P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.28 No.4 October - December 2023
Page: [20-25]
Original research article

Marginal Regression Models for Mixed Bivariate Responses

Chijioke Joel Nweke[1,*], Kelechi Charity Ojide[1], George Chinanu Mbaeyi[1], Fidelis Ifeanyi Ugwuowo[2]

[1]*Department of Mathematics and Statistics, Alex Ekwueme Federal University, Ndufu-Alike 48213, Nigeria*
[2]*Department of Statistics, University of Nigeria, Nsukka 410001, Nigeria*

Received 7 May 2023; Received in revised form 8 November 2023
Accepted 1 December 2023; Available online 27 December 2023

ABSTRACT

A new framework for marginal regression model with bivariate responses from different distributions was proposed in this study. It adopts a Generalized Estimating Equation (GEE) approach of model estimation. A framework for mixture of response variables from different distributions was proposed for “Normal and Poisson”, “Normal and Bernoulli”, and “Poisson and Bernoulli”. Application on the proposal framework was examined in measuring the effect of certain hospital inputs on hospital performance in three selected tertiary health institutions.

Keywords: Bernoulli and normal; Generalized estimating equation; Marginal model; Mixed responses; Poisson

1. Introduction

Modeling useful and informative relationships within possible groups of research variables of interest has been a common and standard practice in diverse fields of study. Several forms of regression models are abundantly available in the literature for this purpose. However, in situations where more than one response variable is considered in a model, most univariate approaches to parameter estimation are deemed inefficient and replaced with appropriate multivariate approaches [1-5]. An even more interesting and daunting situation in multivariate modeling is when the response variables comprise a mixture of discrete and continuous variables [6]. Such mixed response variables are common in health sectors, social sciences and economics [6, 7]. In modeling the above scenarios, there is always an assumption of independence of observations.

However, such assumptions only make sense in cross-sectional studies, not in the case of time series and longitudinal data [8]. This would then necessitate an approach that will account for dependence among

***Corresponding author:** chijioke.nweke@funai.edu.ng
doi: 10.14456/scitechasia.2023.68

Repeated observations over time in a longitudinal observed response.

When such a response is drawn from a population that is normally distributed, an extended generalized linear model is applied [2]. Fitzmaurice [9] identified three broads, but quite distinct, classes of regression models for longitudinal data: (i) marginal or population averaged models, (ii) random-effects or subject-specific models, and (iii) transition or response conditional models. These models differ not only in how the correlation among the repeated measures is accounted for, but also have regression parameters with discernibly different interpretations. When the interest is to ascertain the effect of the covariate on the average population of the response, the marginal model is considered more appropriate [9-11]. Furthermore, for a marginal model, using the Generalized Estimating Equation (GEE) approach developed by [12, 13] to model estimation is considered more appropriate than the Maximum Likelihood Method for various reasons [9]. This study proposes the framework for mixed response variables in a panel data regression analysis when a marginal model approach is adopted.

2. Methodology

Let y_{it} be the i^{th} panel with t^{th} time response variable and X_{it} be the corresponding vector of covariates. If y_{it} is from an exponential family of distribution, then:

$$f\left(y_{it};\theta,\varphi\right)=\exp\left\{\frac{y_{it}\theta_{it}-b(\theta_{it})}{\alpha(\varphi)}+c\left(y_{it},\varphi\right)\right\}, \quad (2.1)$$

where θ is the location parameter of the distribution, $\alpha(\varphi)$ is the scale or dispersion parameter, $c\left(y_{it},\varphi\right)$ is a function that represents the normalizing term, $b'\left(\theta_{it}\right)=E\left(y_{it}\right)=\mu_{it}$ represents the mean of y_{it} and $b''\left(\theta_{it}\right)=Var\left(y_{it}\right)$ is the variance of y_{it}.

Hardin and Hilbe [10] gave the Generalized Estimating Equation for Marginal Model as:

$$\Psi(\beta)=\left[\left\{\sum_{i=1}^{n}\sum_{t=1}^{n_i}\left(\frac{y_{it}-\mu_{it}}{\alpha(\varphi)Var\left(\mu_{it}\right)}\right)\left(\frac{\partial\mu_{it}}{\partial\eta_{it}}\right)x_{jit}\right\}\right]_{p\times 1}=[0]_{p\times 1}, \quad (2.2)$$

where y_{it} and μ_{it} retains its definition in Eq. (2.1), $x_{jit}=j^{th}$ is the covariate at panel i in time t, η_{it} represents the link function that relates the parameter μ_{it} to covariates, $Var\left(\mu_{it}\right)$ is the variance function as a function of μ_{it} and $\alpha(\varphi)$ is the scale parameter, β is the marginal regression model parameters, and n is the number of panels while p is the number of covariates. Eq. (2.2) in terms of the panels can be represented as:

$$\Psi(\beta)=\left[\left\{\sum_{i=1}^{n}X_{ji}^{T}D\left(\frac{\partial\mu}{\partial\eta}\right)\left[Var\left(\mu_i\right)\right]^{-1}\left(\frac{Y_i-\mu_i}{\alpha(\varphi)}\right)\right\}\right]_{p\times 1}=[0]_{p\times 1}, \quad (2.3)$$

where $\mathbf{X}_i=\left(X_{i1},X_{i2},\ldots,X_{in_i}\right)^T$ for $j=1,2,\ldots,p$ is the matrix of covariates for panel i, $\mathbf{Y}_i=\left(Y_{i1},Y_{i2},\ldots,Y_{in_i}\right)$ is a vector of the response variable for panel i, $D(\cdot)$ is the diagonal matrix.

The variance function is $Var\left(\boldsymbol{\mu}_i\right)$ diagonal matrixes which can be decomposed thusly:

$$Var\left(\boldsymbol{\mu}_i\right)=\left[D\left(Var\left(\mu_{it}\right)\right)^{1/2}R\left(\boldsymbol{\alpha}\right)_{n_i\times n_i}D\left(Var\left(\mu_{it}\right)\right)^{1/2}\right]_{n_i\times n_i}. \quad (2.4)$$

Hence, $Var\left(\mathbf{Y}\right)=\varphi V\left(\boldsymbol{\mu}_i\right)$ where $D\left(\mu_{it}\right)$ is a diagonal matrix with $Var\left(Y_{it}\right)$ along the diagonal and $R\left(\boldsymbol{\alpha}\right)=Corr\left(\mathbf{Y}_i\right)$ is the correlation matrix as a function of $\boldsymbol{\alpha}$.

2.1 Estimation of multivariate marginal model

2.1.1 Mixed Normal and Poisson responses

Let $(\mathbf{Z_i}, \mathbf{Y}_i)$ be mixed Poisson (discrete) and normal (continuous) responses with n_i observations for n independent individuals (units or panels). Where $\mathbf{Y_i} = Y_{i1}, Y_{i2}, ..., Y_{in_i}$ and $\mathbf{Z}_i = Z_{i1}, Z_{i2}, ...,$ Z_{in_i}. Let $\mathbf{X}_i = \left(X_{i1}, X_{i2}, \ldots, X_{in_i}\right)$ be a $p \times n_i$ matrix of covariates, where $X_{ij} = p \times 1$ vector. Then from Eq. (2.1)

$$f\left(z_i / X_i\right) = \exp\left\{z_i \ln \lambda_i - \lambda_i - \ln z_i !\right\}, \quad (2.5)$$

where $\lambda_i = E\left(Z_i\right) = \mu_{1i} = Var\left(Z_i\right)$ is i^{th} panel mean and variance for the response variable Z and

$$f\left(y_i / Z_i; x_i\right) = \exp\left\{-\frac{1}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left[y_i - \mu_{2i} - \Gamma\left(z_i - \mu_{1i}\right)\right]^2\right\}, \quad (2.6)$$

where $\mu_{2i} = E\left(Y_i\right)$ is i^{th} panel mean for the response variable Y, σ^2 is the variance of Y and Γ represents the parameter of regression of Y_i on X_i.

Therefore, the joint distribution is given by:

$$\begin{aligned} f\left(z_i, y_i\right) &= f\left(z_i\right) f\left(y_i / z_i\right) \\ &= \exp\left\{z_i \ln \lambda_i - \lambda_i - \ln z_i !\right\} \\ &\times \exp\left\{-\frac{1}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(y_i - \mu_{2i} - \Gamma\left(z_i - \mu_{1i}\right)\right)^2\right\}. \end{aligned} \quad (2.7)$$

Building a Marginal Regression Model for Poisson and Normal responses, we have:

$$\ln E\left(Z_i\right) = \ln \lambda_i = \mathbf{X}_i^T \boldsymbol{\beta}_1, \quad (2.8)$$

and

$$E\left(Y_{ij} / Z_i\right) = \mathbf{X}_{ij}^T \boldsymbol{\beta}_2 + \Gamma_1\left(Z_{ij} - \mu_{1ij}\right) + \Gamma_2 S_i, \quad (2.9)$$

where $S_i = \sum_{t=1}^{n_i}\left(Z_{it} - \mu_{1it}\right), \boldsymbol{\beta} = \left(\boldsymbol{\beta_1}, \boldsymbol{\beta_2}\right)$ are sets of regression parameters, $\boldsymbol{\Gamma} = \left(\Gamma_1, \Gamma_2\right)$ are parameter vectors that induce correlation between $\mathbf{Y_i}$ and $\mathbf{Z_i}$ (With Γ_1 characterizing the association between the responses on the observation within a panel, whereas Γ_2 characterizes this association for different observations within the same panel.

For convenience of notation, let

$$\mathbf{Q_{ij}} = \left(X_{ij}^T, Z_{ij} - \mu_{1ij}, S_i\right)^T, \quad (2.10a)$$

and

$$\boldsymbol{\delta} = \left(\beta_2, \Gamma_1, \Gamma_2\right)^T. \quad (2.10b)$$

Hence, Eq. (2.9) becomes:

$$E\left(Y_{ij} / Z_i\right) = \mathbf{Q_{ij}}\boldsymbol{\delta}. \quad (2.11)$$

Applying Eq. (2.3), we have

$$\Psi\left(\beta\right) = \sum_{i=1}^{n}\begin{pmatrix} \frac{\partial E\left(Z_i\right)}{\partial \beta_1} & \frac{\partial E\left(Y_i / Z_i\right)}{\partial \beta_1} \\ \frac{\partial E\left(Z_i\right)}{\partial \delta} & \frac{\partial E\left(Y_i / Z_i\right)}{\partial \delta} \end{pmatrix} \times Cov^{-1}\begin{pmatrix} Z_i \\ Y_i / Z \end{pmatrix}\begin{pmatrix} z_i - E\left(Z_i\right) \\ y_i - E\left(Y_i / Z_i\right) \end{pmatrix}. \quad (2.12)$$

Hence Eq. (2.12) yields the following:

$$\frac{\partial E\left(Z_i\right)}{\partial \beta_1} = X_i^T E\left(Z_i\right) = X_i^T Var\left(Z_i\right), \quad (2.12a)$$

$$\frac{\partial E\left(Z_i\right)}{\partial \boldsymbol{\delta}} = 0, \quad (2.12b)$$

$$\begin{aligned} \frac{\partial E\left(Y_i / Z_i\right)}{\partial \beta_1} &= \left(\Gamma_1\right)\left(-X_i^T\right)E\left(Z_i\right) + \left(\Gamma_2\right)\left(-X_i^T\right)E\left(Z_i\right) \\ &= \left(\Gamma_1\right)\left(-X_i^T\right)Var\left(Z_i\right) + \left(\Gamma_2\right)\left(-X_i^T\right)Var\left(Z_i\right) \\ &= -\left(\Gamma_1 + \Gamma_2\right)\left(X_i^T\right)E\left(Z_i\right) = -\left(\Gamma_1 + \Gamma_2\right)\left(X_i^T\right)Var\left(Z_i\right), \end{aligned} \quad (2.12c)$$

$$\frac{\partial E\left(Y_i / Z_i\right)}{\partial \boldsymbol{\delta}} = \mathbf{Q_{ij}} = \left(X_i^T, Z_i - \mu_{1i}, S_i\right)^T. \quad (2.12d)$$

Putting the repeated nature of the observation in consideration and from Eq. (2.4), we have

$$Cov\left(Z_i\right) = V_{1i} = \boldsymbol{\Delta}_i^{1/2} R\left(\boldsymbol{\alpha}\right)\boldsymbol{\Delta}_i^{1/2}, \quad (2.12e)$$

and

$$Cov\left(Y_i / Z_i\right) = V_{2i} = \sigma^2 R\left(\boldsymbol{\alpha}\right), \quad (2.12f)$$

where $\boldsymbol{\Delta_i}$ is a diagonal matrix with elements $Var\left(Z_i\right) = \mu_{1i} = E\left(Z_i\right)$ and $\sigma^2 = Var\left(\mathbf{Y}_i\right)$.

Substituting Eqs. (2.12a)-(2.13f) to Eq. (2.12), we have system of equation given by:

$$\Psi(\beta)=\sum_{i=1}^{n}\begin{pmatrix} X_i^T Var(Z_i) & -(\Gamma_1+\Gamma_2)X_i^T Var(Z_i) \\ 0 & \mathbf{Q_i}\end{pmatrix}$$
$$\times\begin{pmatrix} V_{1i}^{-1} & 0 \\ 0 & V_{21i}^{-1}\end{pmatrix}\begin{pmatrix} z_i - E(Z_i) \\ y_i - E(Y_i)\end{pmatrix}=(\mathbf{0}). \quad (2.14)$$

2.1.2 Mixed Normal and Bernoulli responses

Let $\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i$ retain the same meaning it has in section 2.1.1 with $\mathbf{Z}_i$ being Bernoulli (discrete) responses with n_i observations for n independent individuals (units or panels). Then,

$$f(z_i / x_i)=\exp\left\{z_i \ln\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right)+\ln(1-\mu_{1i})\right\}, \quad (2.15)$$

and

$$f(y_i / z_i; x_i)=\exp\left\{-\frac{1}{2}\ln\left(2\pi\sigma^2\right) -\frac{1}{2\sigma^2}\left[y_i-\mu_{2i}-\Gamma(z_i-\mu_{1i})\right]^2\right\}. \quad (2.16)$$

Therefore, the joint distribution is given by

$$\begin{aligned} f(z_i, y_i) &= f(z_i) f(y_i / z_i) \\ &= \exp\left\{z_i \ln\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right)+\ln(1-\mu_{1i})\right\} \\ &\times\exp\left\{\begin{aligned}&-\frac{1}{2}\ln\left(2\pi\sigma^2\right)\\ &-\frac{1}{2\sigma^2}\left(y_i-\mu_{2i}-\Gamma(z_i-\mu_{1i})\right)^2\end{aligned}\right\}. \end{aligned} \quad (2.17)$$

Following the same procedure as in section 2.1.1, the Marginal Regression Model for Bernoulli and Normal response is given by

$$\ln\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right)=\text{logit}(\mu_{1i})=\text{logit}\left[E(\mathbf{Z}_i)\right]=\mathbf{X}_i^T\boldsymbol{\beta}_1, \quad (2.18a)$$

and

$$E(Y_{ij} / Z_i)=\mathbf{X}_{ij}^T\boldsymbol{\beta}_2+\Gamma_1(Z_{ij}-\mu_{1ij})+\Gamma_2 S_i. \quad (2.18b)$$

Therefore, Eq. (2.14) for mixed Bernoulli and Normal has $Var(Z_i)=\mu_{1i}(1-\mu_{1i})$ and $\boldsymbol{\Delta}_i =$ a diagonal matrix with elements $Var(Z_i)=\mu_{1i}(1-\mu_{1i})$.

2.1.3 Mixed Poisson and Bernoulli

Let $\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i$ retain the same meaning they have in section 2.1.1 with $\mathbf{Z}_i$ being Bernoulli and $\mathbf{Y}_i$ being Poisson responses with n_i observations for n independent individuals (units or panels). Then,

$$f(z_i / x_i)=\exp\left\{z_i \ln\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right)-\ln(1-\mu_{1i})\right\}, \quad (2.19)$$

and

$$f(y_i / z_i; x_i)=\exp\left\{\begin{aligned}&y_i \ln(\mu_{1i}\mu_{2i})\\ &-(\mu_{1i}\mu_{2i})-\ln(y!)\end{aligned}\right\}. \quad (2.20)$$

Therefore, the joint distribution is given by as:

$$\begin{aligned} f(z_i, y_i) &= f(z_i) f(y_i / z_i) \\ &= \exp\left\{z_i \ln\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right)+\ln(1-\mu_{1i})\right\} \\ &\times\left\{\exp\left[y_i \ln(\mu_{1i}\mu_{2i})-(\mu_{1i}\mu_{2i})-\ln(y!)\right]\right\}. \end{aligned} \quad (2.21)$$

Following the same procedure in sections 2.1.1 and 2.1.2, the Marginal Regression Model for Bernoulli and Poisson responses is given by

$$\ln\left(\frac{\mu_{1i}}{1-\mu_{1i}}\right)=\text{logit}(\mu_{1i})=\text{logit}\left[E(\mathbf{Z}_i)\right]=\mathbf{X}_i^T\boldsymbol{\beta}_1, \quad (2.22)$$

and

$$\ln E(Y_{ij} / Z_i)=\ln(\mu_{1i}\mu_{2i})=\mathbf{X}_{ij}^T\boldsymbol{\beta}_2. \quad (2.23)$$

Therefore, Eq. (2.14) for mixed Bernoulli and Poisson has $Var(Z_i)=\mu_{1i}(1-\mu_{1i})$ and $\boldsymbol{\Delta}_i =$ a diagonal matrix with elements $Var(Z_i)=\mu_{1i}(1-\mu_{1i})$.

The correlation matrix $R(\boldsymbol{\alpha})$ also called working correlation with vector parameter $\boldsymbol{\alpha}$ when properly specified equals $Cov(\bullet)$ and gives an efficient estimate. The closer the estimate of $\hat{R}(\boldsymbol{\alpha})$ is to the population $R(\boldsymbol{\alpha})$, the more efficient the parameter estimate of the model is.

To ascertain the appropriate working correlation matrix $R(\boldsymbol{\alpha})$ to account for the correlation of the repeated observation and the relevant covariates, we adopt a goodness-of-fit model test. Hardin and Hilbe [11] recommended the use of Quasi-likelihood information criterion (QIC) which is analogous to Akaike Information Criterion (AIC) for likelihood-based models. The information criterion is given by

$$QIC = -2\wp\left[g^{-1}\left(\mathbf{X}\beta_{\mathbf{R}}\right)\right] + 2trace\left[\mathbf{A}^{-1}\left(\beta_{\mathrm{R}}\right)V_{MS,R}\right], \tag{2.19}$$

where $2\wp\left[g^{-1}\left(\mathbf{X}\beta_{\mathbf{R}}\right)\right]$ is the value of the quasi-likelihood computed using coefficients from the model with hypothesized correlation structure $\mathbf{R}, \mathbf{A}$ is the variance matrix obtained by fitting an independent model, $V_{MS,R}$ is the modified sandwich estimate of the variance from the model with hypothesized correlation structure $\mathbf{R}$. Details of different existing working correlation can be seen in [10, 12].

3. Results and Discussion

We will illustrate the above methodology using data collected on number of outpatients, number of inpatients, number of active hospital beds, number of doctors, number of nurses, number of other medical staff and number of discharges. The derived variable that was adopted is Average Length of Stay (ALS). The response variables are number of discharges (Poisson), Average Length of Stay (ALS) (normal) while the covariates are number of doctors, number of other medical staff, number of active beds and type of hospital (Federal owned hospital = 1, Others = 0). The data was collected from Federal Teaching Hospital Abakaliki, University of Nigeria Teaching Hospital Enugu, and Enugu State University Teaching Hospital from 2010 to 2016. See Aloh et al. [14]. The average length of stay is derived as

$$\text{ALS} = \frac{\text{Occupied Bed} - \text{Total Inpatient Days}}{\text{Number of Discharges}} \times 100\%.$$

Table 1. The fit of ALS and Number of discharges on selected input variables plus type of hospitals (mixed Poisson and Normal).

Coefficient	Independent Working correlation		Exchangeable working correlation		AR(1) working correlation		Unstructured working correlation	
	Number of Discharges	ALS	Number of Discharges	ALS	Number of Discharges	ALS	Number of Discharges	ALS
β_0	7.507* (0.0124)	10.420* (1.1698)	7.743* (0.0130)	8.903* (1.5165)	8.139* (0.0117)	10.103* (0.8464)	7.478* (0.0095)	11.079* (0.2515)
β_1	-0.147* (0.0076)	-1.902* (0.9536)	-0.294* (0.0104)	-1.187 (0.7832)	-0.314* (0.0119)	-1.381 (0.9562)	-0.241* (0.0068)	-1.951* (0.5241)
β_2	0.004* (5.93E-5)	-0.009 (0.0089)	0.003* (5.26E-5)	0.0100* (0.0049)	0.001* (4.01E-5)	0.014* (0.0074)	-0.004* (5.59E-5)	0.043* (0.0015)
β_3	-0.002* (6.17E-5)	0.17* (0.0010)	0.001* (5.67E-5)	-0.002 (0.0028)	0.0001* (1.22E-5)	-0.004* (0.0014)	0.002* (1.58E-5)	-0.013* (0.0016)
β_4	0.001* (1.95E-5)	-0.005 (0.0028)	9.980E-5* (1.75E-5)	-0.005 (0.0107)	0.001* (3.57E-5)	-0.009 (0.0065)	0.004* (4.48E-5)	-0.015* (0.0022)
Γ_1	-	0.0001* (0.0001)	-	-0.001* (9.57E-5)	-	0.0001* (8.23E-5)	-	0.0001* (2.24E-5)
Γ_2	-	0.0001* (3.09E-5)	-	0.0001* (7.95E-6)	-	-5.29E-5* (11.47E-5)	-	0.0001* (7.12E-6)
QIC	27234.89	53.82	25313.12	50.897	36306.12	58.79	25996.80	59.24
QIC for Univariate		137.136		195.336		157.409		2876.927

( )-standard error, X_1 = Type of hospital, X_2 = BED, X_3 = Number of Doctors, X_4 = Number of other medical officers, ALS = Average length of stay, AR = Autoregressive, *-significant at 5%.

Table 1 is the result of the estimate of the multivariate marginal model with mixed Poisson and Normal responses. An independent, exchangeable AR(1) and unstructured working correlation was adopted. The result showed that all the covariates under consideration have a significant effect on the average population of number of discharges for all different working correlations, but such was not the case with average length of stay (ALS). Based on QIC, the exchangeable working correlation gave the best fit for the model estimate. For each working correlation, the multivariate model was found to be significantly better than the univariate. This is evident by comparing the values of the QIC for the univariate model against that of the multivariate.

4. Conclusion

A bivariate marginal regression model for mixed responses is examined in this study. A framework for mixed "Normal and Poisson", "Normal and Bernoulli" and "Poisson and Bernoulli" using generalized estimating equation methodology was derived. The model was applied in modeling the effects of the input variables on the output variables; where average length of stay serves as a normal response while number of discharges serves as a Poisson response. It was discovered that the multivariate marginal regression model performed far better than the univariate equivalent.

References

[1] Inouye G, Yang E, Allen G, Revikumar P. A Review of Multivariate Distributions for Count Data Derived from the Poisson Distribution. Wiley Interdisciplinary Rev Computational Statistics, 2017;9(3):1398.

[2] McCullagh P, Nelder JA. Generalized Linear Models, 2nd eds. Chapman & Hall/CRC, UK 1989.

[3] Rancher AC, Christensen WF. Methods of Multivariate Analysis. 3th eds. John Wiley and sons, USA 2012.

[4] Todorov H, White ES, Gerber S. Applying univariate vs. multivariate statistics to investigate therapeutic efficacy in (pre)clinical trials: A Monte Carlo simulation study on the example of a controlled preclinical neurotrauma trial. PLOS ONE, 2020; 15(3): e0230798.

[5] Yoo K, Rosenberg MD, Noble S, Scheinost D, Constable RT, Chun MM. Multivariate approaches improve the reliability and validity of functional connectivity and prediction of individual behaviors. Neuroimage, 2017;197:212-23.

[6] Fitzmaurice GM, Laird NM. Regression Models for a Bivariate Discrete and Continuous Outcome with Clustering. Journal of the American Statistical Association 90, 1995:845-52.

[7] Ying Y, Kang J, Mao K, Zhang J. Regression models for mixed Poisson and continuous longitudinal data. Stat Med. 2007;26(20):3782-800.

[8] Wang M. Generalized Estimating Equations in Longitudinal Data Analysis: A Review and Recent Developments. Advances in Statistics, 2014;(1):1-11.

[9] Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, editors. Longitudinal data analysis. CRC press; 2008 Aug 11.

[10] Liu X. Methods and applications of Longitudinal Data Analysis. Higher Education Press., Elsevier, Oxford OX5 1GB, UK, 2016.

[11] Hardin JW, Hilbe JM. Generalized Estimating Equations. 2nd eds. Chapman & Hall/CRC, UK 2013.

[12] Liang K-Y, Zeger SL. Longitudinal Data Analysis Using Generalized Linear Models. Biometrika, 1986;73(1):13-22.

[13] Zeger SL, Liang K-Y. Longitudinal data analysis for discrete and continuous outcomes. Biometrics, 1986;42:121-30.

[14] Aloh HE, Onwujekwe OE, Aloh OG, Nweke CJ. Is bed turnover rate a good metric for hospital scale efficiency? A measure of resource utilization rate for hospitals in Southeast Nigeria. Cost Eff Resour Alloc. 2020 Jul 1;18:21.