Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computational Statistics & Data Analysis 50 (2006) 3663–3680
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NLMIXED

James M. McMahon[a,*], Enrique R. Pouget[a,b], Stephanie Tortu[c]

[a]*National Development and Research Institutes, 71 West 23rd Street, 8th fl., New York, New York 10010, USA*
[b]*Yale University, School of Epidemiology and Public Health, New Haven, CT, USA*
[c]*Louisiana State University, School of Public Health, New Orleans, LA, USA*

Received 29 June 2005; received in revised form 22 August 2005; accepted 22 August 2005
Available online 15 September 2005

Abstract

In the social and health sciences, data are often structured hierarchically, with individuals nested within groups. Dyads constitute a special case of hierarchically structured data with variation at both the individual and dyadic level. Analyses of data from dyads pose several challenges due to the interdependence between members within dyads and issues related to small group sizes. Multilevel analytic techniques have been developed and applied to dyadic data in an attempt to resolve these issues. In this article, we describe a set of analyses for modeling individual- and dyad-level influences on *binary* outcomes using SAS statistical software; and we discuss the benefits and limitations of such an approach. For illustrative purposes, we apply these techniques to estimate individual- and dyad-level predictors of viral hepatitis C infection among heterosexual couples in East Harlem, New York City.

Keywords: Dyadic data; SAS NLMIXED; HLM; Multilevel modeling; Binary outcomes

1. Introduction

Many processes under study in the health sciences, such as treatment delivery, child-care, and disease transmission, involve interpersonal relationships and mutual influence involving two persons (e.g., physician–patient, parent–child, wife–husband). Conventional

* Corresponding author. Tel.: +1 212 845 4553; fax: +1 917 438 0894.
E-mail address: mcmahon@ndri.org (J.M. McMahon).

0167-9473/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2005.08.008

methods for inferential data analyses, including analysis of variance (ANOVA) and general linear regression, assume that observations obtained from each individual are independent. When such analyses are applied to data obtained from interacting dyads, the assumption of independent observations may be violated, leading to underestimation of standard errors and invalid inferences (i.e., increased Type I error).

To overcome the problem of nonindependence in the case of distinguishable dyad members, such as female–male couples, researchers often conduct separate analysis for each member class. For example, in a study of the effects of spousal support on health indicators, Heffner et al. (2004) performed separate analyses for females and males. Although this approach maintains independence of observations, it can obscure class interactions that may be of theoretical interest. In the example cited, although theory suggested that couple-level effects might moderate health status through interactions with spousal support and other variables, such effects could not be assessed in separate analyses. Several techniques have been developed to circumvent the problem of nonindependence while permitting estimation of dyad-level effects, including interactions (e.g., Kenny, 1996; Gonzalez and Griffin, 1999; Newsom, 2002).

Multilevel linear modeling (MLM) is one such technique developed and applied to dyadic data analysis (Barnett et al., 1993; Raudenbush et al., 1995; Windle and Dumenci, 1997; Kenny and Cook, 1999; Kashy and Kenny, 2000; Hoff, 2005). To make multilevel modeling techniques more accessible to data analysts, Campbell and Kashy (2002) have provided a practical guide for MLM analysis of dyadic data with continuous outcomes using two commercial software programs—SAS PROC MIXED and HLM. Here, we extend the work of these authors by providing a guide for *nonlinear* multilevel modeling of dyadic data with *binary* outcomes using NLMIXED and other procedures in SAS.

In Section 2, we briefly introduce multilevel modeling techniques and discuss limitations of this approach to analysis of dyadic data with binary outcomes. We also present model equations for both conditional and unconditional multilevel models and discuss methods for determining the appropriate use of the multilevel approach with dyadic binary response data. Section 3 lists the statistical assumptions of multilevel modeling with binary outcomes. In Section 4, we identify different types of dyadic data and discuss the implications of data properties and structure for multilevel modeling. An illustrative analysis involving risk factors for viral hepatitis C within heterosexual couples is introduced in Section 5. In Section 6, we provide practical guidelines for data preparation, using examples from the couples' risk study. In Section 7, we describe SAS PROC NLMIXED and provide a step-by-step guide for performing multilevel modeling analysis and interpretation using data from the couples risk study as exemplar. We conclude with a discussion, including limitations of the approach, in Section 8.

2. Multilevel modeling approaches to dyadic analysis with binary outcomes

Multilevel linear modeling refers to a family of regression estimation techniques applied to data organized into hierarchically structured clusters, such as students (level-1) nested within classrooms (level-2) (Raudenbush and Bryk, 2002). Dyadic data represent a special case of hierarchically clustered data, with individuals nested within dyads. Multilevel

analysis combines the effects of variables at different levels into a single model, while accounting for the interdependence among observations within higher-level units. In a two-level MLM, separate linear regressions are performed on observations within each level-2 cluster; and these first-order regression estimates (intercepts and slopes) are then used as outcomes in regression models involving level-2 units. These separate level-1 and level-2 regression models can be combined into a single multilevel model that may contain both fixed and random effects. Fixed effects are model components that assume no random variance or sampling error (e.g., group means, experimental conditions), and are constant across units of a given level; whereas random effects are model components that estimate population variance including sampling error (e.g., residuals, unobservable random quantities), and exhibit variation across units of a given level according to an error distribution. Multilevel models containing both fixed and random effects are referred to as mixed models.

Research into the statistical properties of mixed model estimates with only two observations per cluster has revealed a number of limitations (Newsom, 2002). One constraint on multilevel analysis of dyadic data is that it is often not feasible to estimate random effects for both the intercepts and slopes simultaneously, due to model identification problems. Specifically, a dyadic multilevel analysis incorporating both random intercepts and slopes will result in an overdetermined model—one with too many parameters to be estimated given the number of covariance elements available (Newsom, 2002). More generally, statistical theory as well as simulation studies indicate that estimates and confidence limits for the level-2 random variance components can be biased if the number of observations per cluster is small (Hox, 1998; Hox and Maas, 2002; Raudenbush and Bryk, 2002). Our own experience with multilevel analyses of dyadic data indicate that inclusion of random slope variance components, in particular, can lead to severe convergence problems. Thus, in practical terms, such models may be limited to inclusion of random intercepts for the random part.

Another limitation of multilevel modeling of dyadic data concerns estimation bias of level-2 coefficients stemming from the within-cluster variance–covariance structure. Carlin et al. (2001) have shown that level-2 coefficient estimates are determined largely by clusters whose members are not all homogeneous on the response variable. Accordingly, dyads in which both members have identical response values do not contribute information to the likelihood function.

Further restrictions apply when the outcome of interest is binary. One common approach to modeling binary responses (0, 1) is the logit link function or log-odds transformation,

$$\eta_{ij} = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right), \tag{1}$$

where p_{ij} is the probability of observing the response, $y_{ij} = 1$, in the ith individual of the jth cluster; and η_{ij} is the log odds of observing the response, which can take any real value.

Analysis of dyadic data using a 2-level random coefficients model with a single predictor at each level and a binary outcome can be represented by three separate equations: the level-1 logistic model with one predictor takes the general form

$$\eta_{ij} = \beta_{0j} + \beta_{1j} x_{ij}, \tag{2}$$

where β_{0j} is the within-dyad intercept in cluster j and β_{1j} is the slope of η_{ij} on x_{ij} in cluster j. The intercept β_{0j} represents the value of η_{ij} when the predictor equals zero. Notice that the individual-level residual term, r_{ij}, is omitted from the model because its variance, denoted σ^2, follows directly from the success probability (i.e., $\sigma^2 = p_j[1 - p_j]$), and is therefore assumed fixed under the log-odds transformation (Snijders and Bosker, 1999; Raudenbush and Bryk, 2002). The intercept term, β_{0j}, in Eq. (2) can be expanded into level-2 fixed and random components:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}, \tag{3}$$

where γ_{00} is the average intercept across dyads, γ_{01} is the slope of the dyad intercepts regressed on the level-2 predictor variable z_j, and u_{0j} is the unique effect of dyad j on the intercept (conditional on z_j). In the current example, the level-2 slope coefficient is viewed as fixed across clusters:

$$\beta_{1j} = \gamma_{10}. \tag{4}$$

Substituting the level-2 equations for terms in the level-1 model yields the combined equation

$$\eta_{ij} = \gamma_{00} + \gamma_{01} z_j + \gamma_{10} x_{ij} + u_{0j}. \tag{5}$$

In this formulation, u_{0j} is the sole random effect in the model and the other terms are the fixed effects.

Before evaluating these conditional models, however, the first step in the analysis is to determine whether there is significant within-cluster interdependence to justify the use of a multilevel approach. In conventional multilevel linear analysis this is typically achieved by evaluating the unconditional model (i.e., one with no predictors) to determine if clustering can account for a significant proportion of the overall variance of the outcome variable y_{ij}. The unconditional linear model takes the form

$$y_{ij} = \beta_{00} + u_{0j} + r_{ij}. \tag{6}$$

The variance τ_{00} of the intercept random effect, u_{0j}, and the variance σ^2 of the level-1 residual term, r_{ij}, are estimated and used to evaluate ρ, the intraclass correlation coefficient (ICC). The ICC is a measure of the extent to which observations within a cluster are related as expressed by the ratio of the between-cluster variance to the total variance

$$\rho = \frac{\tau_{00}}{\sigma^2 + \tau_{00}}. \tag{7}$$

In multilevel linear modeling with continuous outcomes, the ICC is a convenient measure of the proportion of the total variance attributed to clustering. However, this standard ICC formulation is not valid in the case of binomial hierarchical models due to the previously discussed properties of the level-1 residual term.

Ridout et al. (1999) evaluated 20 different methods for estimating the ICC with binary outcomes and found that the kappa-type method originally proposed by Fleiss and Cuzick (1979) performed best under a variety of conditions. For dyadic data with binary outcomes

the simple Pearson correlation coefficient (equivalent to the Phi coefficient) is a close approximation to the Fleiss–Cuzick method, which has been shown to provide reliable estimates for ρ using data with a reasonably large number of clusters ($j > 50$), each of relatively small size (Zou and Donner, 2004). This method has been referred to as the *pairwise intraclass correlation coefficient* (PICC) by Donner and Koval (1980). A favorable property of the Pearson-type PICC as applied to dyadic data is that its lower limit is always -1, according to the general formula for determining the PICC lower boundary: $-1(n-1)$, where n is the number of subjects per cluster (Kenny et al., 1998). It should be noted that whereas the conventional ICC with linear unconditional models will give an estimate of the proportion of the total variance in the outcome attributed to clustering, the Pearson-type PICC will provide only a measure of dyadic interdependence.

Due to the propensity for Type II errors, Kenny and colleagues (Kenny and Kashy, 1991; Kenny et al., 1998) recommend the use of a liberal alpha level (.20–.25) when interpreting the Pearson-type PICC to determine whether within-dyad interdependence is evident. A liberal alpha level is also preferred because it is possible that group effects only emerge after controlling for covariates (Snijders and Bosker, 1999). Alternatively, a modified variance formula for the PICC in the case of constant cluster size was derived by Zou and Donner (2004); and these authors also suggest the use of a modified Wald method to construct confidence intervals appropriate for cases in which ρ is close to 0 or 1 or the number of dyads is small (Donner and Zou, 2002).

3. Statistical assumptions

There is no single definitive set of assumptions that apply to all multilevel logistic models. The primary assumptions that are relevant to multilevel models involving binary outcomes using the logit link function (as shown in Eq. (10)) are that (a) the probability of success ($y_{ij} = 1$) is identical for individuals within clusters, (b) observations between clusters are independent, whereas pairs of observations within clusters have a common correlation, (c) each random effect is independent and follows a generalized distribution that can be estimated using maximum likelihood, (d) random effects and model predictors at all levels are independent, and (e) an appropriate model linking y_i and u_i exists with a joint probability density function. A more comprehensive discussion of the statistical assumptions under nonlinear multilevel modeling can be found in Raudenbush and Bryk (2002).

4. Types of dyadic variables and data structure

In multilevel models, dyadic interdependence is accounted for by modeling variance and covariance within and across dyads. Variance constraints on particular types of dyadic variables are thus an important issue in multilevel analysis of dyadic data. Kashy and Kenny (2000) have identified three types of dyadic variables based on their locus of variance: *between*-dyads, *within*-dyads, and *mixed*. Between-dyad variables measure shared experiences or behavior and do not vary within the dyad, but do vary across dyads. Examples include relationship duration, number of doctor–patient visits, and frequency of intercourse

between sexual partners. Since variation occurs across dyads only, between-dyad variables are restricted to level-2 analysis. Within-dyad variables are those that vary within each dyad, but sum to the same value across dyads. Examples include gender in heterosexual couples and proportion of childcare performed by adult caregivers. Within-dyad variables are restricted to level-1, since there is no variation across dyads. Mixed variables can vary both within and between dyads. Examples include age, monthly caloric intake, and psychiatric assessment scores. Mixed dyadic variables can be included as level-1 or level-2 variables in multilevel models. Recoding is required prior to use of mixed variables as level-2 predictors; for example, by calculating dyad means.

Kenny and colleagues (Kenny, 1996; Kenny and Cook, 1999; Kashy and Kenny, 2000) have also provided a conceptual and analytical framework for modeling between-, within- and mixed-dyad variables while accounting for dyadic interdependence. The actor–partner interdependence model (APIM) can be used to analyze data from dyads or small groups in a variety of research designs (Kenny et al., 2002). In the APIM framework, each dyad member is considered an *actor* as well as a *partner* in the dyad. Each individual outcome can be influenced by actor effects (level-1), partner effects (level-1), interactions between actor and partner effects (level-2), and by dyad-level effects (level-2). This framework does not require dyads to be distinguishable, since both actor and partner effects and interactions are assessed for both members.

5. Example: viral hepatitis C infection among heterosexual couples

To illustrate the application of multilevel modeling to dyadic data with binary outcomes, we employed data from a recent epidemiological study conducted to examine risk for hepatitis C and other viral infections among drug-using couples in East Harlem, New York City (Tortu et al., 2003; McMahon et al., 2003). Heterosexual couples reporting recent substance use were recruited from East Harlem and administered risk assessment surveys and screened for viral hepatitis C antibodies. Protocols for this study have been described in detail elsewhere (McMahon et al., 2003).

Hepatitis C virus (HCV) is the most common blood-borne pathogen in the United States with nearly 2% of the general population infected. HCV is a major cause of chronic liver disease and the leading indication for liver transplantation worldwide. Injection drug use (IDU) and high-risk sexual activity are the most common risk factors associated with HCV infection (Ramadori and Meier, 2001; CDC, 2004). Currently, there is no vaccine against HCV and therapeutic treatments for chronic active infection are limited.

To illustrate the use of PROC NLMIXED to model actor, partner, and dyadic effects on a binary outcome, we examined the effects of four predictors on actor HCV antibody reactivity, including (1) actor gender, (2) actor injection drug use, (3) partner age, and (4) recent dyadic sexual behavior. In the APIM framework, actor gender (binary) is a *within*-dyad variable, because variation occurs within- but not between-dyads; actor IDU status (binary) and partner age (continuous) are *mixed* variables, in that they can vary both within and between couples; and recent shared sexual behavior (continuous) is a *between*-dyad variable because variation occurs between couples, not within each couple (members within each couple should report the same number of acts of intercourse between them).

Table 1
Description of variables used in the analysis ($N = 265$ couples)

	Females	Males	Dyads
Dependent variable			
Actor anti-HCV status (aHCV; % reactive) (binary, level-1, mixed)	50.6%	54.7%	
Independent variables			
Actor self-reported IDU status (aIDU; % ever injected) (binary, level-1, mixed)	61.5%	65.3%	
Partner self-reported age (pAGE; mean years) (binary, level-1, mixed)	40.3	39.1	
Number of within-couple unprotected acts of intercourse in previous 30 days (dSEX; mean acts) (count, level-2, between)			14.7

It must be emphasized that this limited model was not intended to be properly specified. Indeed, the model is likely misspecified by the omission of several known risk factors for hepatitis C infection, and our results should be viewed as illustrative rather than substantive.

Of the 353 couples recruited for the initial study, 265 provided conclusive HCV antibody test results and were included in the present analyses. In our example, the prefix *a* denotes actor variables, the prefix *p* denotes partner variables, and the prefix *d* denotes dyad variables. Table 1 describes the variables used in this illustrative analysis. Before proceeding with the analysis, we first discuss several data preparation issues, including variable coding and centering.

6. Variable coding and centering

Between-dyad variables (such as frequency of intercourse reported by each couple) require that a single value be entered in the dataset for each dyad. However, in cases in which data are collected from both members of each dyad, responses can disagree (i.e., measurement error). One way to resolve this problem is to employ the response of one dyad member only (e.g., the female member of each couple); a second method is to use within-dyad means. The latter method was used in the current analysis, such that dyad means were used for the variable *dSEX*.

Binary independent variables are dummy coded (0, 1) to facilitate interpretation of results. Dummy coding is preferred to other coding schemes (such as effect coding) because the 0 value on which the intercept is interpreted corresponds to a real state (Snijders and Bosker, 1999). Thus, actor gender (aGEN) was coded 1 for females and 0 for males; and actor IDU status (aIDU) was coded 1 for respondents who reported a history of injection drug use and 0 for those who reported no such history.

To further facilitate interpretation of the intercept, level-1 continuous mixed variables were centered on the grand mean. Kreft et al. (1995) provide a good discussion of the consequences of various centering schemes and justification for their use. In our example, we centered partner age by subtracting the grand mean from each variate ($x_j - \bar{x}$). Hence, the intercept, β_{0j}, represents the expected outcome (log odds) for an individual in dyad j whose age is equal to the grand mean. Without centering, β_{0j} would denote the outcome

Table 2
APIM-structured dataset with variables from example analysis

Obs	id	aGEN	aHCV	pHCV	aIDU	pIDU	aAGE	pAGE	dSEX
1	1	0	0	1	1	1	44	54	9
2	1	1	1	0	1	1	54	44	9
3	2	0	1	1	1	0	36	40	4
4	2	1	1	1	0	1	40	36	4
5	3	0	0	0	0	0	38	37	15
6	3	1	0	0	0	0	37	38	15

measure at age 0, a value without meaningful interpretation. The suffix *c* denotes variables that were centered on the grand mean (e.g., aAGEc, dSEXc). Unlike several other software packages, SAS procedures have no automated variable centering capability, so variables must be centered prior to execution of PROC NLMIXED.

7. SAS PROC NLMIXED

Previous versions of SAS software have provided a variety of procedures for constructing multilevel mixed models. The MIXED procedure was developed to handle linear multilevel random effects models with continuous outcomes. Two subsequent SAS macros—GLIMMIX and NLINMIX—were written to extend the capabilities of PROC MIXED to include nonlinear mixed models (Littell et al., 1996; Wolfinger, 1997). Although these macros provide several different estimation options, most are iteratively fit to a set of generalized estimating equations (GEE), which have been shown to work poorly for data with considerable heterogeneity and a small number of responses per cluster (Breslow and Clayton, 1993; Wolfinger and O'Connell, 1993; Ten Have and Localio, 1999). Moreover, GLIMMIX and NLINMIX provide only approximate parameter estimations. In contrast, the NLMIXED procedure, which was introduced in version 7 of SAS, delivers maximized, and theoretically exact, integrated likelihood estimates based on an adaptive Gaussian quadrature (Pinheiro and Bates, 1995; Agresti et al., 2000). This estimation method provides superior parameter estimates especially when variance components are large or not normally distributed (Kuss, 2002). In addition, with PROC NLMIXED the conditional distribution of the data given the random effects can take any general form that can be programmed using the SAS language including normal, binary, Poisson, and binomial. Kuss (2002) has provided a concise description and comparison of the various SAS procedures that can be used to analyze mixed models with binary outcomes; and Patefield (2002) has described several nonlinear model applications of PROC NLMIXED using data with large samples.

7.1. Data structure for NLMIXED

In order to be read properly by the NLMIXED procedure within the framework of APIM, the data need to be structured as shown in Table 2. This table contains a number of different actor, partner, and dyadic variables to illustrate the APIM structure more fully. The dataset must contain one record for each level-1 individual in the sample. Depending on the variable

type—within, between, or mixed—the dataset will consist of columns containing *actor*, *partner* (individual, level-1) and *dyadic* (couple, level-2) data. Thus, each subject's *actor* data will be the same as his or her partner's *partner* data. This structure has been referred to as *pairwise reverse column scores* (e.g., actor and partner age variables in Table 2). Couple-level data are identical for individuals within each dyad. NLMIXED assumes that data are ordered both within and between clusters as shown in Table 2.

Below we provide example code in SAS for restructuring a dataset containing subject actor variables into the APIM format. The first six records from the couples HCV data were read into a SAS dataset named *new*. Note that a dyad identifier variable (id) and a within-dyad class variable (in this case, aGEN where 0 = female and 1 = male, in ascending order) are required in the raw dataset.

```
data new;
   input id aGEN aHCV aIDU aAGE aSEX;
   datalines;
   1   0   0   1   44   10
   1   1   1   1   54   7
   2   0   1   1   36   4
   2   1   1   0   40   3
   3   0   0   0   38   18
   3   1   0   0   37   12
run;
```

The next program restructures the raw data into APIM-format by merging the initial dataset with a second dataset in which gender is sorted in descending order; and the between-dyad variable dSEX is constructed using dyad means.

```
proc sort data = new out = APIM;
   by id descending aGEN;
data APIM;
   merge APIM (rename = (aHCV  = pHCV aIDU  = pIDU aAGE  = pAGE aSEX
 = pSEX)) new;
   dSEX  = (aSEX + pSEX)/2;
```

Finally, the continuous variables aAGE, pAGE, and dSEX were grand mean centered and read into a SAS dataset named *couplesHCV*.

```
proc means data = APIM;
   var aAGE dSEX;
   output out  = meandata mean  = mAGE mSEX;
data couplesHCV;
   if_N_  = 1 then set meandata;
   set APIM;
   aAGEc = round(aAGE − mAGE);
   pAGEc = round(pAGE − mAGE);
   dSEXc = round(dSEX − mSEX,  .1);
run;
```

7.2. Test of within-dyad interdependence using complete dataset

As discussed in Section 2, the first step in the analysis is to determine whether there is significant within-cluster interdependence to warrant the use of a multilevel approach. The Pearson correlation coefficient for paired binary responses provides a measure of within-dyad interdependence, based on the PICC method discussed previously. The Pearson coefficient can be obtained in SAS using either the FREQ or CORR procedures (the former gives confidence limits whereas the latter reports p-values). The correlation between female and male dyad pairs for anti-HCV positivity was obtained using the following SAS code:

```
proc freq data = couplesHCV;
   where aGEN = 0;
   table aHCV*pHCV / measures cl;
run;
```

With APIM-structured data, inclusion of the WHERE clause ensures that each dyad record will be read only once, thus attaining the appropriate degrees of freedom for inferential tests (note that substituting "where aGEN = 1" on this line will produce identical results). The TABLE statement identifies the variables on which the analysis is performed; and the MEASURES and CL options stipulate that the output will list measures of association, include the Pearson correlation coefficient, and 95% confidence limits. Using the complete dataset ($N = 265$ couples), this analysis produced a Pearson correlation coefficient of 0.298 with an asymptotic standard error of 0.059 and confidence limits of 0.184 and 0.413. Based on this result, we may reject the null hypothesis of no interdependence between members within dyads on anti-HCV positivity. We conclude that the use of a multilevel modeling approach to the data is warranted in the present example.

It should be noted that both the FREQ and CORR procedures in SAS employ conventional methods for calculating the variance, confidence limits, and p-values for the Pearson correlation coefficient, not the modified formulae recommended by (Zou and Donner, 2004; Donner and Zou, 2002). In our example, the sample prevalence for anti-HCV positivity was not close to 0 or 1 and the number of dyads was reasonably large; therefore, the Zou–Donner modified formulae gave nearly identical variance and confidence limits to those generated by PROC FREQ. A SAS Interactive Matrix Language (IML) program for calculating the Zou–Donner parameter estimates is provided in Appendix A.

7.3. Evaluation of conditional random intercepts model

Next, the (conditional) random intercepts model was evaluated using the following PROC NLMIXED code:

```
proc nlmixed data = couplesHCV qpoints = 20 tech = newrap;
   parms beta0 = -1.7940 beta1 = -0.1762 beta2 = 3.0536 beta3 = 0.0562 beta4 =
0.0162 s2u = 0.0418;
   eta = beta0 + beta1*aGEN + beta2*aIDU + beta3*pAGEc + beta4*dSEXc + u;
   mu = exp(eta) / (1 + exp (eta));
   model aHCV ~ binary (mu);
   random u ~ normal(0, s2u) subject = id;
run;
```

The preceding SAS program evaluates the multilevel model:
Level-1 model:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}\text{aGEN}_{ij} + \beta_{2j}\text{aIDU}_{ij} + \beta_{3j}\text{pAGEc}_{ij}. \tag{8}$$

Level-2 models:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}\text{dSEXc}_j + u_{0j},\\ \beta_{1j} &= \gamma_{11},\\ \beta_{2j} &= \gamma_{12},\\ \beta_{3j} &= \gamma_{13}.\end{aligned} \tag{9}$$

Combined model:

$$\eta_{ij} = \gamma_{00} + \gamma_{11}\text{aGEN}_{ij} + \gamma_{12}\text{aIDU}_{ij} + \gamma_{13}\text{pAGEc}_{ij} + \gamma_{01}\text{dSEXc}_j + u_{0j}. \tag{10}$$

The PROC NLMIXED statement invokes the SAS procedure, and the DATA = option inputs the APIM-structured dataset *couplesHCV*. The QPOINTS = option stipulates the number of quadrature points to be used to obtain integral approximations over the random effects. Carlin et al. (2001) have shown in simulation studies that a reasonably large number of quadrature points (i.e., 20) is required to ensure convergence on model parameter estimates. Indeed, in our example, running the SAS program without the QPOINTS = 20 option led PROC NLMIXED to adaptively select one (1) quadrature point, which resulted in a Type II error involving the estimate for the level-2 random variance component, $s2u$. More complex models may require an even greater number of quadrature points. However, if we require 20 quadrature points for each random effect, the total number of quadrature points required will be 20^k, where k is the number of random effects to be estimated. Such a large number of quadrature points can lead to excessive demands on memory and computation time.

The TECH = option specifies the optimization technique used in parameter estimation. PROC NLMIXED provides seven alternative optimization techniques, which vary in degree of reliability and performance. The most reliable optimization techniques compute the Hessian matrix, but these techniques are also the most computationally demanding. The default technique is the dual quasi-Newton (QUANEW) algorithm, which affords a good compromise between reliability (it computes an approximation of the Hessian) and performance (it converges relatively quickly for models of small to medium complexity). If computation time is not an issue, the Trust Region method (TRUREG) or the Newton–Raphson algorithm (NEWRAP) tends to be more reliable (SAS, 1999).

The PARMS statement defines the free parameters and their starting values, which are crucial to achieve convergence. Good starting values can be obtained by running other SAS procedures. For example, starting values for the intercept and slope parameters (*beta*0 through *beta*3) can be obtained using PROC GENMOD with a marginal modeling approach such as the generalized estimating equations (GEE) formulation; whereas initial values for $s2u$ (between-cluster variance, or variance due to dyads) can be obtained from PROC MIXED (see Appendix B).

The ETA statement specifies the combined multilevel model in linear form. Variable u represents the level-2 random effect. The MU = statement specifies the logistic mixed model. The MODEL statement defines the dependent variable and its distribution and indicates the model. The RANDOM statement defines the single random effect, u, and specifies that it follows a normal distribution, with a mean of 0 and variance $s2u$. The SUBJECT = argument indicates the distinction between level-1 and level-2 hierarchies; in this case, the couple *id* number.

7.4. NLMIXED results for random intercepts model

The preceding NLMIXED code was run in SAS Release 8.02 TS Level 02M0; and the output is presented in Table 3.

The Specifications table lists the various options and specifications employed in the analysis, including the dataset employed, the dependent variable and its distribution, the random effects and their distributions, the subject variable used to identify level-2 membership, and the optimization and estimation methods. The Dimensions output provides information about the sample size and its hierarchical structure, and the number of parameters estimated in the model. The starting values for the model parameters and the negative log likelihood given these values are listed in the Parameters output. Convergence was achieved according to the GCONV criterion in six iterations, as documented in the Iteration History table. GCONV is the default termination criterion for model optimization based on relative gradient convergence. The Fit Statistics table lists the final maximized value of the −2 Log Likelihood, and three other model fit statistics: the Akaike Information Criteria (AIC and AICC) and Bayes Information Criterion (BIC). The AIC, AICC and BIC apply various corrections for model complexity to the log likelihood model fit (Cherkassky and Ma, 2003).

The Parameter Estimates table lists the six free parameters, their maximum likelihood estimates, standard errors, and inferential statistics. Degrees of freedom are equal to $j - 1$. The estimate for the model intercept (*beta*0) is −2.35, and represents the log-odds of anti-HCV reactivity for an individual with 0-values on the predictor variables. The log-odds can be converted back to a probability using the inverse exponential transformation:

$$p_{ij} = \left(\frac{e^{\eta_{ij}}}{1 - e^{\eta_{ij}}} \right), \tag{11}$$

where $e^{\eta_{ij}}$ is the exponent of the log-odds. Thus, the probability of anti-HCV positivity is .087 for an actor who is not an injection drug user (aIDU = 0), with a partner of average age (pAGE = 39.7 years), and who has engaged in an average number of acts of unprotected intercourse in the last 30 days (dSEX = 14.7).

Estimated coefficients for the level-1 predictors aGEN, aIDU and pAGEc are beta1 = −0.219, beta2 = 4.042 and beta3 = 0.066, respectively. Examination of the confidence limits and p-values indicate that both beta2 and beta3 are significant at the .05 level, whereas beta1 is not. This suggests that gender had no significant effect on the probability of testing anti-HCV positive, taking into account the other predictors and the level-2 random variance. To make the interpretation of beta2 and beta3 statistics more intuitive, the coefficients and 95% confidence limits can be converted to adjusted odds ratios (AOR) by taking their

Table 3
SAS output listing from PROC NLMIXED code

The NLMIXED procedure

Specifications

Data Set	COUPLESHCV
Dependent Variable	aHCV
Distribution for Dependent Variable	Binary
Random Effects	u
Distribution for Random Effects	Normal
Subject Variable	id
Optimization Technique	Newton-Raphson
Integration Method	Adaptive Gaussian Quadrature

Dimensions

Observations Used	530
Observations Not Used	0
Total Observations	530
Subjects	265
Max Obs Per Subject	2
Parameters	6
Quadrature Points	20

Parameters

beta0	beta1	beta2	beta3	beta4	s2u	NegLogLike
−1.794	−0.1762	3.0536	0.0562	0.0162	0.04178	264.187725

Iteration history

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	16	260.334592	3.853133	11.49504	−6.05858
2	24	258.901004	1.433588	2.728152	−2.31961
3	32	258.565145	0.335859	0.546435	−0.57772
4	40	258.536421	0.028725	0.072913	−0.05392
5	48	258.536111	0.00031	0.000997	−0.00061
6	56	258.536111	4.428E-8	5.86E-8	−8.86E-8

NOTE: GCONV convergence criterion satisfied.

Fit statistics

-2 Log Likelihood	517.1
AIC (smaller is better)	529.1
AICC (smaller is better)	529.2
BIC (smaller is better)	550.6

Parameter estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr>|t|	Alpha	Lower	Upper	Gradient
beta0	−2.3458	0.3747	264	−6.26	<.0001	0.05	−3.0837	−1.6080	2.342E-8
beta1	−0.2187	0.2515	264	−0.87	0.3854	0.05	−0.7139	0.2765	9.256E-9
beta2	4.0416	0.5053	264	8.00	<.0001	0.05	3.0466	5.0365	5.153E-8
beta3	0.06592	0.02039	264	3.23	0.0014	0.05	0.02577	0.1061	−5.69E-8
beta4	0.02031	0.008921	264	2.28	0.0236	0.05	0.002743	0.03787	−3.92E-8
s2u	2.0291	0.9356	264	2.17	0.0310	0.05	0.1869	3.8713	−5.86E-8

exponents—exp(beta2) = 56.94, CI: 21.05, 154.01; exp(beta3) = 1.07, CI: 1.03, 1.11. This means that (holding the other terms constant) the odds of being anti-HCV positive for an injection drug user are nearly 60 times that of a non-IDU. The significant partner age effect indicates that for each 1-year increase in an actor's partner's age, the odds of that actor being anti-HCV positive increase by about .07.

The coefficient for the level-2 predictor dSEXc (beta4) can be interpreted in much the same manner. The estimate of beta4 = 0.02 (AOR = 1.02), $p = .024$, indicates that for each additional act of unprotected intercourse between actor and partner per 30 days, the odds of being anti-HCV positive increase by about .02. However, as discussion in Section 2, interpretation of level-2 coefficients in random effects models with binary responses must be viewed with caution (Neuhaus and Kalbfleisch, 1998).

The estimate for the level-2 random effect (s2u) in this model is 2.029 (p-value = 0.031). Attainment of statistical significance for this parameter is typically interpreted as indicating the existence of unexplained variance in the outcome variable associated with one or more level-2 unobservable random quantities. However, these test statistics must be interpreted with caution because sampling distributions can be skewed (Wolfinger, 1999).

8. Discussion and conclusions

The SAS procedures outlined in this paper provide a practical guide for evaluating multilevel mixed models with binary outcomes using data from distinguishable dyads. One of the strengths of the SAS NLMIXED procedure is the flexibility it allows for specifying a variety of models, which may include any combination of actor, partner, and dyad-level effects, within-level and cross-level interaction terms, and random components. Interaction terms may be added to NLMIXED models directly using the SAS multiplication operator (*). For example, an interaction term for actor IDU status and partner age can easily be added to the model specified in Section 7.3 simply by adding the term beta5*aIDU*pAGEc into the ETA = statement (with an appropriate seed value for beta5 specified in the PARMS statement). Additionally, the wide array of options available for parameter estimation and hypothesis testing (e.g., density function, degrees of freedom, optimization, integration, and convergence) permit model evaluation to be tailored to specific data properties and research requirements; although this same flexibility can be somewhat daunting for those unfamiliar with the various combinations of options.

Along with these strengths, NLMIXED also has several limitations. Currently, only one random statement is supported in PROC NLMIXED, so that nonlinear mixed models cannot be assessed at more than two levels. Parameter estimation in NLMIXED is limited to maximum likelihood (ML) solutions. The restricted maximum likelihood (REML) method, which corrects for the downward bias of ML estimates, is not available in NLMIXED. The REML method is especially useful for adjusting standard errors of the level-2 random effects (Maas and Hox, 2004); and there is no alternative option in NLMIXED for robust correction of standard errors, such as the Huber–White adjustment.

Other limitations derive from the use of the APIM. The practice of using dyad means for level-2 variables from between-dyad scores ignores within-dyad measurement error. A related problem is how to handle between-dyad data in which dyad members disagree on

a dichotomous variable. It is not clear that using a mean of .5 in such cases will provide an optimum solution. One potential solution to these problems is to apply simulation techniques to model within-dyad uncertainty in the results (e.g., Phillips, 2003; Steenland and Greenland, 2004).

A similar approach might also be helpful in dealing with another problem related to analysis of APIM-structured data. Due to the structure of APIM data, it is not possible to include both an actor outcome variable and its corresponding partner variable as a predictor in the same model. In our hepatitis C analysis, for example, we could not include partner HCV reactivity (pHCV) as a main effect in the model because it contains the same data as the outcome variable—actor HCV reactivity (aHCV). We could not, therefore, test various interactions of theoretical interest, such as the relationship between couple's frequency of unprotected sex and actor anti-HCV positivity moderated by partner HCV status. However, if uncertainty estimates based on sensitivity and specificity parameters of the HCV testing procedures are incorporated into the model (e.g., Neuhaus, 2002; McInturff et al., 2004), both actor dependent and paired partner independent variables can be evaluated in the same model.

Research focusing on the effects of dyadic interactions and processes on health outcomes has increased in recent years. Analyses of dyadic data pose special challenges due to small sample sizes and interdependence of observations within dyads. Several authors have outlined procedures for conducting dyadic data analysis with continuous responses using commercially available software packages. In this paper, we outline a step-by-step guide to estimate multilevel effects on *binary* outcomes in data from dyads using PROC NLMIXED and other SAS software procedures. Use of the analytic tools described in this guide were illustrated within the APIM framework using an epidemiological example concerning risk factors for viral hepatitis C infection among heterosexual couples. Our aim was to provide a worked example of these procedures as context for a discussion on some of the main concepts, issues, options, and interpretations involving nonlinear multilevel modeling of dyadic data. Along with the set of procedures for analysis of continuous outcomes provided by Campbell and Kashy (2002), the methods outlined here can be extended to include analysis of data from a wide variety of generalized distributions. It is hoped that this guide will serve as a starting point for researchers who wish to explore theories about dyadic effects on binary outcomes using multilevel nonlinear models.

Acknowledgements

Support for this work was provided by grants from the National Institutes of Health (NIH), National Institute on Drug Abuse (NIDA) to Dr. James McMahon (R01 DA15641) and Dr. Stephanie Tortu (R01 DA12805). The authors thank Dr. Janet Rice, Department of Biostatistics, Tulane University, for her helpful remarks on the manuscript. Dr. Peter Flom and the NDRI Statistics Support Group provided insightful comments on an earlier version of the paper. Jeanine Botta provided technical assistance with the manuscript.

Appendix A. SAS PROC IML code for calculating the Zou–Donner modified variance estimate for the Pearson-type PICC.

```
proc iml;
p = 0.298;          *Pearson correlation coefficient from PROC FREQ;
j = 265;            *Number of dyads (clusters in analysis);
n = 2;              *Number of subjects per dyad (cluster size);
t = 0.526;          *Proportion of outcome successes for entire sample;

*Zou and Donner (2004 modified formula for variance of Pearson-type ICC);

varp = ((1 - p)/j)*((2/(n*(n - 1))) - ((3 - (1/(t*(1 - t))))*p)
      +((n - 1)/n)*((4 - (1/(t*(1 - t))))*(p**2)));

print varp;

run;
```

Appendix B.

SAS PROC MIXED and PROC GENMOD code and selected output for parameter starting values to be used in PROC NLMIXED. The random intercepts' variance estimate is listed in the Covariance Parameter Estimates table under the Mixed Procedure output line UN(1,1). Coefficient estimates are listed in the Analysis of GEE Parameter Estimates table under the GENMOD Procedure output.

```
proc mixed data = couplesHCV method = REML;
  class id;
  model aHCV = aGEN aIDU pAGEc dSEXc /solution;
  random intercept /subject = id type = un;
run;
```

The Mixed Procedure

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	id	0.04178
Residual		0.1227

```
proc genmod data = couplesHCV descending;
  class id;
  model aHCV = aGEN aIDU pAGEc dSEXc / dist = bin link = logit;
  repeated subject = id / type = un;
run;
```

The GENMOD Procedure

Analysis Of GEE Parameter Estimates

Empirical Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr>|Z|
Intercept	−1.7940	0.2421	−2.2685	−1.3195	−7.41	<.0001
aGEN	−0.1762	0.1898	−0.5483	0.1959	−0.93	0.3534
aIDU	3.0536	0.2824	2.5001	3.6071	10.81	<.0001
pAGEc	0.0562	0.0184	0.0200	0.0923	3.05	0.0023
dSEXc	0.0162	0.0069	0.0027	0.0297	2.35	0.0190

References

Agresti, A., Booth, J.G., Hobert, J.P., Caffo, B., 2000. Random effects modeling of categorical response data. Sociological Methodology 30, 27–80.

Barnett, R.C., Marshall, N., Raudenbush, S.W., Brennan, R., 1993. Gender and the relationship between job experiences and psychological distress: a study of dual-earner couples. J. Personality Social Psychology 64, 794–806.

Breslow, N., Clayton, D., 1993. Approximate inference in generalized linear mixed models. J. Amer. Statist. Assoc. 88, 9–25.

Campbell, L., Kashy, D.A., 2002. Estimating actor partner and interaction effects for dyadic data using PROC MIXED and HLM: a user-friendly guide. Personal Relationships 9, 327–342.

Carlin, J.B., Wolfe, R., Brown, C.H., Gelman, A., 2001. A case study on the choice, interpretation and checking of multilevel models for longitudinal binary outcomes. Biostatistics 2, 397–416.

Centers for Disease Control and Prevention, 2004. Hepatitis Surveillance Report No. 59. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Cherkassky, V., Ma, Y., 2003. Comparison of model selection for regression. Neural Computation 15, 1691–1714.

Donner, A., Koval, J.J., 1980. The estimation of intraclass correlation in the analysis of family data. Biometrics 36, 19–25.

Donner, A., Zou, G., 2002. Interval estimation for a difference between intraclass kappa statistics. Biometrics 58, 209–214.

Fleiss, J.L., Cuzick, J., 1979. The reliability of dichotomous judgments: unequal numbers of judges per subject. Appl. Psychological Measurement 3, 537–542.

Gonzalez, R., Griffin, D., 1999. The correlational analysis of dyad-level data in the distinguishable case. Personal Relationships 6, 449–469.

Heffner, K.L., Kiecolt-Glaser, J.K., Loving, T.J., Glaser, R., Malarkey, W.B., 2004. Spousal support satisfaction as a modifier of physiological responses to marital conflict in younger and older couples. J. Behavioral Medicine 27, 233–254.

Hoff, P.D., 2005. Bilinear mixed-effects models for dyadic data. J. Amer. Statist. Assoc. 100, 286–295.

Hox, J.J., 1998. Multilevel Modeling: When and Why? Classification, Data Analysis, and Data Highways. Springer, Berlin.

Hox, J.J., Maas, C.J.M., 2002. Sample sizes for multilevel modeling. Social science methodology in the new millennium. Proceedings of the Fifth International Conference on Logic and Methodology, Leske & Budrich, Opladen.

Kashy, D.A., Kenny, D.A., 2000. The analysis of data from dyads and groups. Handbook of Research Methods in Social and Personality Psychology. Cambridge University Press, New York.

Kenny, D.A., 1996. Models of non-independence in dyadic research. J. Social Personal Relationships 13, 279–294.

Kenny, D.A., Cook, W., 1999. Partner effects in relationship research: conceptual issues, analytic difficulties, and illustrations. Personal Relationships 6, 433–448.

Kenny, D.A., Kashy, D.A., 1991. Analyzing interdependence in dyads. Studying Interpersonal Interaction. Guilford, New York.

Kenny, D.A., Kashy, D.A., Bolger, N., 1998. Data analysis in social psychology. The Handbook of Social Psychology. McGraw-Hill, New York.

Kenny, D.A., Mannetti, L., Pierro, A., Livi, S., Kashy, D.A., 2002. The statistical analysis of data from small groups. J. Personality Social Psychology 83, 126–137.

Kreft, I.G.G., de Leeuw, J., Aiken, L., 1995. The effect of different forms of centering in hierarchical linear modeling. Multivariate Behavioral Research 30, 1–20.

Kuss, O., 2002. How to use SAS for logistic regression with correlated data. SAS Users Group—27th Annual SAS Users Group International Conference or SUGI 27, Paper 261-27.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., 1996. SAS System for Mixed Models. SAS Institute Inc., Cary, North Carolina.

Maas, C.J.M., Hox, J.J., 2004. Robustness issues in multilevel regression analysis. Statist. Neerlandica 58, 127–137.

McInturff, P., Johnson, W.O., Cowling, D., Gardner, I.A., 2004. Modelling risk when binary outcomes are subject to error. Statist. Med. 23, 1095–1109.

McMahon, J.M., Tortu, S., Torres, L., Pouget, E.R., Hamid, R., 2003. Recruitment of heterosexual couples in public health research: a study protocol. BMC Medical Research Methodology 4, 1–12.

Neuhaus, J.M., 2002. Analysis of clustered and longitudinal binary data subject to response misclassification. Biometrics 58, 675–683.

Neuhaus, J.M., Kalbfleisch, J.D., 1998. Between- and within-cluster covariate effects in the analysis of clustered data. Biometrics 54, 638–645.

Newsom, J.T., 2002. A multilevel structural equation model for dyadic data. Structural Equation Modeling 9, 431–447.

Patefield, M., 2002. Fitting non-linear structural relationships using SAS procedure NLMIXED. The Statistician 51, 355–366.

Phillips, C.V., 2003. Quantifying and reporting uncertainty from systematic errors. Epidemiology 14, 459–466.

Pinheiro, J.C., Bates, D.M., 1995. Approximations to the log-likelihood function in the nonlinear mixed-effects model. J. Comput. Graphical Statist. 4, 12–35.

Ramadori, G., Meier, V., 2001. Hepatitis C virus infection: 10 years after the discovery of the virus. European Journal Gastroenterology Hepatology 13, 465–471.

Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. Sage Publications, Thousand Oaks.

Raudenbush, S.W., Brennan, R.T., Barnett, R.C., 1995. A multivariate hierarchical model for studying psychological change within married couples. J. Family Psychology 9, 161–174.

Ridout, M.S., Demetrio, C.G., Firth, D., 1999. Estimating intraclass correlation for binary data. Biometrics 55, 137–148.

SAS, 1999. SAS Procedures Guide. Version 8. SAS Inc., Cary, NC.

Snijders, T.A.B., Bosker, R.J., 1999. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Sage Publications, London.

Steenland, K., Greenland, S., 2004. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. Amer. J. Epidemiology 160, 384–392.

Ten Have, T.R., Localio, A.R., 1999. Empirical Bayes estimation of random effects parameters in mixed effects logistic regression models. Biometrics 55, 1022–1029.

Tortu, S., McMahon, J.M., Pouget, E.R., Hamid, R., 2003. She said, he said: drug-using women's perceived versus actual HIV risk from primary male partners. Presented at the 131st Annual Meeting, of APHA, November 15–19, 2003, San Francisco.

Windle, M., Dumenci, L., 1997. Parental and occupational stress as predictors of depressive symptoms among dual-income couples: a multilevel modeling approach. J. Marriage and the Family 59, 625–634.

Wolfinger, R.D., 1997. Comment: experiences with the SAS macro NLINMIX. Statist. Med. 16, 1258–1259.

Wolfinger, R.D., 1999. Fitting nonlinear mixed models with the new NLMIXED procedure. SAS Users Group—24th Annual SAS Users Group International Conference or SUGI 24, Paper 287.

Wolfinger, R.D., O'Connell, M., 1993. Generalized linear mixed models: a pseudo-likelihood approach. J. Statist. Comput. Simulation 48, 233–243.

Zou, G., Donner, A., 2004. Confidence interval estimation of the intraclass correlation coefficient for binary outcome data. Biometrics 60, 807–811.