SPATIAL AUTOCORRELATION ESTIMATION OF MALARIA MORBIDITY AMONG POLYTECHNIC STUDENTS IN NIGERIA

Adeboye N. Ol.[1]* and Ajibode, I. A[2].
Department of Mathematics & Statistics, Federal Polytechnic, Ilaro P.M.B 50, Ilaro, Ogun State, Nigeria
nureni.adeboye@federalpolyilaro.edu.ng; +2348033348141

ABSTRACT

While malaria effect on people of all ages is quite immense, the most serious impact of malaria is on students of tertiary institutions who lives and operate mostly in mosquito infested environment. The frequent rate of malaria incidence for this category has been largely responsible for poor academic performance, whereby majority struggle through schooling amidst unstable healthy condition. This research thus attempt to estimate spatial autocorrelation of malaria morbidity among the students of Yaba College of Technology and Federal Polytechnic Ilaro. As a result, Ordinary Least Square (OLS), Spatial Lag Regression (SLR) and Spatial Error Regression Models (SERM) were employed for modeling the secondary data collected from the institutions medical centers. Moran I test statistic was introduced to check for the presence of spatial autocorrelation among the adopted explanatory variables of headache, pain, fever and cold for the two locations. OLS was discarded due to its low significant level while SLR model was adopted based on its Akaike Information Criterion (AIC) and Log Likelihood test as the best measure of spatial autocorrelation existence in malaria morbidity between the two selected institutions, thereby given credence to the fact that malaria incidence in Nigerian tertiary institutions have a universal effect on students.

Keywords: Malaria Morbidity, Ordinary Least Square, Spatial Error Regression, Spatial Lag Regression, Youth Employment.

INTRODUCTION

Malaria is a mosquito-borne infectious disease which contributes substantially to the poor health situation in Africa. According to Ayeni (2011), Afolabi *et al. (*2012), WHO (2012) and Cheesbrough (2016), it is on record that sub-Saharan Africa accounts for 90% of the world's 300 to 500 million cases of malaria and 1.5 to 2.7 million deaths annually. About 90% of these deaths in Africa are of young children, suggesting some serious demographic consequences for the continent. Malaria is a great burden on the health system in Africa, as it is responsible for 20 to 40 % of outpatient visits and 10 to 15% of hospital admissions, according to the World Health Organization (WHO, 1999). In Sub-Saharan Africa, 10.8% of all disability-adjusted life years (DALYs) were lost to malaria in 1990. According to the World Bank report, malaria accounted for an estimated 35 million DALYs lost in Africa in 1990 due to ill health and premature death. The estimate was 39 million DALYs in 1998 and 36 million DALYs in 1999 (WHO, 1999; 2010; 2012). Furthermore, while malaria contributed 2.05% to total global deaths in 2000, it was responsible for 9.0% of all deaths in Africa (WHO, 2012). With regards to school children, malaria, is a major cause of absenteeism in endemic countries. It is estimated that about 2% of children who recover from cerebral malaria suffer brain damage including epilepsy (WHO/UNICEF, 2005). Hence, among young children, frequent episodes of severe malaria may harm their learning abilities and educational attainment (Clifford *et al.,* 2016). This is a threat to human capital formation, which constitutes a key factor in youth employment.Malaria is therefore a massive problem that affects all segments of society.

Poverty contributes to the prevalence of malaria in Africa. According to Pattanayak*et al* (2014) many of the world's poorest people live in areas of high rate of malaria. These people do not have access to effective health care due to financial constraint. Brown *et al.* (2002) called malaria as a disease of poverty. The economic status of a vulnerable country plays another role in determining the equipped and control measures in case of epidemics (Gerristen*et al.* 2008). A survey in Zambia found a substantially higher prevalence of malaria infections among the poorest population group.

While its effect on people of all ages is quite immense, the most serious impact of malaria is on students of tertiary institutions who lives and operate mostly in mosquito infested environment due to financial constraints. The frequent rate of malaria incidence for this category has been largely responsible for poor academic

performance, whereby majority struggle through schooling amidst unstable healthy condition. Looking at the immediate environment of the densely populated urban and rural dwellers of Yaba, Lagos and Yewa South respectively, it is clear that these communities exhibit breeding places for mosquitoes which are carriers of malaria parasite. The bushy environment, unhygienic condition in which people live and unavailability of functional drainage system are major setback to malaria control in these areas. Since the areas cannot be exonerated from malaria epidemic, it is however necessary to develop a predictive model that will help in the measure of spatial autocorrelation existence in malaria morbidity between the two selected institutions, thereby given credence to the fact that malaria incidence in Nigerian tertiary institutions have a universal effect on students, thus predicting the prevalence of this disease in future and assist the Government in formulating policies that will help in curbing the deadly disease among academic communities.
Perhaps the impact of malaria has not been demonstrated in quantitative terms that might convince policy makers to devote the needed attention and resources to combating this dreadful disease. Moreover, there will be need to establish scientifically the degree of relationship that exist between malaria infection and its incidence (the symptoms causing this disease) within the considered location. This study is an attempt to fill this gap with the estimation of spatial autocorrelation between the two selected locations. According to Anselin (2012) and Ord (2013), spatial autocorrelation represents the correlations between the values of a random variable at a location and the values of the same variable at "neighboring" locations. Thus, the adopted variables in this research are headache, pain, fever and cold for the two locations.

MATERIALS AND METHODS

The data for this study was obtained mainly from records kept by Federal Polytechnic Ilaro School's Clinic, Ogun State as well as from Yaba College of Technology Military School's Clinic, Lagos State between January 2002 and December 2017. It comprises of monthly malaria cases as well as the symptoms that causes malaria.
The analytical technique employed in this research is spatial data analysis. Spatial data is characterized by having "location" or "spatial" effects, where there are spatial heterogeneity between and spatial homogeneity within neighboring clusters; thus "spatial dependence" is exhibited among these clusters. When these characteristics are ignored, Ordinary Least Square (OLS) parameter estimates are bias, inconsistent and R^2 value does not provide a goodness of fit since the assumption of independent error terms is violated. This necessitates the adoption of Spatial Lag Model (SLM) and Spatial Error Model (SEM).

Spatial Lag Model

A spatial lag model is expressed as

$$y = \rho w_y + X\beta + \varepsilon \tag{1}$$

where ρ is a spatial autoregressive coefficient, ε is a vector of error terms, β is the regression coefficient and the spatial lag term W_y is correlated with the disturbances as given in the reduced form:

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1}\varepsilon \tag{2}$$

$usingtheargument$: $(I - \lambda W)y = Z\beta + u$ and $y = (I - \lambda W)^{-1} Z\beta + (I - \lambda W)^{-1} u$

When $\lambda \neq 0$ and $\rho = 0$, the model becomes:

$$y = \lambda W\gamma + Z\beta + u, \quad \lambda < 1 \tag{3}$$

With $u \sim N.\mathrm{i.i.d}(0, \sigma_u^2, I_n)$;thus, equation (3) is the Spatial Lag Model without autocorrelation.

The Spatial Error Model

Setting $\lambda = 0$ and $\rho \neq 0$ with the regressors Z and the weights W non-stochastic in the spatial lag model, the Spatial Error Model (SEM) becomes

$$y = Z\beta + u \tag{4}$$

$$u = \rho W u + \varepsilon |\rho| < 1 \qquad (5)$$

If $\varepsilon | X \sim N.\mathrm{i.i.d}(0, \sigma_\varepsilon^2, \mathrm{I_n})$, then we have that $u = (I - \rho W)^{-1} \varepsilon and E(u) = 0$

$$E(uu^T) = \sigma_\varepsilon^2 (I - \rho W)^{-1} (I - \rho W^T)^{-1} = \sigma_\varepsilon^2 \Omega \qquad (6)$$

Equation (6) considers both heteroscedastic and autocorrerlated error terms. Notice that from equation (5) we have:

$$(I - \rho W)u = \varepsilon$$

Model (4) and (5) can thus be written as:

$$(I - \rho W)y = (I - \rho W)Z\beta + (I - \rho W)u$$

$$y = \rho W y + Z\beta - WZ\rho\beta + \varepsilon$$

$$y = \rho W y + Z\beta - WZ\gamma + \varepsilon \qquad (7)$$

Equation (7) is over-parameterized due to restriction $\gamma = \rho\beta$ and the term Wy is correlated with the error term, thus producing endogeneity.

From equation (7), we have

$$(I - \rho W)y = Z\beta - WZ\gamma + \varepsilon$$

And so

$$y = (I - \rho W)^{-1}(Z\beta - WZ\gamma) + (I - \rho W)^{-1}\varepsilon \qquad (8)$$

Thus, equation (8) is the spatial error model.

RESULTS AND DISCUSSION

Three different analyses were performed. First, Ordinary Least Squares (OLS) regression was carried out as a reference model after which estimation by means of maximum likelihood of the specified spatial regression lag model and a spatial error regression model were carried out.

The results are presented as follows:

Table 1: Ordinary Least Square Regression

Coefficients	Estimate	Std. Error	Pr(>|t|)
Intercept	-0.41947	0.10020	0.000354
Headache	0.30926	0.13065	0.026719
Pain	1.34775	0.09823	1.46e-12
Fever	0.64214	0.11865	1.69e-05

Cold	0.31067	0.12750	0.022983

Table 2: Moran I Test Statistic

Sample Estimates:

Observed Moran I	Expectation	Variance	Moran I statistic standard deviate	p-value
-0.03766159	-0.08355617	0.05155787	0.20212	0.0419

Based on OLS output of table 1, the considered explanatory variables (i.e. headache, pain, fever and cold) were all significant at 5% level of significance. The F-test equally provides the overall goodness of fit for the regression model at 5% level of significance. However, the research main focus supersedes model fitting, hence the conduct of Moran I test statistic for the detection of spatial autocorrelation among the residuals as presented in table 2, which show evidence of a positive and highly significant spatial correlation in the regression residuals that motivates further analysis. The p-value for the tested null hypothesis that "there is no spatial autocorrelation" is 0.0419 which is less than the p-value. Therefore the null hypothesis is rejected and this necessitates the estimation of spatial lag and spatial error models presented in tables 3 and 4, which are capable of dealing with spatial dependencies of which the best is selected based on some adopted criteria.

Table 3: Spatial Lag Regression Model

Coefficients	Estimate	Std. Error	Pr(>|t|)
Intercept	-0.356789	0.108904	0.001052
Headache	0.322950	0.117135	0.005832
Pain	1.328451	0.089557	< 2.2e-16
Fever	0.626894	0.106742	4.281e-09
Cold	0.355770	0.122099	0.003571

Table 4: Spatial Error Regression Model

Coefficients	Estimate	Std. Error	Pr(>|t|)
Intercept	-0.414869	0.089543	3.601e-06
Headache	0.301869	0.118261	0.010694
Pain	1.349869	0.316829	< 2.2e-16
Fever	0.642779	0.106545	1.610e-09
Cold	0.316829	0.114088	0.005485

Table 5: Model Diagnostics

	Spatial Error		Spatial Lag	
Test	**Estimate**	**p-value**	**Estimate**	**p-value**
LR test value	0.034334	0.8530	0.98935	0.3199

Asymptotic standard error	0.15057	0.82967	0.00023376	0.31565
Wald statistic	0.046279	0.82967	1.0069	0.03156
Log likelihood	16.752273		27.229783	
AIC	0.49545		-0.45957	

The models specified for the spatial lag and spatial error regression from tables 3 and 4 respectively are given as

$$\hat{y} = -0.356789 + 0.322950\mathbf{headache} + 1.328451\mathbf{pain} + 0.626894\mathbf{fever} + 0.355770\mathbf{cold} \quad (9)$$

$$\hat{y} = -0.414869 + 0.301869\mathbf{headache} + 1.349869\mathbf{pain} + 0.642779\mathbf{fever} + 0.316829\mathbf{cold} \quad (10)$$

The two models are uniformly related and thus present situations where malaria morbidity will be negatively inclined (i.e reduced) when the four symptoms are kept constant while their significant presence posed increasing effects of malaria morbidity on any treated patients. The result of Wald test shows the spatial lag model goodness of fit and its overall significant while that of Spatial error model is not significant based on its P-value which is greater than 0.05. However, while emphasizing the important feature of a spatial dependence of malaria morbidity, it cannot be said to have been completely removed. Meaning the problem of residual spatial autocorrelation is still present in the models based on LR test for residuals which shows that there is still some positive and significant residual correlation.

Finally, the results of information criterion presented in table 5 implied that that spatial lag regression model (SLM) is better than spatial error regression model (SEM) since SLM has highest Log likelihood and lowest Akaike Information Criterion compared to that of SEM.

CONCLUSIONS

The findings of this research give credence to the high significant impacts of headache, pain, fever and cold on malaria morbidity and their uniformity of effects on the spatial dependencies of the two polytechnics is quite revealing. It is a clear indication of spatial autocorrelation existence in malaria morbidity between the two selected institutions, thereby given credence to the fact that malaria incidence in Nigerian tertiary institutions have a universal effect on students. Thus, an unhealthy students populace that are constantly experiencing the attack of mosquito bites which is a major known transmitter of malaria plasmodium cannot be considered a healthy population with sound mind for excellent academic performance. Finally, the research has shown that spatial lag regression model (SLM) is better than spatial error regression model (SEM) in the modeling of spatial autocorrelation effects of malaria morbidity among suspected ecosystem. While malaria effect on people of all ages is quite immense, the most serious impact of malaria is on students of tertiary institutions who lives and operate mostly in mosquito infested environment due to financial constraints. The frequent rate of malaria incidence for this category has been largely responsible for poor academic performance, whereby majority struggle through schooling amidst unstable healthy condition. Looking at the immediate environment of the densely populated urban and rural dwellers of Yaba, Lagos and Yewa South respectively, the results of this research have shown clearly based on the evaluated significance level of all the adopted malaria symptoms that these communities exhibit breeding places for mosquitoes which are carriers of malaria parasite.

REFERENCE

Afolabi, I. S., Akpokene, O. N., Fashola, D. V., Famakin, T. C (2012). Comparative Evaluation of the Nutritional benefits of some underutilized Plants Leaves. *Journal of Natural Product and Plant Resources, 2(2), 261-266*

Anselin, L., (2012). Local indicators of spatial association –LSA. *Geographical Analysis, 27 (2),93-115.*

Ayeni, A. O**.**, (2011). Malaria Morbidity in Akure, Southwest Nigeria: A Temporal observation in Climate Change scenario. *Trends Applied Science. (Res. 6): 488-494.*

Brown, V., Issak, M.A., Rossi, M., Barboza, P. and Paugam, A. (2002). "Epidemic of malaria in North Eastern Kenya", *Lancet 352, 1356–1357.*

Castilla R.E., and Sawyer D.O.(1993). Malaria rates and fate: A socioeconomic study of malaria in Brazil. *Social Science & Medicine. 37(9):1137-1145.*

Cheesbrough, M**.** (2005). *District Laboratory Practice in Tropical Countries,* Cambridge University Press, Cambridge University Press, Cambridge UK, 2nd Edition.

Clifford, A., Lang, L., Chen, R., Anstey, K. J., Seaton, A. (2016). Exposure to Air Pollution and Cognitive Functioning across the life course - a systematic literature review. *Environmental Research, 147,383-398*

Gerristsen, A.; Kruger P; Van der Leo, M. and Grobusch, M. (2008). 'Malaria incidence in Limpopo Province, South Africa, 1998 – 2007', *Malaria Journal,* 7(1) http://www.malariajournal.

Martin, H. (2013). Genomic and Epigenomic Landscapes of Adult de novo acute Myeloid Luekemia, New England Journal of Medicine, 368(22), 2059-2074

Mills. A, Lubell. Y. and Hanson. K, (2008*). "Malaria eradication: the economic, financial and Institutional Challenge",* Malaria journal 7, (suppl I), sII. URL: http:// www. Malaria journal. Com /7/ s1/s11.

Pattanayak V., Guilinger, J. P., Liu, D. R. (2014). Determining the Specificities of TALENs, Cas9, and other Genome Editing Enzymes. Methods in enzymology, 546, 47-78. Doi: 10.1016/B978-0-12-801185-0.00003-9.

Nchinda, T.C. (2005). Malaria: A Reemerging Disease in Africa, World Health Organization, Geneva.

Ord, J.K. (2013) 'Estimation Methods for Models of Spatial Interaction', *Journal of the American Statistical Association, 70, 120-6.*

Upton, G. and B. Fingleton (2009) *Spatial data analysis by example.* Volume 1. Wiley, Chichester.WHO (2014). Epidemiology of Severe Falciparum Malaria, *Tropical Medicine & International Health, 19, 7-131.*

World Health Organization (2012).World Health Report, 2012: *making a difference. http:/www.who.int/whr/2012/*

World Health Organization (2010). *Basic Malaria Microscopy. Part 1 Learner's Guide*, WHO Geneva, Switzerland, 2nd Edition

WHO/UNICEF (2005). World Malaria Report. WHO, Geneva Switzerland. ISBN. 92 4 1593199. *https://www.who.int.malaria.atoz.*

WHO/UNICEF (2003). Africa Malaria Report. WHO/CDS/MAL/2003.1093.2003. Geneva.

World Health Organization (1999). World Health Report, 1999; *making a Difference*. *http:/www.who.int/whr/1999/*