Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.5, No.4, 2015
www.iiste.org

On The Statistical Analysis of Ante-Natal Care Use in Nigeria: Multilevel Logistic Regression Approach

A.A Akinrefon[1*], O.I Adeniyi[2], A.O Adejumo[3] O.S Balogun[4], E. Torsen[5]
[1,4,5] Modibbo Adama University of Technology, Yola, Adamawa State
[2,3]University of Ilorin, Ilorin, Kwara State
[*]e-mail: akinrefonadesupo@yahoo.com

Abstract
Ante-Natal Care usage among women has a great impact on the survival rate of both mother and the new born child. Inadequate usage influences the mortality rate of both woman and infant thus affecting both economic and health status of any population, Nigeria in particular. Attaining and even surpassing the minimum number of Ante Natal Care (ANC) visits as specified by World Health Organization, WHO helps to ensure healthy outcomes for women and newborns. Multilevel logistic regression models were applied to data involving women who make ANC visits who participated in the 2013 National Demographic Health Survey (NDHS) in Nigeria.

Results show that location, women education, religion and wealth index of participant were significant factors that influence the probability of women meeting and/or surpassing the required number of ANC visits required by WHO. We also found that husband's level of educational attainment and the frequency of information gathering by such women also influenced their chances of making more ANC visits.
Keywords: Ante-Natal Care, Multilevel, Logistic Regression, Intra-class correlation

Introduction:

Ante Natal Care (ANC) is the care that a woman receives during pregnancy. This helps to ensure healthy outcomes for women and newborns (WHO/UNICEF, 2003). ANC is a key entry point for a pregnant woman to receive a broad range of health promotion and preventive health services including nutritional support, prevention and treatment of anemia; prevention, detection and treatment of malaria, tuberculosis and sexually transmitted infections (STI's), HIV/AIDS (particularly prevention of HIV transmission from mother to child). ANC is an opportunity to promote the benefits of skilled attendance at birth and to encourage women to seek postpartum care for themselves and their newborn. ANC provides an ideal time to counsel women on the benefits of child spacing. www.accesstohealth.org.

ANC coverage in Africa is a success story since over two thirds of pregnant women (about 70%) have at least one ANC contact (WHO/UNICEF, 2003). However, to reach or achieve full life-saving potentials which ANC promises for women and babies, **at least four (4)** visits providing essential evidence-based interventions is required. Traditional ANC uses risk approach to classify which women are more likely to experience complications, and assumes that more visits mean better outcomes for mother and baby. However, many women with risk factors will not develop complications while women without risk factors may do so. Lilford and Chard, (1983); Vanneste et.al (2000); Yuster (1995).

The World Health Organization (WHO) recommends a minimum of four (4) ANC visits for women whose pregnancies are progressing normally. Maternal health care services utilization is important for early detection of mothers who are at risk of illness and mortality during pregnancy.

Reducing maternal mortality ratio in all countries by 2015 is one of the Millennium Development Goals (MDGs). As a result, Government Health Sector policies in Nigeria has been geared towards achieving this. Worldwide, an estimated half a million deaths occur in women due to complications of pregnancy and childbirth. Of these, 55,000 maternal deaths occur in Nigeria (Nigeria Health Review, 2006; National Planning Commission, NPC, 2001). Nigeria accounts for 10% of the global estimates of maternal deaths. Rural areas are the worst hit accounting for 828 deaths per 100,000 live births and 531 deaths per 100,000 in urban areas. Nigeria Health Review (2006) reports inadequate motherhood services such as ANC as a major cause of maternal death.

Several factors have been identified in literature to attract ANC utilization: Adewoye et al (2013) found knowledge, awareness or health knowledge, to play a key role in ANC utilization especially in Government owned health facilities. Nwosu et al (2012), found women educational level and household wealth status to contribute to ANC utilization. Okonofua et al (1992) in their study of risk factors that affect maternal mortality in Ile-Ife, Nigeria, identified young age of mothers and socio-economic status (wealth) as contributing factors.

The aim of this study is to check the effect of these factors on ANC utilization in Nigeria. We would also investigate the contribution of husband's educational level, regularity of information gathering by women and place of residence on propensity of women to Ante-Natal Care services in Nigeria. Furthermore, very few research have utilized multilevel logistic model in their research. Given the structure of demographic survey which are large scale surveys and often are hierarchical data structure based on multistage stratified cluster sampling, the appropriate approach to analyzing such survey data should be based on the nested sources of variability which come from different levels of the hierarchy. Traditional logistic regression would be inappropriate when the variance of the residuals errors is correlated between individual observations as a result of the nested structures which is often the case in multistage data.

THE MULTILEVEL MODEL

Multilevel modelling is used in the analysis of data that have hierarchical or clustered structure. Such data arise routinely in various fields as in educational research in which pupils are nested in schools; in family studies where children are nested within families; in medical research in which patients are nested in hospitals, and in biomedical research in which teeth are nested within different people's mouth in dental analysis (Joop and Cora, 2005).

The specific research design may also give rise to clustered data. Because of cost, time and efficiency consideration, stratified multistage samples are the norm for sociological and demographic surveys. For such samples, the clustering of the data is a nuisance which should be taken into consideration. This cluster sampling scheme often introduces multilevel dependency or correlation among the observations that can have implications for model parameter estimates. For multistage clustered samples, the dependence among observations often comes from several levels of the hierarchy. The appropriate approach to analyzing large-scale survey data which follows a hierarchical structures is based on nested sources of variability which come from different levels of the hierarchy. When the variance of the residual errors is correlated between individual observations as a result of these nested structures, traditional logistic regression is not appropriate (Khan and Shaw, 2011).

Multilevel models generally assumes that the grouping criterion is clear and that variables can be assigned unambiguously to their appropriate level (Joop and Cora, 2005). Multilevel models are set to design variables from different levels simultaneously, using a statistical model that includes the various dependencies.

MULTILEVEL REGRESSION MODEL

In the research literature, multilevel regression model has become known under a variety of names as; hierarchical linear model, random coefficient model, variance component model and mixed (linear) model. Often, it assumes hierarchical data, with one response variable measured at the lowest level and explanatory variables at all other existing levels.

The model is often viewed conceptually as a hierarchical system of regression equations. Hierarchical linear models allow for the simultaneous investigation of the relationship within a given hierarchical level, as well as the relationship across levels. Two models are developed in order to achieve this: one that reflects the relationship within lower level units, and another that models how the relationship within lower level units vary between units (thus correcting for the violation of aggregating or disaggregating data (Hofmann, 1997). This is applicable when lower lever units are nested within higher-level units.

Assume we have data on j groups with N_j individual in each of the groups. On the individual level, we have the response variable Y_{ij} and p explanatory variables $X_1,\ldots,X_p$ and q explanatory variables at the group level $Z_1,\ldots Z_q$.

With only one predictor variable at the individual level, we have the separate regression equation in each group as

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \quad (1)$$

The regression coefficient β carries a subscript j for the groups, which implies that the regression coefficients may vary across groups. The variation in the regression coefficients β_j is modelled by explanatory variables and random residual terms at the group level as

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{ij} \quad (3)$$

Substituting (2) and (3) in (1), we have

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij} + u_{ij} X_{ij} + u_{0j} + e_{ij} \quad (4)$$

Introducing the q explanatory variables at the individual level and the q explanatory variables at group level, the model (4) becomes

$$Y_{ij} = \gamma_{00} + \sum_p \gamma_{p0} X_{pij} + \sum_q \gamma_{0q} Z_{qj} + \sum_q \sum_p \gamma_{pq} Z_{qj} X_{pij} + \sum_p u_{pj} X_{pij} + u_{0j} + e_{ij} \qquad (5)$$

The γ are the usual regression coefficients while the u terms and the e are the residuals at the group and the individual level respectively. The regression coefficients are the fixed part of the model because they do not change over the groups while the residual error terms are the random or stochastic part of the model, $e_{ij} \sim N(0,\sigma^2)$, while the u_{0j} and u_{1j} are assumed to be independent of the individual level error term and to have a multivariate normal distributions with means of zero.

The proportion of the variance in the population explained by the grouping structure is estimated by the intra-class correlation ρ. The model with no explanatory variable at both levels (intercept only model) is used to estimate the ρ. The model is given as

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \qquad (6)$$

The intra-class correlation coefficient is estimated by the

$$\rho = \frac{\sigma_{u0}^{\ 2}}{\sigma_{u0}^{\ 2} + \sigma_e^{\ 2}} \qquad (7)$$

Where $\sigma_{u0}^{\ 2}$ is the variance of the second level residuals and $\sigma_e^{\ 2}$ is the variance of the individual level residuals.

Multilevel Logistic Regression

When the response variable is dichotomous or a proportion, the assumption of continuous scores and normality are not met, also the homoscedastic assumption of errors is violated.

Consider a two-level model for binary outcomes with an explanatory variable at each of the two levels. This model is equivalent to model 1 except for the outcome variable. Let π_{ij} be the observed proportion of individuals in category 1 of group j, $\pi_{ij} = P(y_{ij} = 1)$. Although y_{ij} has a binomial distribution, logit (P_{ij}) has a distribution that is approximately Normal. The intercept-only model is given by

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} \qquad (8)$$

The individual level error term e_{ij} is not in (8) because in binomial distribution, the variance of the observed proportion depends only on the population proportion; therefore, the individual level variance is determined by the predicted value for the P_{ij} and does not enter the model as a separate term (Joop and Cora, 2005).

Adding a predictor of the individual level, the model (8) becomes

$$\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \beta_{1j} x_{ij} \qquad (9)$$

The regression coefficient is assumed to vary across the group, which is modelled by the predictor of the group level.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \qquad (10)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{ij} \qquad (11)$$

Substituting (10) and (11) in (9), we have

$$\log\left(\frac{P(y_{ij}=1)}{P(y_{ij}=0)}\right)=\gamma_{00}+\gamma_{10}X_{ij}+\gamma_{01}Z_j+\gamma_{11}Z_jX_{ij}+u_{ij}X_{ij}+u_{0j} \quad (12)$$

The probability can be written as

$$\pi_{ij}=\frac{\exp\left(\gamma_{00}+\gamma_{01}z_j+\gamma_{10}x_{ij}+u_{0j}\right)}{1+\exp\left(\gamma_{00}+\gamma_{01}z_j+\gamma_{10}x_{ij}+u_{0j}\right)} \quad (13)$$

Intra-class Correlation

The intra-class Correlation ρ indicate the proportion of the variance explained by the grouping structure in the population. In the multilevel logit model, $\sigma_e^{\ 2}=\pi^2/3$ by assumption, so that the intra-class correlation is

$$\rho=\frac{\sigma_{u0}^{\ 2}}{\sigma_{u0}^{\ 2}+\frac{\pi^2}{3}} \quad (14)$$

This is the intra-class correlation for the latent response variable.

APPLICATION TO DATA ON NUMBER OF ANTENATAL VISITS

The 2013 Nigerian Demographic Health Survey (NDHS) data set is used for this study. This is based on multistage stratified cluster sampling, thus an approach based on nested sources of variability would be most appropriate in assessing the attitude of pregnant mothers/women to ANC use/visits in Nigeria. NDHS in nationally representative cutting across the 36 states and the Federal Capital Territory of the Federation. The response variable in this study is "number of ANC visits during pregnancy". For this study, it is classified into two non-overlapping categories such that those with counts of visits less than four and those with at least four visits to ANC from the two groups hence, the use of a multilevel logistic regression for modelling.

The number of antenatal visits made by a pregnant woman makes is grouped into two (2) non-overlapping categories of either the minimum of four (4) as specified by WHO or less than four (4) thus making the outcome variable dichotomous.

ANALYSIS AND RESULT:

The survey by design provides information at individual and state levels, the hierarchical structure of the data set in this study is as described below:

At individual level, the explanatory variables at the individual-level include: place of residence (rural and urban), education level of the woman, (and of her husband) (no education, primary, secondary and higher degree), wealth index (poor, middle and rich), religion (Christianity, Islam and Others), and frequency of listening or watching T.V., newspaper or radio.

At the State level, each individual belongs to one of the 37 states in the Country. The only explanatory variables at this level is the geopolitical zone (North-Central, North-East, North-West, South-East, South-West, and South-South).

The number of antenatal visit of woman i from state j is defined by a binary indicator

$$y_{ij}=\begin{cases}1 & y_{ij}\geq 4\\ 0 & y_{ij}<4\end{cases}$$

For this research work we shall consider the following three cases:

Case 1: Intercept Only Multilevel Logistic Model (Null Model)

We first fit a simple model with no predictors, that is, an intercept-only model that predicts the probability of making a minimum of four (4) ANC visits. The model is

$$\ln\left(\frac{y_{ij} \geq 4}{y_{ij} < 4}\right) = \beta_{0j} + u_{0j} \tag{15}$$

Case 2: Model with the individual-level factors

The model is

$$\log\left(\frac{pr(y_{ij} \geq 4)}{pr(y_{ij} < 4)}\right) = \beta_{0j} + \beta_1 Rural + \beta_2 Wife\_\Pr imary + \beta_3 Wife\_Secondry + \beta_4 Wife\_HigherEducation$$
$$+ \beta_5 Husband\_Secondary + \beta_6 Husband\_HigherEducation + \beta_7 Islam + \beta_8 Traditional + \beta_9 Poor$$
$$+ \beta_{10} Middle\_class + \beta_{11} NewsPaper >= 1/wk + \beta_{12} Radio >= 1/wk + \beta_{13} Television >= 1/wk + u_{0j} \tag{16}$$

Case 3: Model with inclusion of the state-level factors

The model is given as

$$\log\left(\frac{pr(y_{ij} \geq 4)}{pr(y_{ij} < 4)}\right) = \beta_{0j} + \beta_1 Rural + \beta_2 Wife\_\Pr imary + \beta_3 Wife\_Secondary + \beta_4 Wife\_HigherEducation$$
$$+ \beta_5 Husband\_Secondary + \beta_6 Husband\_HigherEducation + \beta_7 Islam + \beta_8 Traditional + \beta_9 Poor$$
$$+ \beta_{10} Middle\_class + \beta_{11} NewsPaper >= 1/wk + \beta_{12} Radio >= 1/wk + \beta_{13} Television >= 1/wk$$
$$+ \beta_{14} North\_East + \beta_{15} North\_West + \beta_{16} South\_East + \beta_{17} South\_South + \beta_{18} South\_West + u_{0j} \tag{17}$$

A stepwise procedure was adopted. The first of which was to fit a null model of overall probability of a woman making not less than four (4) antenatal visits excluding all explanatory variables.

The second step was the inclusion of the factors at the individual level while the factor at the state level was added at the third stage of the analysis. Urban dwellers, Christians, No Education, North Central and those whose make use of TV, radio and Newspapers less than once a week are considered as reference variables The results of the fixed effects (measures of association) are shown as coefficients odds ratios (ORs) and their corresponding p-value.

Table 1 in Appendix, presents the results for a sequence of three models: the intercept only model, a model with the individual level factors and a model with the state level factors.

For case 1, the intercept is estimated as 0.562 which refer to the underlying distribution established by the logistic link function. To determine the expected proportion, the inverse transformation for the logistic link function is given by

$$g(x) = \frac{e^x}{1+e^x} \tag{18}$$

as is used. This gives an expected proportion of 0.637. The intra-state correlation coefficient implied by the estimated intercept component variance, 36.02% of the variance in number of ANC visits is attributed to the state-level.

For case 2, Women who live in rural areas seem to be disadvantaged in access to ANC as they are less likely to make at least 4 visits to ANC services. The negative in the coefficients for women of Islamic and traditionalist background indicates that they are less likely to make 4 or more antenatal visits compared to their Christian counterparts. It thus seems that Christian women meet up the WHO standard than others. Women with primary, secondary and higher education are about 1.661, 2.187 and 4.221 times more likely to meet the WHO standard for ANC respectively than those with no education. One would note that this odds improved with advancement in education. Husband's education also has a significant effect on the propensity of women to ANC visits. Wives whose husbands have secondary and Higher education are about 1.448 and 1.933 times more likely to meet and/or exceed the minimum 4 ANC visits respectively than those whose husbands have below secondary education. Clearly, the rich are more likely to meet this benchmark than the poor and middle class citizens. On information gathering and ANC visits, radio and Television are the only significant source. Women who listen to radio and watch TV at least once a week are 1.259 and 1.319 times more likely to meet up with the minimum four (4) ANC visit than those who are irregular/ non-users of these media.

For the case 3, from the state level factors, North-East and South-East alone are not significant. The odds of women from the North-West, South-South and South-West making at least four antenatal visits are 0.416, 0.455 and 2.796 times that of their counterparts from the North-Central respectively. Women from the South-West seem to make more use of Antenatal services than others.

Both individual level factors and the state level factors are significant in understanding the variation in the number of antenatal visits women make to health facilities and the effect of geographical variation on this.

DISCUSSION AND CONCLUSION:

The results from the multilevel logistic analysis show that women in the various geopolitical zones vary in their tendencies to make at least four (4) ANC visits, thus, both the individual and the state level factors influence the number of ANC visits women make. The multilevel logistic analysis showed evidence that number of ANC visits is significantly associated with wealth index, location of residence, educational attainment (of the woman and also of her husband), religion and frequency and source of information gathering.

Husbands have a great role to play in ensuring that their wives attend ANC services regularly as such educational campaigns should be geared towards husbands and wives

REFERENCES

Adewoye K.R., Musa I.O, Atoyebi O.A, Babatunde O.A (2013). Knowledge and utilization of ANC services by women of Child Bearing Age in Ilorin East Local Government Area, North Central Nigeria. International Journal of Science and Technology. www.ejournalofsciences.org.

Hofmann D.A. (1997). An Overview of the Logic and Rationale of Hierarchical Linear Models. Journal of Management.23:723. DOI: 10.1177/014920639702300602. http://jom.sagepub.com

Joop J.H. and Cora J.M.M. (2005). Multilevel Analysis. Encyclopedia of Social Measurement. Vol.2. 785-793.

Khan H.R. and Shaw J.E.H (2011). Multilevel logistic Regression Analysis Applied to Binary Contraceptive.

Lilford R.J and Chard T. (1983). Problems and Pitfalls of risk assessment in ANC. British Journal of Obstetrics and Gynecology 90; 507-510.

Nigerian Health Review (2006)

NPC (2001): Children's and Women's rights in Nigeria: A wake-up call. Situation Assessment and Analysis. National Planning Commission.

Nwosu E.O., Urama N.E, Uruakpa C. (2012). Determination of ANC Services Utilization in Nigeria. www.iiste.org. Vol. 2 No. 6.

Okonofua, F.E, Abejide O.R, Makanjuola R.O. (1992).Maternal mortality in Ile-Ife, Nigerian: A Study of risk factors. Stud. Family Plann. 23(5): 319-324.

Vanneste A.M, Ronsmans C., Chakraborty J., Andre De Francisco (2000). Prenatal screening in rural Bangladesh: From prediction to care. Health Policy and Planning 15(1): 1 – 10. Oxford University Press.

World Health Organization (WHO)/United Nations Children's Fund (UNICEF) (2003). ANC in developing Countries: Promises, Achievements and Missed Opportunities – An Analysis of Trends, levels and differentials 1990 – 2001. WHO: Geneva.

www.accesstohealth.org

Yuster E.A. (1995). Rethinking the role of the risk approach and ANC in Maternal Mortality reduction. International Journal of Gyneacology and Obstetrics 50 (Suppl. 2): 559 – 561.

ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.5, No.4, 2015

Table 1: Estimated coefficients, Odds Ratio for Fixed and Random Part of Model

MODEL	INTERCEPT ONLY			INDIVIDUAL LEVEL			STATE LEVEL FACTOR		
	Coeff.	Odds ratio	Sig.	Coeff.	Odds ratio	p-value	Coeff.	Odds ratio	p-value
FIXED PART									
Intercept	0.562	1.754	0.014	0.416	1.517	0.016	0.461	1.586	0.059
Location									
Rural				-0.315	0.730	<0.001	-0.311	0.733	<0.001
Education									
Wife: Primary				0.508	1.661	<0.001	0.504	1.656	<0.001
Secondary				0.787	2.197	<0.001	0.782	2.185	<0.001
Higher				1.440	4.221	<0.001	1.434	4.195	<0.001
Husband: Sec				0.370	1.448	<0.001	0.370	1.448	<0.001
Higher				0.659	1.933	<0.001	0.661	1.937	<0.001
Religion									
Islam				-0.266	0.766	<0.001	-0.243	0.784	<0.001
Traditional				-0.410	0.663	0.030	-0.410	0.664	0.030
Wealth.Index									
Poor				-0.878	0.416	<0.001	-0.877	0.416	<0.001
Rich				-0.291	0.747	<0.001	-2.900	0.718	<0.001
Freq of Info.									
Paper>=1/wk				0.041	1.041	0.727*	0.043	1.044	0.708*
Radio.=1/wk				0.231	1.259	<0.001	0.232	1.261	<0.001
TV>=1/wk				0.277	1.319	<0.001	0.274	1.316	<0.001
Zone									
North East							-0.095	0.910	0.789*
North West							-0.876	0.416	0.007
South East							0.664	1.943	0.068*
South South							-0.787	0.455	0.020
South West							1.028	2.796	0.003
RANDOM PART									
U_{0j} (S.E)	1.361(0.163)			0.909 (0.111)			0.591 (0.074)		