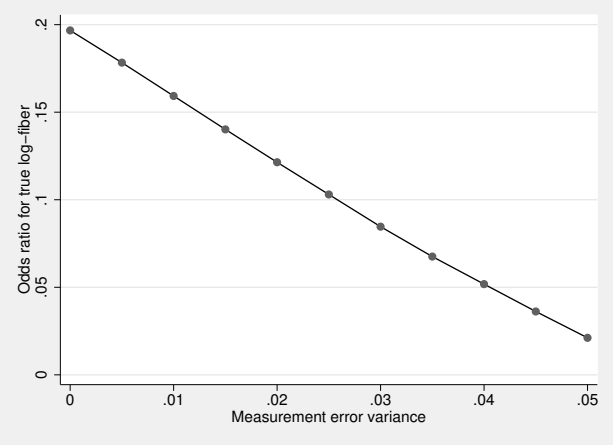

Figure 4: Sensitivity analysis of odds-ratio for different assumed measurement error variances (supposing there are no replicate measurements).

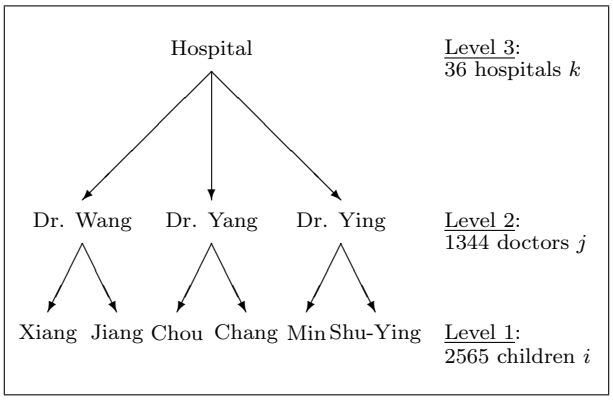

Figure 5: Three-level structure of antibiotics data.

'naive' estimated odds-ratio $\exp(-1.63) = 0.20$ (assuming no measurement error) and the 'corrected' estimated odds-ratio increases as the measurement error variance increases.

The diet and heart disease application has been discussed in more detail elsewhere (35,37), where several extensions of conventional covariate measurement error modelling are also described.

It is demonstrated in the Appendix that the `cme` wrapper for `gllamm` described in (37) makes estimation of the model with a normal true exposure distribution extremely easy. Commands for producing nonparametric maximum likelihood estimates in `gllamm` are also given in (37).

Multilevel modelling of nominal data: Abuse of antibiotics

Acute respiratory tract infection (ARI) is a common disease among children, pneumonia being a leading cause of death in young children in developing countries. In China the standard medication for ARI is antibiotics, which has led to concerns about antibiotics misuse and resultant drug resistance. As a response, the WHO introduced a program of case management for ARI in children under 5 years old in China in the 1990's.

We will analyze data on physicians prescribing behavior of antibiotics in two Chinese counties, one of which was in the WHO program whereas the other was not. These data have previously been analyzed by Yang (38). The multilevel structure of the prescribing data is displayed in Figure 5.

Medical records were examined for medicine prescribed and a 'correct' diagnosis determined from symptoms and clinical signs. Based on the WHO case management criteria, the antibiotic prescription for child i by doctor j in hospital k, y_{ijk}, was classified into three categories (denoted 'abuse' if there were no clinical indications):

1: Correct use of antibiotics (reference-category)

2: Abuse of one antibiotic

3: Abuse of several antibiotics

Prescription behavior is obviously a complex process, and the present analysis is merely intended to shed some light on basic issues such as the impact of the patient's status at arrival and the physician's experience.

We consider the following covariates ($\mathbf{x}_{ijk}$) at different levels:

- Child-level:
  - [Age]: age in years (0-5)
  - [Temp]: body temperature, centered at 36^oC
  - [Paymed]: dummy variable for patient pays for medication (yes=1, no=0)
  - [Selfmed]: dummy for self medication before seeing doctor (yes=1, no=0)
  - [Wrdiag]: dummy for diagnosis classified as wrong (yes=1, no=0)
- Doctor-level:
  - [DRed]: doctor's education (ordinal with six categories from self-taught to medical school)
- Hospital-level:
  - [WHO]: dummy for hospital in WHO program (yes=1, no=0)

Disregarding the multilevel structure of the data for a moment, the natural model for a multicategory response variable is polytomous logistic regression (39). The probability of the realized category, say a, can be written as

$$\Pr(y_{ijk}=a) = \frac{\exp(V^a_{ijk})}{\sum_{b=1}^{3}\exp(V^b_{ijk})}, \qquad (16)$$

where

$$V^a_{ijk} = g^a_0 + \mathbf{x}'_{ijk}\mathbf{g}^a$$

is a linear predictor with category-specific intercepts g^a_0 and category-specific effects of covariates that do not vary over categories $\mathbf{g}^a$. Since the first category serves as reference, $g^1_0 = 0$ and $\mathbf{g}^1 = \mathbf{0}$ for identification.

An alternative specification of the polytomous logistic regression model, based on so-called random utility models, is often used in econometrics and psychometrics but is unfamiliar among epidemiologists. Random utilities U^a_{ijk} are introduced for each category a, where we emphasize that the term utility should be broadly construed as 'attractiveness' of the category in some sense. The utilities are modelled as

$$U^a_{ijk} = V^a_{ijk} + \epsilon^a_{ijk},$$

where ϵ^a_{ijk} is a random term, assumed to be independently distributed across children, doctors, hospitals and categories with a Gumbel distribution

$$g(\epsilon^a_{ijk}) \;=\; \exp\left\{-\epsilon^a_{ijk} - \exp(-\epsilon^a_{ijk})\right\}.$$

The realized category is then viewed as arising from *utility maximization* where the utility of the realized category $U^{a^*}_i$ is larger than the utilities of the other categories. Remarkably, the familiar logistic regression model (16) arises if and only if the random utilities are Gumbel distributed (40).

In an effort to handle the multilevel nature of the prescription problem, we specify a three-level random intercept polytomous regression model

$$V^a_{ijk} = g^a_0 + \gamma^{a(2)}_{0jk} + \gamma^{a(3)}_{0k} + \mathbf{g}^{a\prime}\mathbf{x}_{ijk}. \qquad (17)$$

Comparing this linear predictor with that for the conventional model, we see that the terms $\gamma^{a(2)}_{0jk}$ and $\gamma^{a(3)}_{0k}$ are also included.

$\gamma^{a(2)}_{0jk}$ is a *category-specific* random intercept varying at level 2, letting the overall 'attraction' of category a differ among doctors over and above the variability explained by the included covariates. Retaining the first category as reference, we let the doctor-level intercepts for categories 2 and 3, $\gamma^{2(2)}_{0jk}$ and $\gamma^{3(2)}_{0jk}$, be bivariate normal with zero means, variances $\mathrm{var}(\gamma^{2(2)}_0)$ and $\mathrm{var}(\gamma^{3(2)}_0)$ and covariance $\mathrm{cov}(\gamma^{2(2)}_0, \gamma^{3(2)}_0)$.

Analogously, $\gamma^{a(3)}_{0k}$ is a category-specific random intercept varying at level 3, permitting the overall 'attraction' of category a to differ among hospitals. The hospital-level intercepts, $\gamma^{2(3)}_{0k}$ and $\gamma^{3(3)}_{0k}$, are bivariate normal with zero means, variances $\mathrm{var}(\gamma^{2(3)}_0)$ and $\mathrm{var}(\gamma^{3(3)}_0)$ and covariance $\mathrm{cov}(\gamma^{2(3)}_0, \gamma^{3(3)}_0)$. Thus, the category-specific random intercepts at a given level are dependent whereas the intercepts are specified as independent across levels.

Estimates for the multilevel polytomous regression model, including only covariates at the child level, are presented in Table 3. Estimates also including covariates varying at the doctor and hospital levels are given in Table 4. The change in log-likelihood indicates that inclusion of these covariates improves the fit considerably. Note that the variances of the intercepts are considerably reduced at the hospital level whereas the variances of the intercepts at the doctor level are fairly similar. In general, the fixed effects do not appear to change appreciably.

The estimates for the fixed effects at the child level all seem to have reasonable signs and magnitudes. For instance, the higher the child's temperature, the lower the risk of abuse. Self-medication also appears to reduce the risk of antibiotics abuse, particularly of several antibiotics. On the other hand, a wrong diagnosis increases the risk. We note that doctor's education appears to reduce the risk of abusing several antibiotics but not the risk of abusing one antibiotic. Importantly, the WHO program seems to have a beneficial effect on antibiotics abuse, most pronounced for several antibiotics.

It could be argued that it would be more parsimonious to treat the antibiotic response as ordinal

Table 3: Abuse of antibiotics – Estimates for multilevel random intercept polytomous regression. No observed covariates at the doctor and hospital levels

	Abuse one vs. None		Abuse several vs. None	
	Est	(SE)	Est	(SE)
Fixed Effects				
Child-level				
g_0^a [Cons]	0.23	(0.32)	-1.64	(0.49)
g_1^a [Age]	0.19	(0.08)	0.09	(0.09)
g_2^a [Temp]	-1.01	(0.12)	-0.27	(0.13)
g_3^a [Paymed]	0.30	(0.31)	0.91	(0.41)
g_4^a [Selfmed]	-0.42	(0.24)	-0.78	(0.29)
g_5^a [Wrdiag]	2.08	(0.23)	1.80	(0.26)
Doctor-level				
g_6^a [DRed]		–		–
Hospital-level				
g_7^a [WHO]		–		–
Random Effects				
Doctor-level				
$\text{var}(\gamma_0^{a(2)})$	0.43	(0.27)	0.51	(0.26)
$\text{cov}(\gamma_0^{2(2)}, \gamma_0^{3(2)})$		-0.47 (0.15)		
Hospital-level				
$\text{var}(\gamma_0^{a(3)})$	2.50	(0.93)	0.23	(0.18)
$\text{cov}(\gamma_0^{2(3)}, \gamma_0^{3(3)})$		0.68 (0.31)		
Log-likelihood		-730.6		

Table 4: Abuse of antibiotics – Estimates for multilevel random intercept polytomous regression. Including observed covariates at the doctor and hospital levels

	Abuse one vs. None		Abuse several vs. None	
	Est	(SE)	Est	(SE)
Fixed Effects				
Child-level				
g_0^a [Cons]	-0.23	(0.55)	-5.72	(0.99)
g_1^a [Age]	0.17	(0.08)	0.07	(0.09)
g_2^a [Temp]	-0.96	(0.12)	-0.27	(0.13)
g_3^a [Paymed]	0.12	(0.32)	0.92	(0.40)
g_4^a [Selfmed]	-0.49	(0.24)	-0.86	(0.29)
g_5^a [Wrdiag]	2.08	(0.23)	1.85	(0.26)
Doctor-level				
g_6^a [DRed]	0.08	(0.11)	-0.62	(0.17)
Hospital-level				
g_7^a [WHO]	-0.88	(0.33)	-2.40	(0.62)
Random Effects				
Doctor-level				
$\text{var}(\gamma_0^{a(2)})$	0.43	(0.22)	0.46	(0.28)
$\text{cov}(\gamma_0^{2(2)}, \gamma_0^{3(2)})$		-0.44 (0.13)		
Hospital-level				
$\text{var}(\gamma_0^{a(3)})$	0.11	(0.12)	0.88	(0.45)
$\text{cov}(\gamma_0^{2(3)}, \gamma_0^{3(3)})$		0.31 (0.20)		
Log-likelihood		-716.2		

instead of nominal, since the categories seem to have an a priori ordering (correct, abuse one, abuse several). However, models for ordinal response are considerably more restrictive than their nominal counterparts.

The multilevel polytomous regression model used above is a special case of a general modeling framework for multilevel polytomous regression (41), given a more introductory treatment in (42). The framework includes category-specific covariates such as the cost or other 'attributes' of the categories, useful for instance in health utilization studies. Multilevel and possibly multidimensional common factors can be included as well as different kinds of random coefficients.

Commands for producing the reported estimates in `gllamm` are given in the Appendix.

DISCUSSION

The purpose of this article has been to convey the usefulness of the general and flexible GLLAMM framework in epidemiology. We have considered three applications; logistic regression for repeated measurement data (course of illness in schizophrenia), logistic regression with covariate measurement error (diet and coronary heart disease), and multilevel modelling of nominal data (abuse of antibiotics).

Importantly, the applications considered do not in any way exhaust the potential of GLLAMMs in epidemiology. Types of applications not covered include:

- Multilevel and multivariate survival analysis with frailties, for discrete time (43,4) and continuous time (4)
- Discrete random growth curve models or latent trajectory models (44)
- Disease mapping (4)
- Endogenous treatment models and joint models of survival and repeated measurements (4)
- Genetic epidemiology, for instance investigating association between a genetic marker and the dichotomous phenotype 'atopy' (asthma, eczema or hayfever) (45)

Although the incidence of gllamm use is increasing in epidemiological research, see for instance (46-50), we hope that this article will further enhance the popularity of this powerful research tool.

ACKNOWLEDGEMENTS

We would like to thank David Clayton and Min Yang for kindly providing us with the heart disease and antibiotics data, respectively.

We also acknowledge helpful comments from an anonymous reviewer and Sven Ove Samuelsen, Lars Christian Stene, Hein Stigum and Aage Tverdal.

APPENDIX: ESTIMATION USING STATA PROGRAMS gllamm AND cme

The commands required to perform many of the analyses reported in this article are given here. We refer to (13) and (51) for explanation of the commands.

Course of illness in schizophrenia

```
* Create interactions
gen month_age = month*age
gen month_gen = month*gen

* Random intercept model
gllamm y month age gender month_age month_gen, /*
  */ i(id) link(logit) family(binom) adapt

* Random coefficient model
gen cons = 1
eq inter: cons
eq slope: month
gllamm y month age gender month_age month_gen, /*
  */ i(id) nrf(2) eqs(inter slope)             /*
  */ link(logit) family(binom) adapt
```

Diet and heart disease

```
* Model assuming normal exposure
cme chd bus (lfib: lfibre1 lfibre2), /*
  */ link(logit) family(binom)
```

See (35) for nonparametric maximum likelihood estimation.

Abuse of antibiotics

The data need to be in 'expanded' form with three records per child, one for each response category, a variable `alt` taking on values 1, 2, and 3 for the response categories, and a dummy variable `choice` indicating which response occurred:

```
doc child alt choice age  ...
1   1      1    0     4
1   1      2    1     4
1   1      3    0     4
1   2      1    0     2
1   2      2    0     2
1   2      3    1     2
```

Estimation in `gllamm`:

```
* Create dummy variabes for the categories
gen categ1 = alt == 1
gen categ2 = alt == 2
eq c1: categ1
eq c2: categ2
gllamm alt age temp ... , i(doc hosp)    /*
  */ nrf(2 2) eqs(c1 c2 c1 c2)           /*
  */ link(mlogit) family(binom)          /*
  */ expanded(child choice m) basecat(3)
```

REFERENCES

1. McCullagh P, Nelder JA. *Generalized Linear Models*. London: Chapman & Hall, 1989.

2. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 1993; **88**: 9–25.

3. Rabe-Hesketh S, Skrondal A, Pickles A. Generalized multilevel structural equation modeling. *Psychometrika* 2004; in press.

4. Skrondal A, Rabe-Hesketh S. *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models.*. Boca Raton, FL: Chapman & Hall/ CRC, 2004.

5. Skrondal A. Interaction as departure from additivity in case-control studies: A cautionary note. *American Journal of Epidemiology* 2003; **158**: 251–258.

6. Skrondal A, Laake P. Regression among factor scores. *Psychometrika* 2001; **66**: 563–576.

7. Muthén BO. A general structural equation model with dichotomous, ordered categorical and continuous latent indicators. *Psychometrika* 1984; **49**: 115–132.

8. Clogg CC. Latent class models. In: Arminger G, Clogg CC, Sobel ME (eds.), *Handbook of Statistical Modelling for the Social And Behavioral Sciences*. New York: Plenum Press, 1995; 311–359.

9. Lindsay BG. *Mixture Models: Theory, Geometry and Applications*, volume 5 of *NSF-CBMS Regional Conference Series in Probability and Statistics*. Hayward, CA: Institute of Mathematical Statistics, 1995.

10. Aitkin M. A general maximum likelihood analysis of variance components in generalised linear models. *Biometrics* 1999; **55**: 117–128.

11. Hogan JW, Laird NM. Model-based approaches to analyzing incomplete repeated measures and failure time data. *Statistics in Medicine* 1997; **16**: 259–271.

12. Panis CWA, Lillard L. Health inputs and child mortality: Indonesia. *Journal of Health Economics* 1994; **13**: 455–489.

13. Rabe-Hesketh S, Pickles A, Skrondal A. Gllamm manual. Technical Report 2001/01, Department of Biostatistics and Computing, Institute of Psychiatry, King's College, University of London, 2001. Downloadable from http://www.gllamm.org/manual.pdf.

14. StataCorp. *Stata Statistical Software: Release 8.0*. College Station, TX: Stata Corporation, 2003.

15. Rabe-Hesketh S, Skrondal A, Pickles A. Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal* 2002; **2**: 1–21.

16. Skrondal A. Design and analysis of Monte Carlo experiments: attacking the conventional wisdom. *Multivariate Behavioral Research* 2000; **35**: 137–167.

17. Rabe-Hesketh S, Skrondal A, Pickles A. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. 2003; Submitted for publication.

18. Dohoo IR, Tillard E, Stryhn H, Faye B. The use of multilevel models to evaluate sources of variation in reproductive performance in dairy cattle. *Preventive Veterinary Medicine* 2001; **50**: 127–144.

19. Rabe-Hesketh S, Touloupulou T, Murray RM. Multilevel modeling of cognitive function in schizophrenics and their first degree relatives. *Multivariate Behavioral Research* 2001; **36**: 279–298.

20. Little RJA. Modeling the drop-out mechanism in repeated measures studies. *Journal of the American Statistical Association* 1995; **90**: 1112–1121.

21. Skrondal A, Rabe-Hesketh S, Pickles A. Informative dropout and measurement error in cluster randomised trials. In: *Proceedings of the XXIth International Biometric Conference (Abstracts)*. Freiburg, 2002; 61.

22. Thara R, Henrietta M, Joseph A, Rajkumar S, Eaton W. Ten year course of schizophrenia - the Madras Longitudinal study. *Acta Psychiatrica Scandinavica* 1994; **90**: 329–336.

23. Diggle PJ, Heagerty P, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. Oxford: Oxford University Press, 2002.

24. Rabe-Hesketh S, Skrondal A. Parameterization of multivariate random effects models for categorical data. *Biometrics* 2001; **57**: 1256–1264.

25. Neuhaus JM. Statistical methods for longitudinal and clustered designs with binary responses. *Statistical Methods in Medical Research* 1992; **1**: 249–273.

26. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**: 13–22.

27. Lindsey JK, Lambert P. On the appropriateness of marginal models for repeated measurements in clinical trials. *Statistics in Medicine* 1998; **17**: 447–469.

28. Rabe-Hesketh S, Everitt BS. *Handbook of Statistical Analyses using Stata (3rd Edition)*. Boca Raton, FL: Chapman & Hall/CRC, 2004.

29. Morris JN, Marr JW, Clayton DG. Diet and heart: postscript. *British Medical Journal* 1977; **2**: 1307–1314.

30. Clayton DG. Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In: Dwyer JH, Feinlieb M, Lippert P, Hoffmeister H (eds.), *Statistical Models for Longitudinal Studies on Health*. New York: Oxford University Press, 1992.

31. Jöreskog KG, Goldberger AS. Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association* 1975; **70**: 631–639.

32. Laird NM. Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association* 1978; **73**: 805–811.

33. Heckman JJ, Singer B. A method of minimising the impact of distributional assumptions in econometric models for duration data. *Econometrica* 1984; **52**: 271–320.

34. Roeder K, Carroll RJ, Lindsay BG. A semiparametric mixture approach to case-control studies with errors in covariables. *Journal of the American Statistical Association* 1996; **91**: 722–732.

35. Rabe-Hesketh S, Pickles A, Skrondal A. Correcting for covariate measurement error in logistic regression using nonparametric maximum likelihood estimation. *Statistical Modelling* 2003; **3**: 215–232.

36. Carroll RJ, Ruppert D, Stefanski LA. *Measurement Error in Nonlinear Models*. London: Chapman & Hall, 1995.

37. Rabe-Hesketh S, Skrondal A, Pickles A. Maximum likelihood estimation of generalized linear models with covariate measurement error. *The Stata Journal* 2003; in press.

38. Yang M. Multinomial regression. In: Leyland AH, Goldstein H (eds.), *Multilevel Modelling of Health Statistics*. Chichester: Wiley, 2001; 107–123.

39. Hosmer DA, Lemeshow SA. *Applied Logistic Regression*. New York: Wiley, 2000.

40. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed.), *Frontiers in Econometrics*. New York: Academic Press, 1973; 105–142.

41. Skrondal A, Rabe-Hesketh S. Multilevel logistic regression for polytomous data and rankings. *Psychometrika* 2003; **68**: 267–287.

42. Skrondal A, Rabe-Hesketh S. Generalized linear mixed models for nominal data. *Proceedings of the American Statistical Association* 2003; in press.

43. Rabe-Hesketh S, Yang S, Pickles A. Multilevel models for censored and latent responses. *Statistical Methods in Medical Research* 2001; **10**: 409–427.

44. Maughan B, Pickles A, Rowe A, Costello R, Angold A. Developmental trajectories of aggressive and non-aggressive conduct problems. *Journal of Quantitative Criminology* 2000; **16**: 199–221.

45. Burton PR, Tiller KJ, Currin LC, Cookson WOCM, Musk AW, Palmer LJ. Genetic variance components analysis for binary phenotypes using generalized linear mixed models (glmms) and gibbs sampling. *Genetic Epidemiology* 1999; **17**: 118–140.

46. Campbell SM, Hann M, Hacker J, Burns C, Oliver D, Thapar A, Mead N, Safran DG, Roland MO. Identifying predictors of high quality care in English general practice: observational study. *British Medical Journal* 2001; **323**: 784–787.

47. McKee MD, Schlechter C, Burton W, Mulvihill M. Predictors of follow-up of atypical and ASCUS Papanicolaou test results in a high-risk population. *The Journal of Family Practice* 2001; **50**: 609–613.

48. Baker D, Hahn M. General practitioner services in primary care groups in England: is there inequity between service availability and population need? *Health & Place* 2001; **7**: 67–74.

49. Kaufman JS, Dole N, Savitz DA, Herring AH. Modeling community-level influences on preterm birth among African-American and white women in central North Carolina. *Annals of Epidemiology* 2003; **13**: 377–384.

50. Margolis DJ, Allen-Taylor L, Hoffstad O, Berlin JA. Diabetic neuropathic foot ulcers: The association of wound size, wound duration, and wound grade on healing. *Diabetes Care* 2002; **25**: 1835–1839.

51. Rabe-Hesketh S, Pickles A, Skrondal A. *Multilevel and Structural Equation Modeling of Continuous, Categorical and Event Data*. College Station, TX: Stata Press, 2004.