Unilag Journal of Mathematics and Applications,
Volume 4, Issue 1 (2024), Pages 1–7.
ISSN: 2805 3966. URL: http://lagjma.edu.ng

APPLICATION OF GEE IN THE FIT OF ORDINAL MARGINAL REGRESSION MODEL FOR TREATMENT RESPONSE OF HYPERTENSION

CHIJIOKE JOEL NWEKE*, GEORGE CHINANU MBAEYI, KELECHI CHARITY OJIDE, CHIJIOKE OGODO OGEH, AND AGHOR UDOCHUKWU

Abstract. This study adopts the ordinal logistic marginal model to fit the response to treatment of hypertensive patients with a view to ascertain any importance in incorporating panels when studying such clinical data. The Generalized Estimating Equation (GEE) provides an appropriate approach for estimating the model parameters. Two versions of ordinal regression models in conjunction with two forms of covariance estimators under three working correlation structures are considered. Results obtained revealed the importance of panels when studying the risk factors and response to treatment of hypertensive patients. With the Quasi-likelihood Information Criterion obtained for ordinal regression and other fitness criteria, the exchangeable working correlation is recommended as most adequate for the given data set.

1. Introduction

Hypertension also known as high blood pressure is a global burden and one of the most common health issue that can lead to myocardial infarction, stroke and renal failure when not detected and treated appropriately [[1], [2]]. The prevalence of hypertension is high in Africa (27%) and Nigeria in particular with a rise from 8.5% to 32.5% of age-adjusted prevalence from 1995 to 2020 [[3]-[5]]. Several studies on the associated risk factors and management of hypertension has been made [[6]-[9]], however there is paucity of studies on the modeling of the treatment response. This gap is what the study seeks to fill by adopting marginal regression model. The data set on response to treatment of High blood pressure patients in Federal Medical Centre (FMC), Owerri motivated this study. The data set contains panel data on treatment response, age, sex, genotype,

2010 *Mathematics Subject Classification.* Primary: 62J12. Secondary: 62J120.
Key words and phrases. Hypertension; Generalized Estimating Equation; Marginal Model; Working Correlation; Treatment Response; Quasi-likelihood Information Criterion
.

Submitted: August 1, 2024. Revised: September 23, 2024. Accepted: October 14, 2024.
* Correspondence.

blood group, blood sugar level as well as blood pressure of High Blood Pressure patients after about 24 hours from the time the patient is placed on certain drug combination (Lisinopril, Amlodiprine, Spirinolactone, Cardyra, Labitalol, Thiazid Diuret, Prazosi and Minozi). The drug combinations formed the panels. Since the outcome (response) variable is categorical, it is expected that the Generalized Linear model be adopted. Specifically, the proportional odds model for ordinal logistic regression described by [10] which provides a useful extension of the binary logistic model to situations where the response variable takes on ordered categorical values will best describe the effect of the covariates on the response. However, the model described above does not put the panel setting of a data set into consideration. When this is ignored, this could lead to bias and inefficient estimate. To remedy this, we adopt a regression model for panel data analysis.

[11] and [12] pointed out three broad, but quite distinct, classes of regression models for panel data. They are: (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 [[11], [13], [14]]. For a marginal model, using the Generalized Estimating Equation (GEE) approach to model estimation is considered more appropriate than Maximum Likelihood Method for various reasons [[11], [15]]. This study therefore seeks to fit proportional odds marginal models to the treatment response of high blood pressure patients who was administered different drug combinations.

2. Materials and Methods

Let y_{ij} be the i^{th} panel (drug combination) of the j^{th} patient response variable and $\mathbf{x}_{ij}$ be the corresponding vector of covariates. If y_{ij} is from an exponential family of distribution, then

$$f(y_{ij};\theta,\phi) = exp\left\{\frac{y_{ij}\theta_{ij} - b(\theta_{ij})}{\alpha(\phi)} - c(y_{ij},\phi)\right\} \tag{2.1}$$

where θ is the location parameter of the distribution, $\alpha(\phi)$ the scale or dispersion parameter, and $c(y_{ij},\phi)$ the normalizing term.

$$E(y_{ij}) = b^{'}(\theta_{ij}) = \mu_{ij} \tag{2.2}$$

$$Var(y_{ij}) = b^{''}(\theta_{ij}) \tag{2.3}$$

$E(y_{ij})$ and $Var(y_{ij})$ are respectively the mean and variance of y_{ij}. [14] gave the Generalized Estimating Equation (GEE) for Marginal Model as:

$$\Psi(\beta) = \left[\left\{\sum_{i=1}^{n}\sum_{j=1}^{n_i}\left(\frac{y_{ij}-\mu_{ij}}{\alpha(\phi)V(\mu_{ij})}\right)\left(\frac{\partial\mu}{\partial\eta}\right)_{ij} x_{kij}\right\}_{k=1,2,\ldots,p}\right]_{p\times 1} = [\mathbf{0}]_{p\times 1} \tag{2.4}$$

where y_{ij} is the response variable of the j^{th} patient in the i^{th} panel, μ_{ij} is the mean of the response of the j^{th} patient in the i^{th} panel, $x_{kij} = k^{th}$ covariate of the j^{th} patient in panel i, η_{ij} is a link function tha relates the parameter μ_{ij} to covariates, $V(\mu_{ij})$ is the function which is a function of μ_{ij}, and $\alpha(\phi)$ is a scale parameter. Putting 2.4 in a matrix term of the panels ($j = 1, 2, \ldots, n_i$), we have

$$\Psi(\beta) = \left[\left\{\sum_{i=1}^{n} \mathbf{X}_{ki}^T D\left(\frac{\partial \mu}{\partial \eta}\right)[V(\boldsymbol{\mu}_i)]^{-1}\left(\frac{\mathbf{Y}_i - \boldsymbol{\mu}_i}{\alpha(\phi)}\right)\right\}_{k=1,2,\ldots,p}\right]_{p\times 1} = [\mathbf{0}]_{p\times 1} \quad (2.5)$$

where $\mathbf{X}_i = (X_{i1}, X_{i2}, \ldots, X_{in_i})^T$ is the matrix of covariates for panel i, $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{in_i})^T$ is a matrix of the response variables for panel i and $D()$ a diagonal matrix.
The variance function $V(\boldsymbol{\mu})$ is a diagonal matrix which can be decompose as:

$$V(\boldsymbol{\mu_i}) = \left[D(V(\mu_{it}))^{\frac{1}{2}} R(\boldsymbol{\alpha})_{n_i\times n_j} D(V(\mu_{it}))^{\frac{1}{2}}\right]_{n_i\times n_j} \quad (2.6)$$

Where $D(V(\mu_{it}))$ is diagonal matrix with $V(Y_{it})$ along the diagonal and $R(\boldsymbol{\alpha}) = Corr(\boldsymbol{Y_i})$ is the correlation matrix as a function of $\boldsymbol{\alpha}$. Available working correlation matrix for panel data are independent, exchangeable, autoregressive, unstructured and the m-dependent correlation matrix. However, when the repeated nature of data is not with respect to time, it is recommended that the exchangeable, m-dependent and unstructured working correlation be used [14].

2.1. **Building a Proportional Odds Marginal Regression Model for Ordinal Response.** Let y_{ij} denote an ordinal response variable with c-levels $(1, \ldots, c)$ and X_{ij} be vector of p-covariates. The proportional odds model is given by:

$$logit(y_{ij} \leq q|\boldsymbol{X}_{ij}) = In\left[\frac{P(y_{ij} \leq q|\boldsymbol{X_{ij}})}{P(y_{ij} > q|\boldsymbol{X_{ij}})}\right] = \alpha_q + \boldsymbol{\beta}'\boldsymbol{X}_{ij} \quad (2.7)$$

Where $\boldsymbol{\beta}$ is the vector of regression coefficients which are constant across the logits, $\alpha_q (q = 1, 2, \ldots, c)$ is the intercept which changes across the logits such that $\alpha_1 < \alpha_2, <, \ldots, \alpha_{c-1}$ and ln is logarithm to base e.
The proportional odds model in 2.7 compares the probability of an equal or smaller response with the probability of larger response, both conditioned on the covariates. The regression coefficients in 2.7 are estimated using the GEE approach in 2.4 and 2.6.

2.2. **Choice of Working Correlation Matrix.** To ascertain the appropriate working correlation matrix $R(\boldsymbol{\alpha})$ in 2.6 to account for the correlation of the repeated observation and the relevant covariates, we adopt a goodness-of-fit model test. [14] recommended the use of Quasi-likelihood information criterion (QIC) which is an analogous to Akaike Information Criterion (AIC) for likelihood based model. The information criterion is given by:

$$QIC = -2\wp|g^{-1}(\boldsymbol{X}\beta_R)| + 2trace|A^{-1}(\beta_R)V_{MS,R}| \quad (2.8)$$

Alternatively, we can use

$$QIC_{HH} = -2\wp|g^{-1}(\boldsymbol{X}\beta_R)| + 2trace|A^{-1}(\beta_I)V_{MS,R}| \quad (2.9)$$

Where $\wp|g^{-1}(\boldsymbol{X}\beta_R)|$ is the value of the quasi-likelihood computed using the coefficients from the model with hypothesized correlation structure $R(\boldsymbol{\alpha})$. A is the variance matrix obtained by fitting an independence model, $V_{MS,R}$ is the modified sandwich estimate of variance from the model with hypothesized correlation structure $R(\boldsymbol{\alpha})$. 2.8 uses the variance matrix of the independence model evaluated at the independence estimates β_I while 2.9 uses variance matrix estimated using regression coefficients for the specified working correlation matrix β_R. Details of different existing working correlation can be seen in [14] and [16].

3. Discussion

The result of application of GEE in the fit of ordinal marginal regression treatment response of hypertension on some risk factors (sex, age and blood pressure at point of admission) are as shown in Table 1 and Table 2 for Robust and Model based covariance estimator respectively. The values are the parameter estimate of the regression model as well as their corresponding standard error and p-value for test of significance. The working correlations used are the independent, exchangeable and unstructured correlation. For Logistic robust estimator model, it could be seen that none of the risk factors is significant ($p > 0.05$) for independent and unstructured model while the age and blood pressure are significant ($p < 0.05$). The same scenario of significance of the estimate was found in the other models (Logistic model based and probit for both robust and model based estimator). On the model adequacy measure using QIC, the unstructured working correlation model has the minimum QIC followed by the Exchangeable working correlation. However, the unstructured working correlation model has none of its estimate significant, hence could not establish any relationship between the risk factors and hypertension. Therefore, the exchangeable working correlation is considered most adequate. This implies that panel using the type of drug administered is of great importance in investigating the effect of risk factors on treatment response of hypertensive patients.

Note: In Table 1 and Table 2, () indicate Standard error, [] contain P-values, Indpt. means Independent, Exch. means Exchangeable, Unstr. means Unstructured

4. Conclusion

The role of effect of drug combination in modeling response to treatment of hypertension has been examined in this study. The drug combination serve as panel, hence panel data regression model was adopted for the modeling. The result revealed that age and blood pressure before administration of drug are significant risk factors in treatment response. An exchangeable working correlation for logistic marginal model is recommended for modeling of this nature. Using QIC, the model based estimator outperformed the robust estimator for logic model and vice versa for probit model. However, adopting model based estimator should be done with caution, because its standard error is usually biased.

TABLE 1. Marginal ordinal regression of treatment response of hypertension on some risk factors (Robust Estimator)

	Logit Model			Probit Model		
Coefficients	Indpt.	Exch.	Unstr.	Indpt.	Exch.	Unstr.
Threshold1	-11.284 (5.583) [0.054]	-12.271 (4.820) [0.008]	-4.024 (2.837) [0.156]	-6.967 (3.489) [0.046]	-7.736 (2.879) [0.007]	-2.728 (1.761) [0.121]
Threshold2	-10.143 (5.748) [0.078]	-11.230 (4.793) [0.015]	-2.847 (2.823) [0.313]	-6.270 (3.434) [0.068]	-7.099 (2.851) [0.013]	-1.999 (1.751) [0.254]
Sex	-0.133 (0.364) [0.715]	-0.187 (0.490) [0.648]	-0.337 (0.267) [0.207]	-0.122 (0.225) [0.588]	-0.132 (0.253) [0.601]	-0.219 (0.173) [0.206]
Age	-0.028 (0.017) [0.097]	-0.029 (0.013) [0.044]	-0.003 (0.009) [0.761]	-0.017 (0.010) [0.086]	-0.018 (0.009) [0.040]	-0.001 (0.005) [0.831]
Age	-0.096 (0.062) [0.119]	-0.108 (0.055) [0.046]	-0.030 (0.033) [0.357]	-0.060 (0.037) [0.109]	-0.068 (0.033) [0.041]	-0.022 (0.020) [0.286]
QIC	-10.1180	-10.7765	-26.1461	-10.0938	-10.1581	-25.9305

TABLE 2. Marginal ordinal regression of treatment response of hypertension on some risk factors (Model based Estimator)

	Logit Model			Probit Model		
Coefficients	Indpt.	Exch.	Unstr.	Indpt.	Exch.	Unstr.
Threshold1	-11.284 (4.889) [0.021]	-12.271 (4.649) [0.001]	-4.024 (3.814) [0.291]	-6.967 (2.929) [0.017]	-7.736 (2.943) [0.009]	-2.728 (2.347) [0.245]
Threshold2	-10.143 (4.850) [0.037]	-11.230 (4.605) [0.019]	-2.847 (3.791) [0.453]	-6.270 (2.914) [0.031]	-7.099 (2.929) [0.015]	-1.999 (2.335) [0.392]
Sex	-0.133 (0.469) [0.777]	-0.187 (0.410) [0.702]	-0.337 (0.474) [0.477]	-0.122 (0.286) [0.671]	-0.132 (0.301) [0.660]	-0.219 (0.289) [0.449]
Age	-0.028 (0.016) [0.082]	-0.029 (0.014) [0.027]	-0.003 (0.015) [0.854]	-0.017 (0.010) [0.083]	-0.018 (0.008) [0.025]	-0.001 (0.009) [0.900]
Age	-0.096 (0.053) [0.071]	-0.108 (0.054) [0.048]	-0.030 (0.045) [0.501]	-0.060 (0.032) [0.063]	-0.068 (0.033) [0.040]	-0.022 (0.028) [0.434]
QIC	-10.1180	-75.2752	-193.5627	-10.0938	-9.7520	-24.2191

Acknowledgment. The authors acknowledge the constructive inputs of two anonymous reviewers as well as the team that facilitated the data collection process.

Authors Contributions. CJN conceptualized the study and wrote the first draft of the manuscript. COO and AU reviewed the manuscript. GCM and KCO analyzed the data. All authors read and approved the final manuscript.

Authors' Conflicts of interest. The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding Statement. This work did not receive any funding.

References

[1] American Family Physician. *JNC 8 Guidelines for the Management of Hypertension in Adult.* 90 (2014), 503–504.
[2] World Health Organization. *Hypertension.* https://www.who.int/newsroom/fact-sheets/detail/hypertension. Accessed 16 Jul 2021.
[3] W.K. Bosu, S.T. Reily, J.M.K. Aheto, E. Zuchelli. *Hypertension in older adults in Africa: a systematic review and meta-analysis.* PLoS One. (2019), 14:e0214934.
[4] F. Ataklte, S. Erqou, S. Kaptoge, B. Taye, J.B. Echouffo-Tcheugui, A.P. Kengne. *Burden of undiagnosed hypertension in sub-saharan Africa: a systematic review and meta-analysis.* Hypertension. 65 (2015), 291–298.
[5] D. Adeloye, E.O. Owolabi, D.B. Ojji, A. Auta, M.T. Dewan, T.O. Olanrewaju. *Prevalence, awareness, treatment, and control of hypertension in Nigeria in 1995 and 2020: a systematic analysis of current evidence.* J. Clin. Hypertens. 23 (2021), 963–977.
[6] D. Shimbo. *Dietary and lifestyle factors in hypertension.* J. Hum. Hypertens. 30 (2016), 571–572.
[7] E.M. Amoah, D.E. Okai, A. Manu, A. Laar, J. Akamah, K. Torpey. *The role of lifestyle factors in controlling blood pressure among hypertensive patients in two health facilities in urban Ghana: a cross-sectional study.* Int. J. Hypertens. 3 (2020), 1–8.
[8] C. Abbafati, K. M. Abbas, M. Abbasi-Kangevari, F. Abd-Allah, A. Abdelalim, M. Abdollahi, et al. *Global burden of 87 risk factors in 204 countries and territories,19902019: a systematic analysis for the global burden of disease study 2019.* Lancet. 396 (2020), 1223–1249.
[9] A.S. Adeke, B.S. Chori, D. Neupane, J.E. Sharman. *Socio-demographic and lifestyle factors associated with hypertension in Nigeria: results from a country-wide survey.* J. Hum. Hypertens. 38 (2022), 365–379.
[10] P. McCullagh. *Regression models for ordinal data.* J. R. Stat. Soc., B: Stat. Methodol. 42 (1980), 109–142.
[11] G. Fitzmaurice, M. Davidian, G. Verbeke, G. Molenberghs. *Handbook of Modern Statistical Methods: Longitudinal Data Analysis.* H. Chapman & Hall/CRC. Boca Raton (2009).
[12] C.J. Nweke, K.C. Ojide, G.C. Mbaeyi, F.I. Ugwuowo. *Marginal Regression Models for Mixed Bivariate Responses.* Sci. Technol. Asia. 28 (2023), 20–25.
[13] L. Xian. *Methods and Applications of Longitudinal Data Analysis.* Elsevier Inc. London (206).
[14] GJ.W. Hardin, J.M. Hilbe. *Generalized Estimating Equations (Second edition).* H. Chapman & Hall/CRC. Boca Raton (2013).
[15] S.R. Lipsitz, K. Kim, L. Zhao. *Analysis of repeated categorical data using generalized estimating equations.* Stat. Med. 13 (1994), 1149–1163.
[16] M. Wang. *Generalized Estimating Equation in Longitudinal Data Analysis: A review and Recent Developments.* Advances in Statistics. (2014), 1–11.

Chijioke Nweke* ,
Department of Mathematics and Statistics, Alex Ekwueme Federal University Ndufu-Alike, Abakaliki, EbonyiState, Nigeria.
E-mail address: chijioke.nweke@funai.edu.ng

George Mbaeyi
Department of Mathematics and Statistics, Alex Ekwueme Federal University Ndufu-Alike, Abakaliki, EbonyiState, Nigeria.
E-mail address: george.chinanu@funai.edu.ng

Kelechi Ojide
Department of Mathematics and Statistics, Alex Ekwueme Federal University Ndufu-Alike, Abakaliki, EbonyiState, Nigeria.
E-mail address: kelechi.ojide@funai.edu.ng

Chijioke Ogeh
Department of Paediatrics, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, EbonyiState, Nigeria.
E-mail address: godstime@yahoo.com

Aghor Udochukwu
Department of Family Medicine, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, EbonyiState, Nigeria.
E-mail address: uaghor@yahoo.co.uk