Available Online

JOURNAL OF SCIENTIFIC RESEARCH

J. Sci. Res. **13** (2), 521-536 (2021)

www.banglajol.info/index.php/JSR

Joint Modeling for Longitudinal Data with Missing Values: A Bayesian Perspective on Human Intelligence

T. Gokul[1*], M. R. Srinivasan[2]

[1]Department of Statistics, University of Madras, Chennai, 600005, India

[2]School of Mathematics and Statistics, University of Hyderabad, Telangana, 500046, India

Received 27 November 2020, accepted in final revised form 17 February 2021

Abstract

Joint modeling in longitudinal data is an interesting area of research since it predicts the outcome with covariates that are measured repeatedly over the time. However, there is no proper methodology available in literature to incorporate the joint modeling approach for count-count response data. In addition, there are several situations where longitudinal data might not be possible to collect the complete data and the Missingness may occur due to the absence of the subjects at the follow-up. In this paper, joint modelling for longitudinal count data is adopted using Bayesian Generalized Linear Mixed Model framework to understand the association between the variables. Further, an imputation method is used to handle the missing entries in the data and the efficiency of the methodology has been studied using Markov Chain Monte-Carlo (MCMC) technique. An application to the proposed methodology has been discussed and identified the suitable nutritional supplements in Bayesian perspective without eliminating the missing entries in the dataset.

Keywords: Longitudinal data; Missing imputation; Joint model; DIC; MCMC.

doi: http://dx.doi.org/10.3329/jsr.v13i2.50479 J. Sci. Res. **13** (2), 521-536 (2021)

1. Introduction

In longitudinal studies, the research interest often lies in understanding the relationship between the longitudinal process and its effect on the risk of an event. Classical models such as separate analysis were performed for these types of data. Consequently, the association between the two longitudinal count outcomes is neglected because the linear mixed model for repeated measurements for count data are conducted separately. Thus, for modeling both components at a time, a class of models named joint models has been developed for analyzing these kinds of data. While longitudinal data, studies the association between the different outcomes through separate analysis, joint modeling allows every outcome to have its own random effects and the association can be obtained from correlation between the random effects components.

Literature is abundant in joint modeling for continuous-continuous, continuous-binary data by Thifiebaut *et al*. [1], Iddi and Molenberghs [2]. Joint modeling for different types

* *Corresponding author:* gokultvasan@gmail.com

of longitudinal outcome has been studied by Efendi *et al.* [3], Horrocks and Van Den Heuvel [4], Li *et al.* [6] and a brief review for modelling longitudinal time to event data has been discussed by Tsiatis and Davidian [5]. In Psychometric studies, Palestro *et al.* [7] provided a systematic approach of fitting joint models (direct and covariance approach) for the longitudinal data and illustrated the same using the neural and behavioural measure of cognition data. However, there is no evident in literature that joint modeling strategy is applied to the count-count response variables. Thus, this study proposed a joint modeling strategy that can be applied to count-count outcome variables.

The problem of missing data requires special attention as it refers to the case where not all data are obtained as planned in the study design. In longitudinal analysis, researchers often face the problem of setting the subjects for every follow-up visit and in several situations observations on the study subject may not be possible to obtain. Hence, there will be missing data in some of the study and poses a major challenge for the analysis of longitudinal data. Multiple imputation (MI) is widely used to handle missing values in longitudinal data. A handful of MI techniques have been proposed in literature to impute incomplete longitudinal studies and these MI approaches have been implemented in various software packages. The main advantage of MI technique is that it is not necessary to use the same model for imputation in data analysis and several software are available for this purpose.

The process of multiple imputation method is explained through the following steps:

STEP 1. Assume an imputation model and simulate "m" possible values for every missing value and it leads to the "m" complete datasets and "m" should be a positive integer usually greater than 1.

STEP 2. For the "m" complete datasets, conduct the statistical analysis and it result in "m" analysis outputs.

STEP 3. Combine those "m" statistical outputs into single output using an appropriate method.

In step 1, from the assumed model, the missing values are generated and this is usually obtained as the predictive distribution given in the observed data. For example, in a regression model, we generate the missing covariates x_{mis} from the predictive distribution $f(x_{mis} \mid y_i, x_{obs})$. However, achieving the predictive distribution requires integrating the unknown parameters thus it lead to improper results in few cases. Step 2 is clear and easier to do in a straightforward way. In step 3, combining the "m" results can be achieved using some mathematical and statistical rule or through combine the test statistics obtained from "m" results [8-10]. This approach is a common strategy to implement multiple imputation in longitudinal backgrounds, where it can be used as the data generating step. A valuable contribution to the development of missing data problem has been reviewed for GLMM by Ibrahim and Molenberghs [11]. Further, Huque *et al.* [12] outlined a special case of multiple imputation with a suitable real time dataset.

In addition, there has been an increasing amount of research is carried out in Bayesian inference for joint mean and variance models. Cepeda and Gamerman [13] discussed the variance model with the combination of Bayesian modeling and the usual regression

analysis. Lin and Wang [14] considered Bayesian joint model for longitudinal data by modeling the mean and covariance vector jointly. Further, Xu and Zhang [15] discussed the Bayesian semi-parametric joint model based on B-spline approximations. Pourahmadi [16,17] proposed the modified cholesky decomposition which decomposes the random effects covariance matrix into two sets of parameters: Generalized Autoregressive Parameters (GARPs) and the Innovation Variances (IVs). Furthermore, Daniels and Zhao [18] discussed the modified cholesky decomposition that reduces the number of parameters in the covariance matrix and is used for the estimation of the covariance matrix to analyze longitudinal Gaussian data. Lee *et al.* [19] proposed a GLMM with the heterogeneous random effects covariance matrix depending on covariates via the modified cholesky decomposition. However, Bayesian analysis for count data through the modified cholesky decomposition served as a motivation for this study. Joint modeling for mean and covariance vector is handled heuristically based on the general linear model by Pourahmadi [16,17]. Further, Zhang and Leng [20] proposed efficient maximum likelihood estimation for joint model based on the cholesky decomposition. On the other hand, with the fast development of Markov Chain Monte-Carlo (MCMC) methods, Bayesian inference for various statistical models has been receiving a lot of attention in recent years [15,21,22]. Therefore, in this paper, an extension of Bayesian joint model for longitudinal count data based on MCMC technique is considered as it gives reliable results even for the small sample size.

Poisson regression modeling becomes a traditional method for analysing univariate count data [36]. However, there has been a competing research for bivariate count data as well. It is worth mentioning that many empirical studies like McHale and Scarf [24], Osiewalski and Marzec [23] apply negative binomial regression because of the nature of count data. Another approach is to model the joint distribution using copulas. Also, Berkhout and Plug [25] proposed relatively flexible dependence structures with an idea of the mixture of independent Poisson distributions on the conditional probabilities. Ghosh *et al.* [26] developed Bayesian Zero Inflated Poisson (ZIP) models for cross-sectional data, using MCMC with data augmentation to obtain posterior samples. Karlis and Tsiamyrtzis [27] illustrated Bayesian estimation approach for the parameters of the bivariate Poisson model and provided the posterior distributions in closed forms.

In this paper, we have built a joint model for count-count response variables, that is General, G-factor as measured by Raven's Coloured Progressive Matrices (RCPM) test and other three Specific, S-factors assessing Verbal Meaning (VM), Arithmetic Score (AS) and Digit Span total (DS), with a view to study the association between cognitive function outcomes, and how it evolves over time using Bayesian strategy. There is also baseline covariate information on each subject, including Gender, Age, Time, Height and Weight and Socio-economic status. The study consists of comparing a suitable joint longitudinal model for general intelligence (G and S factors) with a set of above covariates under consideration. In this context, study on improvement of mental skills measured is correlated with the working ability and this association can only be studied if both outcomes are modeled jointly using a suitable GLMM. For ease of exposition, we

focus on Poisson models, but the methods can easily be extended to other count distributions, such as the negative binomial. Further, this study made a comparison of complete case using Bayesian mechanism with different priors.

2. Bayesian Generalized Linear Mixed Models for Longitudinal Count Data

Let Y_{1it} and Y_{2it} be the count bivariate response for subject $i,(i=1,...,N)$ at time $t,(i=1,...,T)$ and let x_{1it} and x_{2it} be the corresponding vector of covariates. We assume that each Y_{1it} and Y_{2it} is conditionally independent given random effects b_{1it} and b_{2it} the responses for different subjects are independent, and the regression model is given by

$$\log \mu_{1it}(b_{1it}) = x_{1it}\beta + b_{1it} \ , \log \mu_{2it}(b_{2it}) = x_{2it}\beta + b_{2it} \quad (1)$$

where $\mu_{it} = (\mu_{1it}, \mu_{2it})$, $b_{it} = (b_{1it}, b_{2it})$, $\mu_{it} = (b_{1it}, b_{2it}) = P((Y_{1it}, Y_{2it}) = 1; b_{1it}, b_{2it})$ and β is the $p \times 1$ vector of fixed effect, $b_{1i} = (b_{1i1},....,b_{1in_i})^T, b_{2i} = (b_{2i1},....,b_{2in_i})^T$. $b_{1i} \sim N(0, \sum_{1i}), b_{2i} \sim N(0, \sum_{2i})$ where $\sum_i = (\sum_{1i,} \sum_{2i})$ is the random effects covariance matrix and b_i is a vector of random effects values for subject i.

To solve the positive-definiteness constraint and the exponentially increasing number of parameters of $\sum_i$, we use the modified cholesky decomposition. We have

$$b_{1i1} = e_{1i1}, b_{2i1} = e_{2i1} \quad (2)$$

$$b_{1i1} = \sum_{j=1}^{t-1} \phi_{i,t} b_{1ij} + e_{1it,} \text{ for } t = 2,...,n_{1i} \ , \ b_{2i1} = \sum_{j=1}^{t-1} \phi_{i,t} b_{2ij} + e_{2it} \text{ , for } t = 2,...,n_{2i} \quad (3)$$

where $(e_{1i} = e_{1i1},.....,e_{1in_i}) \sim N(0, D_{1i}), (e_{2i} = e_{2i1},.....,e_{2in_i}) \sim N(0, D_{2i})$ with $D_i = (D_{1i,} D_{2i})$, $D_{1i} = (\sigma^2_{1i1},.....,\sigma^2_{1in_i}), D_{2i} = (\sigma^2_{2i1},.....,\sigma^2_{2in_i})$. From equations (2) and (3), we have the following matrix form as

$$T_{1i} b_{1i} = e_{1i}, T_{2i} b_{2i} = e_{2i} \ , \quad (4)$$

where $T_i = (T_{1i,} T_{2i})$ is a lower triangular matrix having ones on its diagonal and $-\phi_{i,tj}$ at its $(t, j)^{th}$ position for $j < t$. Then we have

$$T_{1i} \sum_{1i} T_{1i}^T = \text{var}(e_{1i}) = D_{1i} \ , T_{2i} \sum_{2i} T_{2i}^T = \text{var}(e_{2i}) = D_{2i} \quad (5)$$

where $D_i = (D_{1i,} D_{2i})$ is diagonal with $\sigma^2_{i,t} = \text{var}(e_{it})$ as its diagonal entries.

The generalized autoregressive parameters (GARP) are represented by ϕ, and $\sigma^2_{i,t}$ denotes the Innovation Variances (IV). For $\sum_i$ to be positive definite the IV must be positive. The parameters, GARP and IV can be modeled using time and/or subject-specific covariate vectors $w_{i,tj}$ and $h_{i,t}$ by setting

$$\phi_{i,tj} = w^T_{i.tj}\gamma, \quad \log(\sigma^2_{it}) = h^2_{i,t}\lambda, \quad (6)$$

where γ is $a \times 1$ vector of unknown dependence parameters, λ is a $b \times 1$ vector of unknown variance parameters, respectively and $w_{i,tj}$ and $h_{i,t}$ are the design vectors and are used in generalized linear mixed models [18,22,28].

2.1. *Likelihood*

The main obstacle for Bayesian inference concerning the bivariate Poisson model is that the likelihood is too complicated. Following Karlis and Tsiamyrtzis [27], let $Y_{1i}, Y_{2i}, i = 1,..n$ is distributed as Poisson $n(\lambda_1 + \lambda_2)$ and likelihood of $\lambda_1 / y_1 \cdot \lambda_2 / y_2$ is

$$L(\lambda_1/Y_1)L(\lambda_2/Y_2) = f(y_1/\lambda_1)f(y_2/\lambda_2) = \prod_{i=1}^{n} e^{-(\lambda_1+\lambda_2+\lambda_3)} \frac{\lambda_1^{y_{1i}}}{y_{1i}!}\frac{\lambda_2^{y_{2i}}}{y_{2i}!} \tag{7}$$

2.2. *Prior*

Let y_{1i} and y_{2i} bivariate, $i = (1,....,n)$ is a random sample from a Poisson distribution with mean λ_1, λ_2, and that the prior distribution of λ_1, λ_2 is a gamma distribution with parameters $\alpha \geq 0$ and $\beta \geq 0$. The gamma distribution is the conjugate prior for the Poisson parameter, and

$$g(\lambda_1 \mid y_1)g(\lambda_2 \mid y_2) \propto \left(\prod_{i=1}^{n} e^{-(\lambda_1+\lambda_2)} \lambda_1^{y_{1i}} \lambda_2^{y_{2i}}\right) \frac{\alpha^{\beta}}{\Gamma(\alpha)} (\lambda_1+\lambda_2)^{\alpha-1} e^{-(\lambda_1+\lambda_2)\beta} \tag{8}$$

$$\propto e^{-(\lambda_1+\lambda_2)(\beta-n)} (\lambda_1+\lambda_2)^{\alpha+n(\bar{y}_1+\bar{y}_2)-1}$$

In addition, we assume that the prior distributions are mutually independent:

$$f(\beta, \lambda_1, \lambda_2) = f(\beta) f(\lambda_1, \lambda_2) \tag{9}$$

2.3. *Posterior*

Let $b_i = (b_1,....,b_n)$ be the random effects. The posterior distribution of all parameters and random effects can be written as

$$f(\beta, \lambda_1, \lambda_2, b_i \mid y_i) \propto \prod_{i=1}^{n} [f(y_{1i} \mid \beta, \lambda_1, b_{1i}) f(y_{2i} \mid \beta, \lambda_2, b_{2i}) f(\beta) f(\lambda_1, \lambda_2)] \tag{10}$$

Let $\beta = (\beta_1^*,....,\beta_n^*)$ be the individual-specific parameters, and let $y = (y_1, y_2)$. Following Davidian and Giltinan [29], the full conditional distributions for random effect in equation (2) can be shown to have the following distributions

For $b_i (i = 1,...,N)$,

$$P(b_i \mid y, \beta, \eta, \lambda_1, \lambda_2) \left\{\prod_{t=1}^{n} (\mu_{1it}^c(b_{1it}, \beta))^{y_{1it}} (1-\mu_{1it}^c(b_{1it}, \beta))^{1-y_{1it}}\right\}$$

$$\times \left\{\prod_{t=1}^{n_i} (\mu_{2it}^c(b_{2it}, \beta))^{y_{2it}} (1-\mu_{2it}^c(b_{2it}, \beta))^{1-y_{2it}}\right\} \phi(b_i \mid \eta, \lambda_1, \lambda_2)$$

Hence, the posterior distribution of $\lambda_1/y_1 \cdot \lambda_2/y_2$ is a gamma distribution with $\tilde{\alpha} = \alpha + n(\bar{y}_1 + \bar{y}_2)$ and $\tilde{\beta} = (\beta + n)$. Followed by the random effect component, since all full conditionals are intractable analytically also not easily generated from, we have to construct suitable proposals for a Metropolis Hastings (MH) step as discussed [30]. The natural (conjugate) choice would be to consider independent gamma distributions. A priori independence might be convenient but not optimal to use always. For example, it might be the case that some prior elicitation procedure provided some dependent structure regarding the $\lambda's$. In such a case we would prefer to have a prior that will provide us the flexibility to incorporate the dependence structure among the parameters. At the same time, we would like to keep the computational complexity relative low by using some form of conjugate prior, which will allow us to have in closed form the posterior distribution. We have used various parameter settings and are as follows:

$a: \alpha_1 = \alpha_2 = 8, \beta_1 = \beta_2 = 1, p = (1/5,1/5,1/5,1/5,1/5)$
$b: \alpha_1 = \alpha_2 = 8, \beta_1 = \beta_2 = 1, p = (0.1,0,0,0,0.9)$
$c: \alpha_1 = \alpha_2 = 8, \beta_1 = \beta_2 = 1, p = (0.5,0,0,0,0.5)$
$d: \alpha_1 = \alpha_2 = 7, \beta_1 = \beta_2 = 1, p = (0.5,0,0,0,0.5)$
$e: \alpha_1 = 6, \alpha_2 = 8, \beta_1 = 1, \beta_2 = 3, p = (1/5,1/5,1/5,1/5,1/5)$

In practice, Gibbs sampling is implemented using WinBUGS. The MCMC algorithm simulates direct draws from the above full conditionals iteratively until convergence is achieved.

2.4. *Deviance information criterion*

The best fit of the models is usually accessed using the model selection criteria. In usual modeling practice there are different model criteria available namely, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), small-sample corrected Akaike Information Criteria (AICC), etc. Similarly, for the Bayesian modeling, there are different model criteria available namely Bayes factor, Deviance Information Criteria (DIC), Predictive loss, etc. Following Spiegelhalter *et al*. [31], in this paper, we make use of the DIC for identifying the best fitted models and it is a combination of goodness of fit and a penalty term. The deviance of the model fit is given by

$$Dev(\theta) = -2\log\{L(\theta \mid y_1) \cdot L(\theta \mid y_2)\}, \tag{11}$$

where $L(\theta \mid y_1) \cdot L(\theta \mid y_2)$ is the likelihood of $y_{1i} = (y_{11},...,y_{1N})^T, y_{2i} = (y_{21},...,y_{2N})^T$ and θ is a vector of parameters. Larger values of the deviance indicate poorer fit. The penalty term measures model complexity and is given by

$$p_D = E\{Dev(\theta) \mid y_1\}E\{Dev(\theta) \mid y_2\} - Dev\{E(\theta \mid y_1) \cdot E(\theta \mid y_2)\}.$$

The value of p_D is called the effective number of parameters. The DIC is defined as

$$DIC = Dev\{E(\theta \mid y_1) \cdot E(\theta \mid y_2)\} + 2_{P_D}$$

$$= -4E\{\{\log L(\theta \mid y_1) \mid y_1\}\{\log L(\theta \mid y_2) \mid y_2\}\} + 2\log\{L(\hat{\theta} \mid y_1) \cdot L(\hat{\theta} \mid y_2)\}. \tag{12}$$

In practice, DIC can be expressed in different ways depending on how $E\{\{Dev(\theta) \mid y_1\}\{Dev(\theta) \mid y_2\}\}$ and $Dev\{\{E(\theta \mid y_1)\}\{E(\theta \mid y_2)\}\}$ are estimated or approximated. The smaller DIC value Celeux [32] preferred to be the best model and the main advantage of this Criterion in the situation of Bayesian model selection is that the DIC can be easily calculated from the MCMC samples than calculating the AIC and BIC Criteria values from MCMC simulation.

2.5. *Multiple imputation for missing data*

Consider a regression model and let β be a $p \times 1$ vector of parameters of interest, and let $y_i = (y_{1i}, y_{2i})$ where $y_{1i} = \{y_{11}, y_{12},...,y_{1n}\}$ and $y_{2i} = \{y_{21}, y_{22},...,y_{2n}\}$ be the bivariate responses and $x = \{x_1, x_2,...,x_n\}$ be the covariates with missing values. Let $\hat{\beta} = h_1(y_i, x_{obs}, x_{mis})$ be the statistic used to estimate β, if no missing entries present in the

data, where h_1 is a known function. Let $V = Var(\hat{\beta}) = h_2(y, x_{obs}, x_{mis})$ be the variance of $\hat{\beta}$, where h_2 is also a known function. As mentioned earlier, the imputation method is to generate x_{mis} from the predictive distribution $f(x_{mis} \mid y_i, x_{obs}, \tilde{\theta})$, where the estimate $\tilde{\theta}$ is obtained using naive approach like the complete-case method. However, in this approach the uncertainty in estimating θ is not promulgated. Thus, the imputed values can be obtained by generating x_{mis} from the predictive distribution $f(x_{mis} \mid y_i, x_{obs})$ through the following Bayesian arguments as discussed in Little and Rubin [9].

The predictive distribution of the missing data given the observed data can be written as $f(x_{mis} \mid y_i, x_{obs}) = \int f(x_{mis} \mid y_i, x_{obs}, \theta) f(\theta \mid y_i, x_{obs}) d\theta$, where θ, the unknown parameter contains β. To generate proper MI of x_{mis} from the predictive distribution $f(x_{mis} \mid y_i, x_{obs})$, we first simulate θ^* from $f(\theta \mid y, x_{obs})$, and then simulate x_{mis}^* from $f(x_{mis} \mid y_i, x_{obs}, \theta^*)$, and then we iterate the procedure until the simulated value gets stabilized. This can be done by the following method, Given a starting value $\theta^{(0)}$, at iteration $k (k = 1,2,3,...)$, simulate $x_{mis}^{(k)} \sim f(x_{mis} \mid y_i, x_{obs}, \theta^{(k-1)}) \propto f(y_i \mid x, \theta^{(k-1)}) f(x \mid \theta^{(k-1)})$ through rejection sampling method as the density functions $f(y_i \mid x, \theta^{(k-1)})$ and $f(x \mid \theta^{(k-1)})$ are known. Then simulate $\theta^{(k)} \sim f(\theta \mid y_i, x_{obs}, x_{mis}^{(k)})$ which is the posterior distribution of θ.

Thus, we obtain a sequence of simulated values $\{(x_{mis}^{(k)}, \theta^{(k)}), k = 0,1,2,...\}$ which is a Markov chain (MC). The MC will converge to the stationary distribution $f(x_{mis} \mid y_i, x_{obs})$ at a certain period. After that, the last values will be (x_{mis}^*, θ^*) where $\theta^* \sim f(\theta \mid y_i, x_{obs})$ and x_{mis}^* is a proper imputation of the missing x_{mis} from the predictive distribution $f(x_{mis} \mid y, x_{obs})$. By repeating this process m times, we get m proper imputations $\{x_{mis}^{(1)}, ..., x_{mis}^{(m)}\}$ for the missing values x_{mis}. Once we generate m imputations, we have m "complete datasets" $\{y_i, x_{obs}, x_{mis}^{(l)}\}, l = 1,2,...,m$. Further, let $\hat{\beta}^{(l)} = h_1(y_i, x_{obs}, x_{mis}^{(l)})$ be the estimate of β and $V^{(l)} = h_2(y_i, x_{obs}, x_{mis}^{(l)})$ be the variance of $\hat{\beta}^{(l)}$ based on the $l-th$ "complete dataset", obtained using available complete-data methods.

Following Wu [8], we combine the m estimates and obtain the overall estimate of β as follows

$$\bar{\beta} = \frac{1}{m} \sum_{l=1}^{m} \hat{\beta}^{(l)} \tag{13}$$

and the overall estimated variance of $\bar{\beta}$ is

$$Var(\bar{\beta}) = \left(1 + \frac{1}{m}\right) B + W, \tag{14}$$

where

$$B = \frac{1}{m-1} \sum_{l=1}^{m} (\hat{\beta}^{(l)} - \bar{\beta})^2, \mathrm{W} = \frac{1}{m} \sum_{l=1}^{m} V^{(l)} \tag{15}$$

are the between-imputation variance and the within-imputation variance respectively. Therefore, in a MI method, the final overall estimate is simply the average of the m

individual estimates from the m imputed "complete" datasets, and the overall variance is the sum of the between imputation variance B and the within-imputation variance W. The between imputation variance B reflects the missing data uncertainty.

3. Data Analysis and Discussion

A controlled Kenya school-children feeding intervention study of 546 individuals was designed to test a cognitive ability among their intake of nutritional supplements. The data were collected from Neumann *et al*. [33] and sample observations are presented in appendix table 7. Each nutrition group was comprised of 9 out of 12 schools with children aged 6-14. School lunch programs are introduced in order to improve the same as they may cause an impact in their health. School-aged children who suffer from severe malnutrition exhibit significantly compromised reasoning and poorer school grades, reduced attentiveness and unresponsive play behaviour, as compared to their adequately nourished peers. In addition, children suffering from mild-to-moderate malnutrition, show significant deficits in intellectual and behavioural functioning. Deficits include compromised development in multiple domains, including verbal and spatial reasoning [34].

The school lunch intervention began at time (round) t = 0 by adding the supplements: Meat, Milk, Control and oil added as Calories to determine the effects of human intelligence outcome measures. Here response as a G factor is collected on RCPM test used to measure general human intelligence. Data were collected at five different points of time and round 1 data is the baseline data collected before the onset of intervention (in other words, pre-intervention). Round 2 was taken as soon as the intervention started, while rounds 3, 4, and 5 were during the second, fourth, and sixth months after intervention started as indicating post intervention scores. Total of 546 children were in this intervention study, out of them 284 are boys and 262 are girls children. 146 children were given calorie supplement, 131 children were given meat supplement, 142 were given milk and 127 were considered as control group in this study. Table 1 gives the descriptive summaries of treatment and outcomes observed at 5 time points. At the end of study, it is found that there are few missing entries present in the data and only 374 individual school children had a full-sequence data resulting from the fact that 172 individuals are missing after the first, second and fifth rounds of school lunch intervention.

The intervention data contains 2 % of missing entries in round 1, 6 % in round 2, 4% in round 3, 6 % in round 4, and 9 % in round 5 in all the response variables in this study. In addition, the treatment milk and meat contains 1 % missing entries in round 1, 2, and 3, 2 % of observation is missing in round 4 and 5. Similarly, calories group contains 1 % missing in round 1, 2, 3, 4 and 3 % missing in round 5. The control group contains 1 % missing data in round 1, 2, 3, 4 and 2 % missing in round 5. Multiple imputation technique has been adopted to nullify the presence of Missingness in the data.

Table 1. Descriptive summary.

		Minimum	Maximum	Mean	Std. Deviation
G factor	Raven's coloured progressive matrices	0	31	18.24	2.98
S factor	Arithmetic score	0	17	7.72	1.78
	Verbal meaning	0	40	30.03	5.28
	Digit span total	0	16	6.31	2.21
Covariates	Age	5	16	7.6	1.41
	Height	101.1	134.95	115.51	6.25
	Weight	14.3	50.15	20.15	2.99
	Head circumstance	45.4	56.6	50.58	1.4
	Socio economic status	28	211	84.06	24.88
	Read test	0	12	6.8	5.24
	Write test	0	11	5.12	4.95

The study involves four longitudinal outcome variables based on human intelligence factors measured on five time points. To capture the relationship between the responses, various assumptions about the distribution of the random effects can be made. There is also baseline covariate information on each subject including Age, Gender, Height, Weight, Head circumstance, Socio economic status intake of food supplements such as Milk, Meat, Calories and Control duration of the follow-up study, measurement of G factor in analytical ability as assessed by RCPM test and S factors involving linguistic ability as assessed by VM, numerical ability as assessed by AS and immediate memory as assessed by DS total. Obviously, it is expected that improvement of children cognitive skills is correlated with the nutrition supplements and this association is studied using GLMM.

Gokul *et al.* [35] proposed a classical joint model, for Kenya school lunch intervention study and suggested that the nutritional supplements show gradual improvement in cognitive behaviour among the students. However, classical estimation of analyzing longitudinal count data may result in influenced estimates. Thus, Bayesian modeling is adopted to understand the strength of association through DIC towards the model selection process and the results are obtained using statistical software SAS and R. Lee [22] proposed a Bayesian modeling of random effects covariance matrices for a generalized linear mixed model with modified cholesky decomposition approach and further we extended into Bayesian joint modeling into count data. In section 2, Bayesian GLMM for longitudinal count data clearly stated the prior and posterior distribution and here we illustrate this with Kenya school lunch intervention data. Upon fitting the joint mixed model, the results in studying the association between G factor in analytical ability as assessed by RCPM test and association with three S factors namely (numerical ability as assessed by AS, linguistic ability as assessed by VM, and immediate memory as assessed by DS total test) in Table 2, 3, 4, 5 and 6 with four different nutritional supplements. We applied the Bayesian approach with five different priors were adopted out of which, three of them were independent gamma priors i.e. Gamma (a_i, b_i) for each parameter $\lambda_i, i = 1,2,3$ with hyper-parameters $a_i = b_i = 0.1,1,10$ respectively, for $i = 1,2,3$ and other two are dependent priors. We denote those priors by $\pi_1, \pi_2, \pi_3, \pi_4$ and π_5 in Table

2. All of them correspond to unit prior mean and varying variance. The first choice provides diffused prior π_5 while for the last one the prior is quite informative as the variance is small and observed from the results.

Table 2. DIC for the cognitive data using the five different priors

Prior	π_1	π_2	π_3	π_4	π_5
Nut. Sup			Calories		
RCPM-AS	13528.05	13532.19	13534.66	13548.03	13523.87
RCPM-VM	14944.74	14948.06	14952.04	14964.03	14939.92
RCPM-DS	12552.49	12556.6	12558.77	12572.6	12548.28
Nut. Sup			Milk		
RCPM-AS	13527.39	13401.4	13403.61	13416.74	13393.87
RCPM-VM	14813.24	14817.81	14819.37	14832.77	14809.92
RCPM-DS	12421.41	12425.57	12427.51	12440.64	12418.28
Nut. Sup			Meat		
RCPM-AS	13176.03	13180.26	13182.15	13191.07	13172.9
RCPM-VM	14674.76	14678.25	14680.56	14689.2	14670.78
RCPM-DS	12283.82	12287.88	12289.88	12298.68	12280.67
Nut. Sup			Control		
RCPM-AS	13790.47	13794.37	13796.61	13811.94	13785.25
RCPM-VM	15187.71	15191.38	15196.78	15209.3	15182.15
RCPM-DS	12788.31	12792.02	12795.57	12809.55	12783.01

Table 2 explains the DIC value for all the joint response variables with different priors and it is evident that the results are quite similar for all the priors in every nutritional supplement. In addition, it is clear that the meat supplement helps in developing the cognitive skills (RCPM, AS, VM and DS) than other nutritional supplement (milk, calories and control). Based on this result, the analysis is extended using Bayesian framework, which deeply studies about the fixed effect and random effect component. This approach allows a complex structure of covariance matrix and satisfies the positive-definiteness of random effects covariance matrix using MCMC algorithm. Based on cognitive function, G factor as RCPM test is a random effects Poisson regression with AR(1) structure of the random effects covariance matrix. S factor such as (AS, VM and DS) are the same model with AR(2) structure of the covariance matrix. Following Lee [22] compared posterior means, DIC values for nutritional supplements for complete data along with multiple imputation are provided in Tables 3 and 4. Additionally, 95 % credible intervals for nutritional supplements using complete data are provided in Tables 5 and 6. It must be noted that the above prior, having a finite mixture representation, can also be seen as a synthesis of more priors elicited by different experts like Karlis and Tsiamyrtzis [27] with different opinions on the parameters of interest.

Table 3. DIC for G factor (RCPM) in association with S factor (AS).

Joint model	Raven's coloured progressive matrices							
Nut. Sup.	Calories		Meat		Milk		Control	
Methods	Complete	MI	Complete	MI	Complete	MI	Complete	MI
Fixed parameters β								
Intercept	0.7071	0.7555	0.5946	0.6236	0.6687	0.7074	0.7478	0.8059
Gender	0.0516	0.0892	0.053	0.079	0.0528	0.0846	0.0516	0.095
Age	0.0186	0.0582	0.0141	0.0371	0.0146	0.0459	0.0198	0.0677
Height	-0.0072	-0.0208	-0.0061	-0.0111	-0.0061	-0.0154	-0.0069	-0.0248
Weight	0.0151	0.0531	0.0139	0.0309	0.0143	0.0418	0.0156	0.0641
Head circ	0.0158	0.1004	0.0162	0.0362	0.0148	0.0671	0.0133	0.1302
SES	0.0039	0.0383	0.0038	0.0148	0.0038	0.0265	0.0039	0.05
Read test	0.0037	0.0153	0.0043	0.0123	0.0038	0.0136	0.0029	0.0163
Write test	0.0085	0.0269	0.0088	0.0228	0.0084	0.0246	0.0078	0.0284
GARP $\gamma_0(AR(1))$ $\gamma_1(AR(2))$	0.6044	0.6125	0.5593	0.566	0.5831	0.5912	0.6069	0.615
IV parameters λ_0	2.1225	2.2056	1.9027	1.975	2.1972	2.2803	2.2728	2.3559
Joint model	Arithmetic score							
Nut. Sup.	Calories		Meat		Milk		Control	
Methods	Complete	MI	Complete	MI	Complete	MI	Complete	MI
Fixed parameters β								
Intercept	-0.2531	-0.1659	-0.3049	-0.2371	-0.2596	-0.1821	-0.148	-0.0511
Gender	0.0751	0.1359	0.0755	0.1247	0.0755	0.1305	0.0736	0.1402
Age	0.0213	0.0941	0.0169	0.0731	0.0169	0.0814	0.0228	0.1039
Height	-0.0071	-0.0379	-0.006	-0.0282	-0.0057	-0.0322	-0.0066	-0.0417
Weight	0.0165	0.0965	0.0156	0.0746	0.0157	0.0852	0.0172	0.1077
Head circ	0.0362	0.25	0.0353	0.1845	0.0341	0.2156	0.0319	0.278
SES	0.0021	0.0833	0.0021	0.0599	0.002	0.0715	0.002	0.0949
Read test	0.0099	0.0287	0.0094	0.0246	0.0096	0.0266	0.0108	0.0314
Write test	0.0068	0.034	0.0066	0.0294	0.0068	0.0318	0.0077	0.0371
GARP $\gamma_0(AR(1))$	0.815	0.8328	0.7895	0.8026	0.7959	0.8184	0.8619	0.8891
$\gamma_1(AR(2))$	-0.0131	-0.005	-0.0243	-0.0162	-0.0318	-0.0237	-0.0062	0.0019
IV parameters λ_0	2.6231	2.7062	2.5904	2.6735	2.896	2.9791	3.6656	3.7487
RPCM-AS (DIC)	13523.87	14313.9	13172.9	13627.8	13393.87	13986	13785.25	14640.5

Table 4. DIC for G factor (RCPM) in association with S factor (VM, DS).

Joint model	Verbal meaning							
Nut. Sup.	Calories		Meat		Milk		Control	
Methods	Complete	MI	Complete	MI	Complete	MI	Complete	MI
Fixed parameters β								
(Intercept)	0.6413	0.7673	0.531	0.6376	0.6048	0.7211	0.6233	0.759
Sex	-0.0138	0.0702	-0.013	0.0594	-0.0132	0.065	-0.0131	0.0767
Age	0.0033	0.1093	-0.001	0.0884	-0.0002	0.0975	0.0032	0.1175
Height	0.0021	-0.0459	0.0028	-0.0366	0.0031	-0.0406	0.0025	-0.0498
Weight	0.0118	0.1338	0.0108	0.1118	0.0112	0.1227	0.0118	0.1443
Head circ	0.0112	0.3542	0.0125	0.2909	0.0105	0.3212	0.0104	0.3857
SES	0.0032	0.1312	0.0031	0.1077	0.0031	0.1194	0.0032	0.1429
Read test	0.0023	0.0283	0.0018	0.0242	0.0022	0.0264	0.0025	0.0303
Write test	0.0065	0.0425	0.0069	0.0385	0.0065	0.0403	0.0063	0.0445
GARP								
$\gamma_0(AR(1))$	0.7249	0.7496	0.6694	0.6821	0.6787	0.6974	0.7407	0.7714
$\gamma_1(AR(2))$	-0.077	-0.0689	-0.1318	-0.1237	-0.0629	-0.0548	-0.0717	-0.0636
IV parameters								
λ_0	2.3942	2.4773	2.3395	2.4226	2.4746	2.5577	2.5401	2.6232
RPCM-VM (DIC)	14939.92	15821.6	14670.78	15201.3	14809.92	15432.4	15182.15	16103.7
Joint model	Digit span total							
Nut. Sup.	Calories		Meat		Milk		Control	
Methods	Complete	MI	Complete	MI	Complete	MI	Complete	MI
Fixed parameters β								
(Intercept)	-0.9793	-0.9697	-1.0107	-1.0205	-0.9103	-0.9203	-0.7537	-0.7344
Sex	0.1945	0.2089	0.1949	0.1977	0.1918	0.2004	0.1918	0.212
Age	-0.0158	-0.0094	-0.0194	-0.0296	-0.0261	-0.028	-0.0078	0.0069
Height	-0.007	-0.0034	-0.0059	0.0063	-0.0042	0.0037	-0.0068	-0.0075
Weight	0.0238	0.0198	0.0229	-0.0021	0.022	0.0075	0.0259	0.0324
Head circ	0.0298	-0.0148	0.0283	-0.0809	0.0251	-0.0518	0.022	0.0097
SES	0.0046	-0.0078	0.0046	-0.0312	0.0042	-0.0199	0.0044	0.0037
Read test	0.0175	0.0219	0.017	0.0178	0.0176	0.0202	0.0201	0.0263
Write test	0.0097	0.0193	0.0095	0.0147	0.0101	0.0175	0.0117	0.0235
GARP								
$\gamma_0(AR(1))$	0.4888	0.4902	0.4349	0.4373	0.4622	0.4656	0.5203	0.5247
$\gamma_1(AR(2))$	0.0348	0.0387	0.0329	0.034	0.0309	0.0334	0.0404	0.0457
IV parameters								
λ_0	1.6504	1.7011	1.2178	1.2469	1.7418	1.7817	1.8184	1.8799
RPCM-DS (DIC)	12548.28	13153.9	12280.67	12583.1	12418.28	12951.3	12783.01	13506.2

Table 5. Credible intervals (95 %) for nutritional supplement based on RCPM and AS.

	Calories	Meat	Milk	Control
	Raven coloured progressive matrix			
(Intercept)	(-8.2506, 1.2869)	(-9.1281, 0.9482)	(-8.3304, 1.499)	(-8.842, 0.8753)
Sex	(-0.1904, 0.3019)	(-0.2053, 0.2835)	(-0.1564, 0.3341)	(-0.1543, 0.3513)
Age	(-0.1041, 0.1519)	(-0.1375, 0.1246)	(-0.1119, 0.1304)	(-0.108, 0.1309)
Height	(-0.0314, 0.0406)	(-0.0478, 0.028)	(-0.0414, 0.0314)	(-0.0379, 0.0345)

	Calories	Meat	Milk	Control
		Raven coloured progressive matrix		
Weight	(-0.0277, 0.0663)	(-0.0277, 0.0713)	(-0.0201, 0.0701)	(-0.0166, 0.0781)
Head circ	(-0.0197, 0.155)	(-0.0086, 0.173)	(-0.0237, 0.15)	(-0.016, 0.1666)
SES	(0.0012, 0.0117)	(0.0012, 0.0119)	(0.0004, 0.0112)	(0.0002, 0.01)
Read test	(-0.0318, 0.0415)	(-0.0175, 0.0606)	(-0.0223, 0.0534)	(-0.0209, 0.0548)
Write test	(-0.0287, 0.0293)	(-0.0326, 0.0262)	(-0.0306, 0.0283)	(-0.0319, 0.0241)
		Arithmetic Score		
(Intercept)	(-3.3922, 2.1301)	(-4.0365, 1.3746)	(-3.8335, 1.917)	(-4.167, 1.214)
Sex	(-0.0623, 0.1895)	(-0.1142, 0.1581)	(-0.0621, 0.2063)	(-0.0568, 0.1878)
Age	(-0.0253, 0.0179)	(-0.0235, 0.0196)	(-0.0194, 0.0212)	(-0.0249, 0.0157)
Height	(-0.024, 0.0065)	(-0.0175, 0.0128)	(-0.0161, 0.016)	(-0.0136, 0.0171)
Weight	(0.0052, 0.0548)	(-0.0067, 0.0456)	(-0.0117, 0.0392)	(-0.0049, 0.0457)
Head circ	(-0.0094, 0.0923)	(-0.005, 0.089)	(-0.0123, 0.0854)	(-0.0123, 0.0828)
SES	(0.0004, 0.0065)	(0.0005, 0.0064)	(0.0003, 0.006)	(0.0007, 0.0059)
Read test	(-0.046, 0.0844)	(-0.0503, 0.0869)	(-0.0527, 0.0805)	(-0.0734, 0.0783)
Write test	(-0.0156, 0.0264)	(-0.0143, 0.0272)	(-0.0106, 0.029)	(-0.0088, 0.0334)

Table 6. Credible intervals (95 %) for nutritional supplement based on VM and DS.

	Calories	Meat	Milk	Control
		Verbal meaning		
(Intercept)	(-5.974, 1.0928)	(-6.3301, 0.74)	(-6.0168, 1.7795)	(-6.7161, 1.2257)
Sex	(-0.1863, 0.1815)	(-0.1692, 0.1739)	(-0.1903, 0.1625)	(-0.2172, 0.1352)
Age	(-0.1007, 0.0802)	(-0.1053, 0.0793)	(-0.1198, 0.0866)	(-0.101, 0.0779)
Height	(-0.0194, 0.0379)	(-0.0151, 0.0425)	(-0.0172, 0.0412)	(-0.0199, 0.041)
Weight	(-0.0137, 0.0518)	(-0.018, 0.0502)	(-0.0138, 0.0547)	(-0.0115, 0.0592)
Head circ	(-0.0175, 0.1169)	(-0.0195, 0.1107)	(-0.0352, 0.1019)	(-0.0302, 0.1095)
SES	(0.0009, 0.009)	(0.0003, 0.0082)	(-0.0003, 0.0084)	(0.0007, 0.009)
Read test	(-0.021, 0.0354)	(-0.0219, 0.032)	(-0.0207, 0.035)	(-0.0203, 0.0389)
Write test	(-0.0131, 0.0291)	(-0.0112, 0.0316)	(-0.0063, 0.0383)	(-0.0126, 0.0311)
		Digit span total		
(Intercept)	(-7.122, 2.571)	(-6.7772, 1.8821)	(-8.5246, 0.7398)	(-7.1138, 1.9635)
Sex	(-0.0329, 0.4216)	(0.0158, 0.4582)	(0.0412, 0.4912)	(-0.0074, 0.4834)
Age	(-0.1106, 0.1183)	(-0.1352, 0.0971)	(-0.1675, 0.0756)	(-0.1135, 0.1298)
Height	(-0.034, 0.0227)	(-0.032, 0.0199)	(-0.0166, 0.0384)	(-0.031, 0.0235)
Weight	(-0.0374, 0.0322)	(-0.041, 0.0292)	(-0.0452, 0.0266)	(-0.0438, 0.0301)
Head circ	(-0.0351, 0.1361)	(-0.027, 0.1325)	(-0.0307, 0.131)	(-0.0381, 0.1291)
SES	(0.0001, 0.0104)	(0.0011, 0.0113)	(0.0007, 0.0114)	(0.0015, 0.0118)
Read test	(-0.0097, 0.0603)	(-0.017, 0.05)	(-0.0139, 0.0527)	(-0.011, 0.0597)
Write test	(-0.0066, 0.0809)	(-0.0184, 0.0677)	(-0.0184, 0.0727)	(-0.0042, 0.0871)

The following are the observations in the study on human intelligence:

(i). In the GARP, the coefficient γ_0 for AR(1) were positive and the credible interval was above zero which implies the significant positive relationship of random effects, the log model is used in equation (6). The coefficient for AR(2) was not significant. It indicates that the random effects covariance matrices had homogeneous AR(1) structures. In this fixed effects, the coefficients such as

gender, age, height, weight, head circumstance, socio economic status, read test, write test were significant because 95% credible intervals did not contain zero. This indicates that the estimated conditional probability of cognitive function was lower for males than females, was lower in SES group than in weight, head circumstance variable, and was higher in read test group than in write test group. The conditional probability of cognitive function increased as human intelligence level increased. Similarly, when comparing the analytical ability by RCPM test in relationship with linguistic ability by VM and immediate memory by DS total, the results revealed that Meat supplements produces smaller criteria values indicating that the supplement of meat helps in improving the intelligence on children. In additional, multiple imputation also reveals the same results as in complete data.

(ii). Lastly, in studying DIC criteria on association between the G and S factors, RCPM and DS showed better association than compared to other G and S factors by considering that the outcomes are independent and assuming that the association between the outcome vectors are captured completely by the association between the random effects considered.

On the whole, it is evident that there exists better association between the analytical ability as assessed by RCPM test and immediate memory as DS total test than other two joint outcomes. Further, the intervention study has shown that calories and control nutritional supplement is comparatively worse than the other two supplements based on Bayesian joint model analysis. It is observed that the result performs similar in complete and missing data mechanism. Moreover, calories and control nutritional supplement on food consumption might affect the mental state and cognitive behaviour deterioration in the condition of school children.

4. Conclusion

In many psychometric data, studying the human intelligence has been an interesting research as it is given least importance in univariate response. However, we illustrated a method to capture the association between bivariate responses that is count-count in nature among the covariates simultaneously through the most powerful tool named MCMC. Hence, this study focused on studying the association between the G and S factors of human intelligence from a school lunch intervention study from rural Kenya through Bayesian perspective. Usually, joint model may lead to poor estimation if the data involves a more number of parameters, whereas the Bayesian inference with different priors provides reliable and better results. In addition, missing entries in the data is often a common problem in the case of longitudinal data analysis. Thus, we adopted a missing imputation technique to handle the missing entries in the data. Further, a comparative study has been made with complete and missing cases and found that both the cases yield similar results that meat supplement turns out to be the suitable nutritional supplement in developing cognitive skills among school children.

Moreover, the Kenya intervention study suggests that understanding the nutritional basis of food effects on cognition will help us for the decision how best to control diet to improve neuron tolerance to insults and promote mental health. Digit Span is a reliable and valid measure of attention assessing 'S' factor of human intelligence. Based on the DIC value of the bivariate analysis, the influence of 'G' factor is quite likely to be felt more on attention (Digit Span) total immediate memory showed better association than the other two 'S' factors deals with higher order cognitive functions involving reasoning abilities on linguistic and numerical domains. However, in practice, there are other factors such as stress, depression, anxiety, parental behaviour, etc., affects the improvement of intelligent factor among the school children. Thus, this analysis can be extended with more covariates related to the study and could be extended with a comparison of different missing data methods such as MAR, MCAR and MNAR techniques.

References

1. R. Thifiebaut, H. Jacqmin-Gadda, G. Chene, C. Leport, and D. Commenges, Comp. Meth. Prog. Biomed. **69**, 249 (2002). https://doi.org/10.1016/S0169-2607(02)00017-2
2. S. Iddi and G. Molenberghs, J. App. Stat. **39**, 2413 (2012). https://doi.org/10.1080/02664763.2012.711302
3. A. Efendi, G. Molenberghs, E. N. Njagi, and P. Dendale, Biomet. J. **55**, 572 (2013). https://doi.org/10.1002/bimj.201200159
4. J. Horrocks and M. J. V. Den Heuvel. Bayesian Analysis **4**, 523 (2009). https://doi.org/10.1214/09-BA419
5. A. A. Tsiatis and M. A. Davidian, Stat. Sinica **14**, 809 (2004).
6. N. Li, R. M. Elashofi, G. Li, and J. Saver, Stat. Med. **29**, 546 (2010). https://doi.org/10.1002/sim.3798
7. J. J. Palestro, G. Bahg, P. B. Sederberg, Z. –L. Lu, M. Steyvers, and B. M. Turner, J. Math. Psy. **84**, 20 (2018). https://doi.org/10.1016/j.jmp.2018.03.003
8. L. Wu, Mixed Effects Models for Complex Data. (Chapman and Hall/CRC, New York, 2009).
9. R. J. A. Little and D. B. Rubin, Statistical Analysis with Missing Data, 2nd Edition (Wiley, New York, 2002).
10. D. B. Rubin, Multiple Imputation for Nonresponse in Surveys. (Wiley, University of Michigan, 1987).
11. J. B. Ibrahim and G. Molenberghs, Test **18**, 1 (2009). https://doi.org/10.1007/s11749-009-0138-x
12. M. H. Huque, J. B. Carlin, and J. A. Simpson, BMC. Med. Res. Meth. **18**, 168 (2018). https://doi.org/10.1186/s12874-018-0615-6
13. E. Cepeda and E. Gamerman, Braz. J. Prob. Stat. **14**, 207 (2000).
14. T. I. Lin, and W. L. Wang, J. Stat. Plan Inf. **141**, 1543 (2011). https://doi.org/10.1016/j.jspi.2010.11.001
15. D. Xu, and Z. Zhang, Stat. Probability Lett. **83**, 1624 (2013). https://doi.org/10.1016/j.spl.2013.02.023
16. M. Pourahmadi, Biometrika **86**, 677 (1999). https://doi.org/10.1093/biomet/86.3.677
17. M. Pourahmadi, Biometrika **87**, 425 (2000). https://doi.org/10.1093/biomet/87.2.425
18. J. M. Daniels and Y. D. Zhao, Stat. Med. **22**, 1631 (2003). https://doi.org/10.1002/sim.1470
19. K. Lee, J. K. Yoo, J. Lee, and J. Hagan, Comp. Stat. Data Anal. **56**, 1545 (2012). https://doi.org/10.1016/j.csda.2011.09.011
20. W. Zhang and C. Leng, Biometrika **99**, 141 (2012). https://doi.org/10.1093/biomet/asr068
21. X. D. Chen and N. S. Tang, Comp. Stat. Data Anal. **54**, 2145 (2010). https://doi.org/10.1016/j.csda.2010.03.022

22. K. Lee, Comm. Stat. App. Meth. **20**, 235 (2013). https://doi.org/10.5351/CSAM.2013.20.3.235
23. J. Osiewalski and J. Marzec, Comp. Stat. **34**, 153 (2019). https://doi.org/10.1007/s00180-018-0828-5
24. A. Mchale and P. Scarf, Statistica Neerlandica **61**, 432 (2007). https://doi.org/10.1111/j.1467-9574.2007.00368.x
25. P. Berkhout and E. Plug, Statistica Neerlandica **58**, 349 (2004). https://doi.org/10.1111/j.1467-9574.2004.00126.x
26. S. K. Ghosh, P. Mukhopadhyay, and J. C. Lu, J. Stat. Plan. Inf. **136**, 1360 (2006). https://doi.org/10.1016/j.jspi.2004.10.008
27. D. Karlis and P. Tsiamyrtzis, Stat. Comp. **18**, 27 (2008). https://doi.org/10.1007/s11222-007-9035-x
28. K. Lee and S. Sung, Comm. Stat. App. Meth. **21**, 169 (2014). https://doi.org/10.5351/CSAM.2014.21.2.169
29. M. Davidian and D. M. Giltinan, Nonlinear Models for Repeated Measurements Data (Chapman and Hall, London, UK, 1995).
30. D. Gamerman, Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference (Chapman and Hall, London, 1997).
31. D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. V. D. Linde, J. Royal Stat. Soc. Series (B-Stat. Meth.) **64**, 583 (2002). https://doi.org/10.1111/1467-9868.00353
32. G. Celeux, F. Forbesy, C. P. Robertz, and D. M. Titteringtonx, Bayesian Anal. **1**, 651 (2006). https://doi.org/10.1214/06-BA122
33. G. Neumann, N. O. Bwibo, S. P. Murphy, M. Sigman, S. Whaley, L. H. Allen, D. Guthrie, R. E. Weisszz, and M. W. Demment, J. Nut. **133**, 3941S (2003). https://doi.org/10.1093/jn/133.11.3941S
34. L. Pope, E. Roche, C. B. Morgan, and J. Kolodinsky, Pre. Med. Rep. **12**, 152 (2018). https://doi.org/10.1016/j.pmedr.2018.09.010
35. T. Gokul, M. R. Srinivasan, and M. Gallo, Modelling Human Intelligence using Mixed Model Approach, in Data Science and Social Research II, ed. P. Mariani et al. (Springer, Switzerland, 2021) pp. 199-217. https://doi.org/10.1007/978-3-030-51222-4_16
36. R. Shanker and K. K. Shukla, J. Sci. Res. **11**, 1 (2019). http://dx.doi.org/10.3329/jsr.v11i1.35745

Appendix

Table 7. Kenya school-children feeding intervention data

id	treatment	rn	rcpm	as	vm	ds	sex	age	height	weight	head_circ	ses	readtest	writetest
1	meat	1	15	5	25	6	girl	7.19	110.15	17.8	50.25	89	12	11
1	meat	2	19	7	39	8	girl	7.19	110.15	17.8	50.25	89	12	11
1	meat	3	21	7	33	7	girl	7.19	110.15	17.8	50.25	89	12	11
1	meat	4	18	8	37	7	girl	7.19	110.15	17.8	50.25	89	12	11
1	meat	5	21	10	37	8	girl	7.19	110.15	17.8	50.25	89	12	11
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
546	control	1	21	11	29	7	boy	7.67	120.15	21.7	49.9	86	12	11
546	control	2	19	10	23	9	boy	7.67	120.15	21.7	49.9	86	12	11
546	control	3	18	9	31	8	boy	7.67	120.15	21.7	49.9	86	12	11
546	control	4	19	9	31	7	boy	7.67	120.15	21.7	49.9	86	12	11
546	control	5	16	10	37	10	boy	7.67	120.15	21.7	49.9	86	12	11