AJAS / SPAS
ISSN 2316-0861

African Journal of Applied Statistics
Vol. 5 (1), 2018, pages 351–376.
DOI: http://dx.doi.org/10.16929/ajas/351.220

ISSN 2316-0861

Evaluating Likelihood Estimation Methods in Multilevel Analysis of Clustered Survey Data

Adeniyi Francis Fagbamigbe[1,*], **Babatunde Bowale Bakre**[1]

[1]Department of Epidemiology & Medical Statistics, Faculty of Public Health, College of Medicine, University of Ibadan, Nigeria.

Received on December 07, 2017; Accepted on June 16, 2018; Published Online on June 18, 2018

Abstract. (Short Abstract) Public health researchers often lay little or no emphasis on multilevel structure of clustered data and its likelihood estimation techniques. This has led to improper inferences. The aim of this research is to evaluate traditional methods and the different multilevel likelihood estimation procedures so as to compare their computational efficiencies.

Key words: Clustered survey; Likelihood; Adaptive Gaussian Quadrature; Penalized quasi likelihood, Modern contraception; Akaike's information criteria.
AMS 2010 Mathematics Subject Classification : 60-07;62H12

Presented Professor Gane Samb Lo, Gaston Berger University, Senegal
Member of the Editorial Board.

*Corresponding author: Adeniyi Francis Fagbamigbe (fadeniyi@cartafrica.org; franstel74@yahoo.com)
Babatunde Bowale Bakre: bakrebabatunde@yahoo.com

Full Abstract. Introduction: Public health researchers often lay little or no emphasis on multilevel structure of clustered data and its likelihood estimation techniques. This has led to improper inferences. The aim of this research is to evaluate traditional methods and the different multilevel likelihood estimation procedures so as to compare their computational efficiencies.

Methodology: We fitted mixed method effect regression model into data on use of modern contraceptive from the Nigeria 2012 National HIV/AIDS and Reproductive Health Survey (NARHS) PLUS II with respondent's characteristics as the independent variables. Also, 600,000 observations was simulated to evaluate the performance of Penalized Quasi-Likelihood (PQL), Non-Adaptive Gaussian Quadrature (NAGQ) and Adaptive Gaussian Quadrature (AGQ) using syntax for Mixed Effects Logit Models (XTMELOGIT) and Generalized Linear Latent and Mixed Models (GLLAMM) in Stata and Generalized Linear Mixed Models (GENLINMIXED) in SPSS.

Result: Full Maximum Likelihood (ML) methods had highest likelihood values with lowest standard error and was considered the best model for both two and three levels logistic regression in both the survey and simulated data. PQL procedure was least biased compared to the other multilevel full FL methods. The full likelihood method had the least $-2logL$, AIC and BIC for the two dataset. Which implies that full likelihood procedure had the best fitted model. Also, current age of the respondents, wealth index, residence, education and religion are significant predictors of modern contraceptive use.

Conclusion: Full ML performed better than quasi likelihood method at both two and three levels for both simulated and survey data. However, PQL appeared to be the best considering whether the estimates were biased or not. In terms of computational time, NAGQ with XTMELOGIT syntax was the fastest for two-levels and three levels model. The cluster-level effect is more significant than zonal level effect on modern contraceptive use in Nigeria.

Résumé (French) Les chercheurs en santé publique accordent souvent peu ou pas d'importance à la structure multi-niveau des données en grappes (clusterized) et à ses techniques d'estimation basées sur la vraisemblance. Cela peut conduire à des inférences incorrectes. Le but de cette recherche est d'évaluer les méthodes traditionnelles et les différentes procédures d'estimation de vraisemblance multiniveaux afin de comparer leur efficacité.

1. Introduction

Likelihood plays important roles in parameter estimation and it is synonymous with probability. It defines the function of parameters included in a statistical model. That is, a set of parameter value given outcome y is the probability of those observed outcome given the parameter values ($l(\theta/y) = p(y/\theta)$). Likelihood is one of the tools used in estimating parameters of multilevel models, including multilevel binary logistic models.

Multilevel model is a statistical model of parameter that varies at more than one level (Leyland & Goldstein (2001); Sampson *et al.* (1997)). This model can be seen as generalization of linear model, although they also extend to nonlinear models. Multilevel model are ideal for research design where the data is collected from study participants who were organized at two or more levels (Maas & Hox (2005); Srikanthan & Reid (2008)). In which case, one level is nested in the other. Usually, the unit of analysis are the individuals (at a lower level) who are nested in within an aggregate unit (at higher level) (Klotz *et al.* (1969); Li *et al.* (2011)). Multilevel (hierarchical) data structure causes correlation among observations within same clusters (Li *et al.* (2011)). Multilevel models present alternative analysis procedures to the famous univariate and multivariate analysis of measures that are collected repeatedly from same individuals. Over the years, the use of multilevel analysis to investigate public health problems has gained significant prominence (DiezRouz & Mair (2011); Leyland & Goldstein (2001)). This growth can be attributed to the need to understand how individuals are related to each other within groups and importance of such in understanding the distribution of health outcomes (DiezRouz & Mair (2011);Oye-Adeniran *et al.* (2004)). The growth has also been aided by increased use of multilevel methods in statistical methods and their applicability to a broad range of scenarios that have multilevel data. However, its use has been fully embraced in most public health research (Bingenheimer & Raudenbush (2004)).

The percent of total variance in the individual-level health outcome and the cluster effects which represent unobserved cluster characteristics that has potentials of affecting individuals outcomes could be large. (Li *et al.* (2011)). This must be viewed in light of the fact that the relevant "levels" are generally grossly mis-specified. So far, the methods of parameter estimation have led to several problems in the best way to carry out multilevel analysis, including under estimation of parameters and biased estimates (John *et al.* (2012)). In this study different methods of estimating multilevel binary logistic model parameters were considered and the best method was determined.

Cluster sampling, whereby samples are not taken randomly from entire population but from clusters, often introduces multilevel dependency and correlation among measurements taken from individuals within same cluster which could substantially affect parameter estimates. The structure of clustered survey data are usually nested and can be analysed using multilevel techniques. Challenges are often encountered when multistage sampling is used in data collection without the use of multilevel analysis. The description of most of "the theoretical and methodological challenges facing contextual analysis" has been made by Blalock (1984). The dependence among observations in multistage-clustered samples often comes from several levels of the hierarchy (Maas & Hox (2005)). In this case, the use of single-level statistical models is no longer valid and reasonable (Leyland & Goldstein (2001) ; Li *et al.* (2011)). The traditional standard logistic regression, that is single-level logistic regression, usually requires a sort of independence among the observations conditional on the independent variables and uncorrelated residual errors. To ensure that appropriate inferences are drawn and that reliable conclusions from clustered survey data is made, it has therefore become necessary to use more effective and more involving modeling techniques like multilevel modeling.

Also, underlying assumptions of ordinary logistic regression are violated when analyzing nested data, hence the best option is multilevel logistic regression analysis (Maas & Hox (2005); Srikanthan & Reid (2008)). This is due to the fact that it considers the variations due to multilevel structure in the data and allows the simultaneous assessment of effects of different levels in the data used in this study. The number of levels, the variance of the random effects and the size of the correlation between random effects may affect the performance of the parameter estimation method. Some methods of estimation could be biased. Therefore, there is need to evaluate these methods and determine the best method. The commonest methods used are Penalized Quasi-Likelihood (PQL), Non-Adaptive Gaussian Quadrature (NAGQ) and Adaptive Gaussian Quadrature (AGQ) and the Maximum Likelihood Estimates (MLE). Early methodology work on multilevel logit model includes use of data from 15 World fertility survey (Goldstein (2003);Hox, J. J. (2002)). Further documentations on multilevel models especially the type of data it allows, sampling, outliers, repeated measures, institutional performance, and spatial analysis have been made (Leyland & Goldstein (2001)).

The robustness, sample sizes and statitical power in multilevel modeling for both categorical and continuous outcome variables has been studied earlier (Bingenheimer & Raudenbush (2004); Goldstein (2003); Li *et al.* (2011); Maas & Hox (2005); Portnoy (1971)). Monte Carlo simulation has been used to "assess the impact of misspecification of the distribution of random effects on estimation of and inference about both the fixed effects and the random effects in multilevel logistic regression models" by Austin (2005). The authors concluded that inferences aboutg fixed effects estimate were not affected by the inherent misspecification of random effects distributions. However, the authors opined that inferences about random effects estimate were influenced by model misspecifications. Simulation studies indicated that increasing number of levels yield better estimates than larger number of individuals per level (Goldstein (2003); Goldstein & Rasbash (1996); Mason *et al.* (1983)). It was concluded in these studies that for second level units with a small sample size, while the estimates of the regression coefficients are unbiased, the standard errors and the variance components are sometimes underestimated (< 30) Maas & Hox (2004). This is not envisaged in the current study since we are using a large dataset.

The use of these statistical methods allows public health researchers to correctly identify factors and causes of disease at different levels. The approach provides opportunity and serves as a tool to investigate disease causation in complex settings.

Contraceptive Use in Nigeria

In 1988, the Nigeria Federal Ministry of Health adopted the "National Policy on Population for Development, Unity, Progress and Self-Reliance" (Essien *et al.* (2010)). It consequently adopted a revised policy in 2004. This was aimed at the reduction of maternal deaths by 75% by 2015 in compliance to the defunct Millennium Development Goal (MDG) which currently operating Sustainable Development Goal (SDG) (United Nations (2015); WHO (2012)). The 1988 National Policy on population encouraged open discussion and promotion of family planning (Essien *et al.* (2010)) as a tactic to encourage the utilization, improve the standard of living, promote health, reduce mortality and morbidity, slow down population growth and control population drift to urban areas.

An evaluation of the policy and the specific targets of the Nigerian Population Policy (NPP) indicated that none of the year 2000 targets could be met (Adekunle *et al.* (2000)). The contraceptive prevalence rate, currently at 14%–16%,(Khan & Shaw (2011)) is far from the targeted 84% of 2005, almost a decade after. Although the total fertility rate fell to about 5 per women from 6.2, this is clearly far from the targeted 4.0. The policy failure have been attributed to poor financing, poor use of contraceptives, lack of political will, poor and uncoordinated implementation strategies, leaving men out of the control, incessant policy changes as a result of political instability (Essien *et al.* (2010)).

Also, the medium of outreach of the NPP to the target population was too narrow. It used only the public sector and clinic-based, physician-controlled family planning programs. This narrowed the coverage and left large gap of unmet demand for contraception (Essien *et al.* (2010), **?**). More worrisome is the level of contraceptive use among sexually active young women. This was unimaginably low, at 7.3% (National Population Commission (Nigeria)) and 10% (Federal Ministry of Health Nigeria (2013)) of modern contraception. The low contraception has escalated number of unwanted pregnancy, unsafe abortions, related morbidities, school drop-outs and maternal mortality (Ankomah *et al.* (2000)). Nevertheless, substantial geographical variations and a slow increasing trend in use of modern contraceptive methods in Nigeria have been documented (**?**; Federal Ministry of Health Nigeria (2013); National Population Commission (Nigeria)). The precarious situation is worrisome and calls for evidence based scientific information that could be used to review strategies that would enhance improved use of modern contraceptive methods in Nigeria.

In the current study, we hypothesized that the various estimation methods will be similar in performance but that the multilevel methods for binary outcomes, particularly the PQL estimates will be more virulent for large random effect variance and large cluster sizes and that it is not more efficient than the AGQ. The main objective in this study is to evaluate the performances of the different multilevel analysis likelihood estimation procedures in determining the factors influencing modern contraceptive use in Nigeria.

2. Methodology

Source of Data

The data used in this study was from the Nigeria 2012 National HIV/AIDS and Reproductive Health Survey (NARHS Plus II) (Federal Ministry of Health Nigeria

(2013)) which was a nationally representative. The survey was carried out to provide information on key HIV/AIDS and reproductive health knowledge and behaviour related issues. Administratively, Nigeria is divided into six geographical zones. Each zone is subdivided into states and each state subdivided into Local Government Areas (LGAs). Using the 2006 census, each LGA is divided into localities and each locality subdivided into census Enumeration Areas (EAs) which are the Primary Sampling Units (PSU), referred to as clusters in the 2012 NARHS Plus II.

The survey utilized three-stage stratified cluster sampling. While all zones and all states were selected and included in the study, 30 clusters were randomly selected from each state irrespective of the number of LGAs. In all, data was collected from 31,235 individual respondents consisting of 15,596 males and 15,639 females from across 6 zones and 1110 clusters. This is illustrated in Figure 1. The units at lower level (level-1) are individuals (ever and never -married women aged 15-49) who are nested within units at higher level (clusters: level-2 which were either rural or urban) and the clusters were nested within the next higher level (zone: level-3 which are the six geopolitical zones in Nigeria). Similar disaggregation of levels can be found in the literature (Maas & Hox (2005); Srikanthan & Reid (2008)). Due to this nested structure, the odds of women to use any modern contraceptive methods are not independent, because individuals from the same cluster may share common exposure to community characteristics.

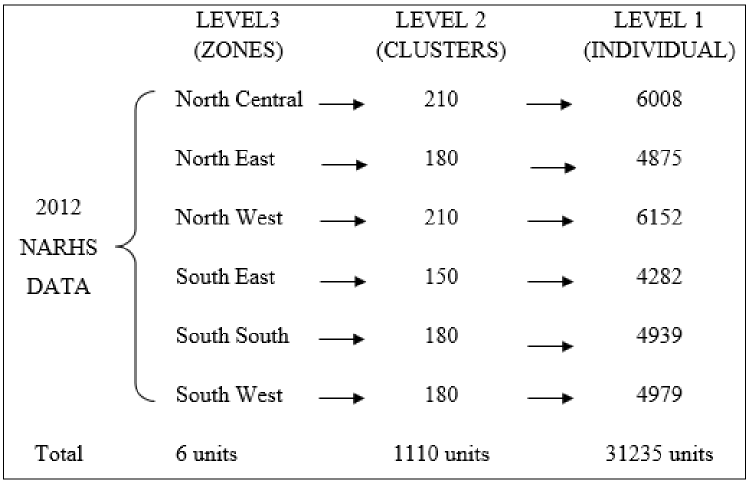

Fig. 1. Hierarchical structure of the 2012 NARHS data

Variables

The response variable in this study is "current use of modern contraception" coded as "1" for current uses and "0" if otherwise. The explanatory variables are respondents' current age, educational attainment, religion, place of residence and wealth index. The choice of this variables were motivated by earlier studies (Adebowale *et al.* (2011); Amin *et al.* (2002); Fagbamigbe *et al* (2015); Hox, J. J. (2010)).

Simulated data

We simulated 600,000 binary data with two possible outcomes (0 & 1) spread across 1200 clusters in the 6 hypothetical geographical zones in which the entire population was divided. Similar procedure has been used in Leckie & Goldstein (2015).

Methods of Analysis

In this research work, the methods for estimating multilevel binary logistic models were based on likelihood estimation. Among the known likelihood methods, Marginal Quasi-Likelihood (MQL) and Penalized Quasi-Likelihood (PQL) are the two most used likelihood approximation procedures. After applying these quasi likelihood methods, the parameters models were then estimated using Iterative Generalized Least Squares (IGLS) consisting of Adaptive Gaussians Quadrature (AGQ) and Laplacian approximation in Goldstein (2003) which is full Maximum Likelihood (ML) estimation procedures in estimating random intercept and fixed effect. In order to ascertain if the effect of the explanatory variable varies across zones and clusters and to allow the examination of both between and within group variability as well as how group level and individual level variables are related to variability at both levels, the following steps were followed in analyzing the data:

1. Fit a simple model with no predictors i.e an intercept-only model that predicts the probability of contraceptive use. The functional form of the model is

$$\ln\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \beta_{ooo} + \mu_{ojk} + \tau_{ook} \tag{1}$$

   where, $\ln\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = Odd$ of using the contraceptive in absence of all explanatory variables, β_{ooo} is the fixed intercept, μ_{ojk} is the level two random intercept and τ_{ook} is the level three random intercept. The random effect are assumed to be normally distributed with $\mu = 0$ and variance σ^2 that is;

$$\begin{bmatrix} \mu_{ojk} \\ \tau_{ook} \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma^2_{\mu o} \\ \sigma^2_{\tau o} \end{bmatrix}\right)$$

   The estimates of parameters and standard errors will be determined using PQL and MQL (AGQ and NAGQ)
2. Assess random intercept and a fixed slope for the variable by using multilevel univariate analysis for both survey and simulated data.
3. Assess all significant factors found in previous univariate analysis that affect contraceptive use.

4. Compare the three methods of parameter estimation by checking standard error and some information criteria in the two dataset.

Modeling a Binary Data

The probability of using modern contraceptive p_{ij} ranges between 0 and 1, let p_{ij} be modeled using a logistic function. The standard assumption is that the outcome variable follows a Bernoulli distribution.

Model for Three Levels Five Predictors Logistic Regression with Random Intercept and Fixed Slope

The commonest basic expansion of a fixed-effects regression model to a multilevel model is to ensure the flexibility of the intercept term, that is to allow it to vary randomly over different groups. This will ensure that the regression slopes remain fixed, while the intercept term is not fixed. The Level-1 model can be expressed as stated in equation (2)

$$\ln\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \beta_{ojk} + \beta_1 x_{1ijk} + \beta_2 x_{2ijk} + \beta_3 x_{3ijk} + \beta_4 x_{4ijk} + \beta_5 x_{5ijk} \quad (2)$$

By randomizing β_{ojk} we obtained,

$$\beta_{ojk} = \beta_{ook} + \mu_{oj} \;\; (level \;\; two) \quad (3)$$

$$\beta_{ook} = \beta_{ooo} + \tau_{ook} \;\; (level \;\; three) \quad (4)$$

The reduced form is derived by the simple substitution of equation (4) into equation (3), which results in

$$\beta_{ojk} = \beta_{ooo} + \tau_{ook} + \mu_{oj} \quad (5)$$

By substituting equation (5) in equation (2), the model resulted in equation (6)

$$\ln\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \beta_{ooo} + \tau_{ook} + \mu_{oj} + \beta_1 x_{1ijk} + \beta_2 x_{2ijk} + \beta_3 x_{3ijk} + \beta_4 x_{4ijk} + \beta_5 x_{5ijk} \quad (6)$$

where

$\ln\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = Odd$ of using modern contraceptive

x_{1ijk} = Current age
x_{2ijk} = Wealth Index
x_{3ijk} = Educational Attainment

x_{4ijk} = Religion
x_{5ijk} = POR (Place of Residence)
$\beta's$ = Fixed effect parameters
$\mu's$ = Random effect parameter of the cluster at level two
$\tau's$ = Random effect parameter of region at level three.

The random intercept denoted by β_{ojk} is an additive function of a grand mean (β_{ooo}) and a group-levels deviation from this mean are μ_{oj} and τ_{ook}. The random effect are assumed to be normally distributed that is;

$$\begin{bmatrix} \mu_{ojk} \\ \tau_{ook} \end{bmatrix} = N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} \sigma^2_{\mu o} \\ \sigma^2_{\tau o} \end{bmatrix}\right)$$

The parameter of equation (11) containing fixed effect, random effect, variance of the random effect and residual variance were simultaneously estimated using iterative method (Bryk & Raudenbush (1992); Bryk & Raudenbush (2002); Kalam & Khan (2002); Mason *et al.* (1983)).

The Gaussian Quadrature Methods

The marginal likelihood of the observed data can be obtained by integrating the distribution of the random effects, marginal likelihood $L(y)$ conditional on the independent variables in the model as shown in equation (7) as demonstrated in Goldstein (2003).

$$L(y) = \int_{-1}^{+1} \pi_{j=1}^{N}\pi_{i=1}^{n_j} f_{y_{ij}/u_j}(y_{ij}/v_j) f_{u_j}(v_j) du_j = \pi_{j=1}^{N}\pi_{i=1}^{n_j} f_{y_{ij}/u_j}(y_{ij}/v) f_u(v) dv \tag{7}$$

Where $L(y)$ depends on the unknown parameters γ_o, γ_1, γ_2, γ_3, σ_o, σ_1, σ_{01}, which are the random effect parameters. Evaluating equation (7) requires the computation of N integrals of dimension m. The likelihood of equation (7) will be maximized with respect to the unknown parameters of the model.

Generally, the integral of equation (8) has no closed form and would have to be computed numerically in R or Matlab. Maximization of the likelihood can be done using the standard methods. The methods usually requires many iterations before convergence could be reached during the iterative maximization procedures. therefore, there is need to use fast but reliable methods to solve equation (8). In this paper we compared the performance of PQL developed in Daniel & Sonya (2011) and full likelihood approaches which are AGQ and NAGQ as described by Hox, J. J. (2010). An alternative approach is to approximate the integral in equation (7) by numerical integration and maximization of the likelihoods. Numerical integration follows the Gauss-Hermite quadrature formula as shown in equation (8). See Goldstein (2003) for details.

$$\int_{-\infty}^{+\infty} h(u)e^{-v^2} du = \sum_{q=1}^{k} h(x_q) w_q \tag{8}$$

where h is a smooth function, $x_1, ..., x_k$ are the quadrature points and $w_1, ...w_k$ are the associated weights summing to one. The larger the k (the number of quadrature points), the better the approximation in (8). The estimator obtained by maximizing the likelihood approximated in this way is called the NAGQ estimator. The performance of the Gaussian quadrature could be improved by using adaptive integration methods which took into account the properties of the integrand. Such methods scale and translate the quadrature locations to place them under the peak of the integrand (Goldstein (2003)), resulting in an estimator known as the AGQ.

Intra-Class Correlation Coefficient

In a multilevel model, the sources of variation could be within-group or between-group Li *et al.* (2011). Thus, in this study, the total variation in individual outcomes can be partitioned into: variability among individual in the same cluster group and in the same geo-political zone and variability between individuals within different cluster and geo-political zone. Thus, when individual within group are very similar to each other, less information is obtained compared to when there are same number of individuals in an unclustered sample. We called the amount of variation in the use of modern contraceptive explained by the cluster variable and geopolitical zone "Intra Class Correlation Coefficient" (ICC). It described the dependencies in the data and measured the extent to which individuals within the same group are more similar to each other than they are to individual in different groups. For binary responses, the ICC is usually stated in terms of the correlation between the latent responses. The logistic distribution for the level one residual e_{ij} implies a variance of $\pi^2/3 = 3.29$. This implies that for three level logistic random intercept model the level three ICC is

$$\rho = \frac{\sigma^2_{\tau o}}{\sigma^2_{\tau o} + \sigma^2_{\mu o} + \pi^2/3}.$$

Where $\sigma^2_{\tau o}$ is the level three constant variance and $\sigma^2_{\mu o}$ is the level two constant variance.

While the ICC for level two is

$$\rho = \frac{\sigma^2_{\mu o} + \sigma^2_{\tau o}}{\sigma^2_{\tau o} + \sigma^2_{\mu o} + \pi^2/3}.$$

We used Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC) to compare the performance of the nested models.

Data Management

We used IBM SPSS version 25 to fit mixed effect model using PQL method by Daniel & Sonya (2011) and Stata 12 was used to estimate single level fixed effect and multilevel fixed and random effect parameter using both AGQ and NAGQ. However, the AGQ estimates was generated using fifteen quadrature point for both Mixed

Effects Logit Models (XTMELOGIT) and Generalized Linear Latent and Mixed Models (GLLAMM) while Generalized Linear Mixed Models (GENLINMIXED) was used to determine the linear mixed effects. Also, NAGQ estimate was generated using Laplacian approximation.

Both GLLAMM and XTMELOGIT were used in multilevel models for the odds. However, the latter can accommodate more complex multilevel structures. GLLAMM is an "user-built" command in Stata and is particularly good for fitting both generalized linear latent and mixed models which belong to a class of multilevel latent variable models (Skrondal & Rabe-Hesketh (2004)). We used the "adapt" option in GLLAMM in Stata to specify the AGQ method as done in earlier studies (Li *et al.* (2011)). Alternatively, XTMELOGIT, an inbuilt Stata command, fits mixed-effects models for binary and binomial outcomes. In which case the mixed models consist of both random and fixed effects. It helps to overcome the incapability of standard logistic regression in estimating fixed effects (Andrew *et al.* (2000)).

Abbrevation	Full Meaning
AGQ	Adaptive Gaussian Quadrature
AIC	Akaike's information criteria
AIDS	Acquired Immune Deficiency Syndrome
BIC	Bayesian information criteria
EA	Enumeration Areas
GENLINMIXED	Generalized Linear Mixed Models
GLLAM	Generalized Linear Latent and Mixed Models
HIV	Human Immunodeficiency Virus
ICC	Intra Cluster Correlation
IGLS	Iterative Generalized Least Square
LGA	Local Government Area
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimates
MQL	Marginal Quasi-Likelihood
NAGQ	Non-Adaptive Gaussian Quadrature
NARHS	National HIV/AIDS and Reproductive Health Survey
PQL	Penalized Quasi-Likelihood
PSU	Primary Sampling Units
SPSS	Statistical Package for Social Sciences
XTMELOGIT	Mixed Effects Logit Models

3. Results

Descriptive statistics

In all, 31235 respondents participated in the survey. This consisted of 15,596 males and 15,639 females from across 1110 clusters nested in the 6 geopolitical zones. Each cluster had average of 30 respondents while number of clusters in each zone ranged 150-210 as shown in Figure 1. The mean age of male and female respondents was 34.0 ± 4.0 and 29.2 ± 9.5 years respectively.

Three Level Intercept Only Multilevel Logistic Model

The outcome of the binary model in the different methods considered were similar except among the standard logistic method and the PQL, For instance the fixed effect intercept was -1.95, -1.987, -2.428, -2.430 and -2.425 for standard logistic, PQL, NAGQ AND AGQ (XTMELOGIT and GLLAMM) respectively. As shown in Table 1 , the fixed and random intercept for three level in all the methods are significant except the random effect at level three. The standard logistic regression model overestimated the parameter compared to the multilevel methods. Table 1 also showed level three model comparison using $-2log-likelihood$, AIC and BIC. Considering the intercept only model for three levels using the quasi and the full ML methods (GENLINMIXED, XTMELOGIT and GLLAMM), AGQ using GLLAMM with fifteen quadrature point have the lowest $-2logL$ (21191.626), AIC (21,197.626) and BIC (21,222.673). The appropriateness of AGQ with GLLAMM syntax was further confirmed in Table 1 as its estimates had the smallest standard error for both fixed and random effect except for level three which is the zone level.

Table 1. Three-level estimates of multilevel analysis using an intercept only, single level and multilevel logistic model to predict modern contraceptive use from the survey data

Model Effect	Standard logistic	PQL	NAGQ XTMELOGIT	AGQ XTMELOGIT	AGQ GLLAMM
Fixed effect	-1.95**	-1.987**	-2.428**	-2.430**	-2.425**
Intercept	-0.171	-0.18	-0.046	-0.047	-0.03
$\sigma^2_{\mu o}$ (BCV)		0.124	1.411717	1.457	1.012
		-0.245	-0.1	-0.104	-0.076
$\sigma^2_{\tau o}$(BZV)		0.003	5.17E-07	2.12E-08	0.5458
		-0.17	-0.001	-0.001	-0.021
Intra CCC		0.037	0.3	0.307	0.321
Intra ZCC		8.77E-04	1.10E-07	4.47E-09	0.115
$-2logL$	23418.35	152294.57	21503.09	21490	21191.626
AIC	23427.35	152300.57	21509.09	21496	21197.626
BIC	23452.397	152325.617	21534.137	21521	21222.673
Iteration	1	3	13	11	3
Computation Time	0.5 minutes	3.5 minutes	3.02 minutes	6.0 minutes	18.1 minutes
N	31235	31235	31235	31235	31235

BCV Between Cluster Variance, BZC Between Zone Variance, CCC Cluster Correlation Coefficient, ZCC Zonal correlation coefficient. **significant at 0.01 *significant at 0.05. Standard error in parenthesis.

Intra Class Correlation for Three Levels Intercept Only Model

In the multilevel methods, the ICC reduced from 32% in AGQ (GLLAMM) to 4% in PQL. The AGQ (XTMELOGIT) have 31% of the total variance explained within the cluster, while NAGQ (XTMELOGIT) and PQL produced 30% and 4% of the total variance explained within the cluster respectively. Also, the random effect for the third level using XTMELOGIT syntax with fifteen quadrature point was approximately zero in the NAGQ and AGQ methods since they had ICC approximately zero. This implied that using geo-political zone as a level was not reliable. However, AGQ (GLLAMM) produced 11% of the total variation explained within the zone while PQL had 10% of the total variance explained among the zones (Table 1).

Two Level Intercept Only Multilevel Logistic Model (Survey Data)

Using a simpler model consisting of only individual and cluster levels, we fitted an intercept-only model that predicted the probability of modern contraceptive use. The estimates of parameters and standard errors are presented in Table 2.

The Predicted Probability of Modern Contraceptive use

With AGQ (XTMELOGIT), the expected log-odds of modern contraceptive use is -2.430, implying odds of $e^{(-2.430)} = 0.088$ and a predicted probability of $1/(1 + e^{(2.430)}) = 0.0809$. For AGQ (GLLAMM) the expected log-odds of contraceptive use is -2.430 with corresponding predicted probability of $1/(1+e^{(2.430)}) = 0.0809$ NAGQ: the expected log-odds of contraceptive use was -2.428. This corresponded to a predicted probability of $1/(1 + e^{(2.428)}) = 0.0811$. PQL: the expected log-odds of contraceptive use is -1.987 and this corresponded to a predicted probability of $1/(1 + e^{(1.874)}) = 0.1331$ Standard logistic model has a predicted probability of $1/(1 + e^{(1.955)}) = 0.1240$ which is the same as the sample ratio of 3874 modern contraceptive users to 28000 non-users. It is the odds-ratio when no predictors have been considered in the model.

Based on the associated standard errors, multilevel methods provided more accurate estimate for the predicted probability than the standard logistic estimation method however estimate from PQL was similar to estimate obtained using the standard logistic regression. Compared to the odds-ratios obtained from the multilevel estimation methods, the standard logistic model odds-ratio seem to have overestimated the parameter. There was a significant difference between the standard logistic estimate and the multilevel logistic estimates. By failing to take into account the possible variability among the clusters (level 2), Compared to multilevel models using PQL, NAGQ and AGQ, the standard logistic model overestimated the odds-ratio by about 2% 19% and 20% respectively (Table 2). Since bias of 5-10% is often considered tolerable (Hox, J. J. (2002)), PQL produced best estimate in terms of bias. The random quantity at cluster level was underestimated by PQL compared to full likelihood method. However, the full likelihood had the smallest standard error.

Convergence of the Estimation Methods in Two Levels for Survey Data

In Tables 1 and 2, the PQL converged faster than other multilevel methods. Table 2 shows that AGQ method using GLLAMM and XTMELOGIT had the smallest $-2logL$ of 21490.132, AIC (21494.133) and BIC (21510.831) among other multilevel binary logistic methods.

Table 2. Two-level estimates of multilevel analysis using an intercept only, single level and multilevel logistic model to predict modern contraceptive use from the survey data

Model Effect	Standard logistic	PQL	NAGQ XTMELOGIT	AGQ XTMELOGIT	AGQ GLLAMM
Fixed effect	-1.955**	-1.874**	-2.428**	-2.430**	-2.430**
Intercept	-0.017	-0.294	-0.046	-0.045	-0.045
$\sigma^2_{\mu o}$(BCV)		0.172 -0.245	1.412 -0.1	1.207 -0.043	1.456 -0.054
Intra CCC		0.0497 (5%)	0.300 (30%)	0.268 (27%)	0.307 (31%)
$-2logL$	23418.35	158912.752	21503.09	21490.132	21490.132
AIC	23422.35	158916.752	21507.09	21494.132	21494.132
BIC	23426.85	158922.75	21523.785	21510.831	21510.831
No of Iterations	1	6	3	2	2
Computation time	0.5 minutes	1.08 minutes	0.92 minutes	1.45 minutes	3.30 minutes
Number of observation	31235	31235	31235	31235	31235
Number of group	1076	1076	1076	1076	1076

BCV Between the Cluster Variance, CCC Cluster Correlation Coefficient, **significant at 0.01 *significant at 0.05. Standard error in parenthesis.

Random Effect of Two Levels Intercept Only Model for Survey Data

The parameters under random effect in Table 2 was the estimated the variances of the random intercepts at level 2 for fitting a two-level intercept-only model. To understand the random effect in this two-level intercept-only model, one can imagine a unique effect for each cluster (level 2) in addition to the fixed intercept of -2.430 (AGQ estimate with XTMELOGIT), -2.425 (AGQ estimate with GLLAMM), -2.428 (NAGQ estimate with XTMELOGIT) and -1.874 (PQL estimate with GENLINMIXED) which is the average of modern contraceptive use in all cluster. The addition of the cluster specific effects makes the model more accurate than the fixed intercept only model. In the random effect model, the cluster effects were assumed to be distributed normally for the purpose of estimation (Li *et al.* (2011)). In Table 2, the estimate of the random effect at level two were higher in the PQL, the NAGQ and the AGQ (XTMELOGIT and GLLAMM syntax). Also the standard error of the random effect in AGQ (XTMELOGIT) was the smallest which implies that it is more efficient.

4. Simulation Study

Three Level Intercept Only Multilevel Logistic Model for Simulated Data

Using $n = 600,000$ simulated data, the parameter estimates in three level analysis shown in Table 3 revealed that fixed and random intercept for three level in all the methods are significant except the random effect at level three. The standard logistic regression model overestimated the parameter compared to the multilevel methods. Considering the intercept only model for three levels using the quasi and the full ML methods (GENLINMIXED, XTMELOGIT and GLLAMM), AGQ using GLLAMM and XTMELOGIT with 15 integration point have the smallest $-2logL(3016.012)$, AIC (3022.012) and BIC (3033.411).

Table 3. Three-level estimates of multilevel analysis using an intercept only, single level and multilevel logistic model predict modern contraceptive use from the simulated data

Model Effect	Standard logistic	PQL	NAGQ XTMELOGIT	AGQ XTMELOGIT	AGQ GLLAMM
Fixed effect Intercept	-2.431** -0.047	-2.873** -0.18	-3.149** -0.168	-3.149** -0.168	–3.149** -0.168
$\sigma^2_{\mu o}$ (BCV)		0.172 -0.245	1.042 -18.16	2.143 -0.457	2.143 -0.457
$\sigma^2_{\tau o}$ (BCV)		0.395 -0.17	1.041 -18.16	0.06 -0.043	6.95E-11 0
Intra CCC		0.147(15%)	0.387(39%)	0.401(40%)	0.394(39%)
Intra ZCC		0.102(10%)	0.194(19%)	0.011(1%)	1.278e-11(0)
$-2logL$	3369.599	152294.57	3017.473	3016.012	3016.012
AIC	3371.599	152300.57	3023.473	3020.012	3022.012
BIC	3378.299	152307.27	3036.872	3033.411	3033.411
Iteration	1	3	5	5	4
Computation time	30sec	1min,30sec	15secs	6 Mins	19mins,37secs
N	600000	600000	600000	600000	600000

BCV Between the Cluster Variance, BZC Between the Zone Variance, CCC Cluster Correlation Coefficient, ZCC Zonal correlation coefficient. **significant at 0.01 *significant at 0.05. Standard error in parenthesis.

Intra Class Correlation for Three Levels Intercept Only Model for Simulated Data

In the multilevel methods, XTMELOGIT syntax with fifteen integration point has highest two-level - cluster correlation coefficient (CCC) (40%) while PQL with GENLINMIXED syntax has the smallest CCC (15%). The AGQ (GLLAMM) with fifteen integration point and NAGQ (Laplacian approximation) have 39% of the total variance explained within the cluster, Also, the random effect for the third level using AGQ (XTMELOGIT and GLLAMM) was approximately zero in which their ICC was about zero at fifteen integration point, but ICC for PQL and NAGQ (Laplacian approximation) was respectively 10% and 19%. Due to inconsistent estimate of the level three random effect, using geo-political zone as a level might not be reliable (Table 3).

Two Level Intercept Only Multilevel Logistic Model for Simulated Data

We also fitted a two level intercept-only model that predicted the probability of modern contraceptive using the 600,000 simulated data. The estimates of parameters and standard errors are presented in Table 4 . The MLE from the standard logistic model of the ratio of modern contraceptive user to Modern contraceptive nonuser was $e^{(-2.432)} = 0.088$. In comparison, the same parameter is estimated to be $e^{(-2.695)} = 0.068$ using PQL and $e^{(-3.149)} = 0.042$ for both NAGQ and AGQ in multilevel model methods. Compared to the odds-ratios obtained from the multilevel estimation methods, the standard logistic model odds-ratio seem to have overestimated the parameter. The difference between the well known standard logistic estimate and the multilevel logistic estimate was significant (See Table 4).

Convergence of the Estimation Methods in Two Levels Simulated Data

In Table 4, the NAGQ (XTMELOGIT with Laplacian approximation) converged faster than other multilevel methods. We showed that AGQ method using GLLAMM and XTMELOGIT with fifteen integration point had the smallest $-2logL$ of 3016.012, AIC (3020.012) and BIC (3033.411) among other multilevel binary logistic methods.

Random Effect of Two Levels Intercept Only Model

The random effect estimates in Table 4 are the estimated variances of the random intercepts at level 2. We found GLLAMM, XTMELOGIT with fifteen integration point and Laplacian approximation had the same fixed intercept (-3.149) but different from quasi likelihood (PQL estimate with GENLINMIXED of -2.695. In Table 4 , the estimate of the random effect at level two were higher in the full ML method compare to the quasi likelihood. Also the standard error of the random effect in NAGQ (XTMELOGIT with one integration point) was the smallest which implies that it is more efficient.

Table 4. Two-level estimates of multilevel analysis using an intercept only, single level and multilevel logistic model to predict modern contraceptive use from the simulated Data

Model Effect	Standard Logistic	PQL	NAGQ	AGQ XTMELOGIT	AGQ GLLAMM
Fixed effect	-2.432**	-2.695**	-3.149**	-3.149**	-3.149**
Intercept	-0.047	-0.234	-0.169	-0.168	-0.168
$\sigma^2_{\mu o}$ (BCV)		1.537 -0.603	2.083 -0.44	2.143 -0.457	2.144 -0.457
Intra CCC		0.318 (32%)	0.387 (39%)	0.394 (39%)	0.394 (39%)
$-2logL$	3369.599	32747.728	3017.472	3016.012	3016.012
AIC	3371.599	32751.727	3021.473	3020.012	3020.012
BIC	3378.299	32758.425	3034.872	3033.411	3033.411
Iteration	1	7	3	2	2
Computation time	0.03 minutes	0.18 minutes	0.08 minutes	0.12 minutes	0.77 minutes
Number of observation	600000	600000	600000	600000	600000
Number of group		1100	1100	1100	1100

BCV Between the Cluster Variance, CCC Cluster Correlation Coefficient. AIC Akaike's Information Criteria, BIC Bayesian Information Criteria. **significant at 0.01 *significant at 0.05. Standard error in parenthesis.

The Predicted Probability of Modern Contraceptive use with Simulated Data

With NAGQ and AGQ (GLLAMM and XTMELOGIT syntax), the estimated probability of modern contraceptive use was 0.042 and for PQL, the expected probability of contraceptive use was 0.063. However, the standard logistic model has a predicted probability of 0.081. Based on $-2logL$, AIC and BIC, AGQ (XTMELOGIT and GLLAMM with 15 integration point) the smallest estimate which implies that the best binary logistic model can be fitted by considering the level effect using AGQ.

5. Determinants of Modern Contraceptive use in Nigeria

Significance of the Covariates in Three Level Model using Survey Datasix

All the covariates considered in the models were significant predictor of modern contraceptive use in Nigeria. As shown in Table 5, the $-2log$ likelihood BIC and AIC estimates for multilevel model were less than that of the standard logistic regression. Furthermore, among the multilevel estimates, AGQ generates a better model compare to PQL and NAGQ.

Convergence of the Estimation Methods when Fixed Effects for Three Level Models are included.

Inclusion of the three levels in the model showed a wide variability in the convergence as well as computational times of the different procedures. The PQL had shortest convergence time followed by NAGQ, Also, AGQ (XTMELOGIT) converged faster than AGQ (GLLAMM).

Variance Component for Three-Level Model.

The random effect of the three levels model in Table 6 showed that the variance among the clusters in AGQ with fifteen integration points is more than variance obtained from all other methods of parameter estimation. The variance among the zones was approximately zero for all the methods. The PQL had the smallest random intercept with largest $-2logL$, AIC and BIC which minimized the reliability of model generated by the method.

Intra Cluster Correlation Coefficient

Based on three level model presented in Table 5, it was apparent that AGQ (GLAMM) had the largest ICCC (0.192) compared with other methods at cluster level. This implied that 19% of the total variance was explained by the variance within the cluster. The PQL had an ICCC of 0.074. Also, AGQ (GLAMM) had the largest IZCC (0.025) which implies that 3% of the total variance was explained by the variance within the zone, whereas AGQ (XTMELOGIT) and NAGQ (XTMELOGIT) had zero intra zone correlation coefficient.

Table 5. Three-level estimates of univariate multilevel quasi likelihood (PQL) and full ML (NAGQ, AGQ with XTMELOGIT and AGQ with GLLAMM) methods from the survey data

Variables	Standard Logistic	PQL	NAGQ	AGQ (XTMELOGIT)	AGQ (GLLAMM)
Constant	-4.789(0.107)**	-4.741(0.145)**	-4.587(0.113)**	-4.587(0.114)**	-4.833(0.227)**
Age					
20-24	1.095(0.075)**	1.096(0.075)**	1.2012(0.079)**	1.203(0.079)**	1.210(0.080)**
25-29	1.372(0.074)**	1.367(0.074)**	1.512(0.078)**	1.524(0.079)**	1.517(0.079)**
30-39	1.249(0.071)**	1.245(0.072)**	1.431(0.076)**	1.433(0.081)**	1.409(0.076)**
40-49	1.063 (0.077)**	1.063(0.077)**	1.202(0.0809)**	1.204(0.081)**	1.173(0.081)**
50-64	0.579 (0.010)**	0.596(0.100)**	0.732(0.105)**	0.733(0.105)**	0.702(0.105)**
Wealth Index					
Poorer	0.408(0.076)**	0.372(0.077)**	0.404(0.084)**	0.405(0.084)**	0.338(0.084)**
Average	0.581(0.074)**	0.526(0.076)**	0.635(0.084)**	0.635(0.087)**	0.504(0.087)**
Wealthier	0.629(0.077)**	0.576(0.079)**	0.715(0.092)**	0.726(0.091)**	0.612(0.092)**
Wealthiest	0.682(0.081)**	0.637(0.084)**	0.864(0.099)**	0.834(0.097)**	0.753(0.098)**
Education					
Quranic	-0.204(0.155)	-0.154(0.156)	-0.160(0.160)	-0.091(0.160)	0.018(0.161)
Primary	0.945(0.080)**	0.912(0.080)**	0.915(0.083)**	0.914(0.083)**	0.747(0.084)**
Secondary	1.241(0.076)**	1.213(0.076)**	1.326(0.077)**	1.325(0.078)**	1.086(0.080)**
Higher	1.507(0.083)**	1.498(0.083)**	1.705(0.084)**	1.705(0.085)**	1.429(0.088)**
Religion					
Non-Catholic	0.655(0.047)**	0.539(0.053)**	0.553(0.058)**	0.550(0.059)**	0.352(0.064)**
Catholic	0.708(0.060)**	0.713(0.066)**	0.662(0.074)**	0.658(0.075)**	0.481(0.080)**
Traditional	0.472(0.2380)*	0.375(0.240)	0.407(0.256)	0.405(0.256)	0.270(0.258)
No Religion	0.875(0.275)**	0.665(0.278)**	0.974(0.292)**	0.973(0.293)**	0.761(0.295)**
Others	-0.184(0.405)	- 0.338(0.406)	0.166(0.321)	0.164(0.321)	0.103(0.294)
POR	-0.182(0.042)**	-0.128(0.046)*	-0.372(0.068)**	-0.371(0.068)**	-0.115(0.078)**
$\sigma^2_{\mu o}$(BCV)		0.213(0.427)	0.670(0.050)	0.682(0.056)	0.684(0.057)
$\sigma^2_{\tau o}$(BZV)		0.049(0.033)	1.84E-08(0.000)	7.00E-07(0.001)	0.103(0.066)
Intral CCC		0.074	0.169	0.172	0.192
Intrar ZCC		0.014	0	0	0.025
-2LOGL	20760.79	20050.88	19949.32	19940.17	19939.15
AIC	20772.79	20086.88	19996.32	19984.17	19983.15
BIC	20787.76	20237.12	20176.4	20168.99	20166.79
Iteration	5	20	12	11	4
Computation	0.6 minutes seconds	0.7 minutes	19.4 minutes	45 minutes seconds	2433.1 minutes
N	31235	31235	31235	31235	31235

Additional legend to Table 5 BCV Between the Cluster Variance, BZC Between the Zone Variance, CCC Cluster Correlation Coefficient, ZCC Zonal correlation coefficient. **significant at 0.01 *significant at 0.05. Standard error in parenthesis. Reference categories: '15-19' for Age, 'urban' for POR, 'No formal education' for Education, 'Islam' for Religion, and 'Poorest' for WI Standard errors are placed in parentheses.

Significance of the Covariates in Two Level Model using Survey Data

In Table 6 , we present the comparison of multilevel logistic regression methods and the standard logistic regression when covariates were included in the models and only two levels considered. Age group of the respondent was significantly associated with modern contraceptive use in the three multilevel binary logistic regression methods at one percent level of significance $(p - value < 0.001)$ as well as the standard logistic regression. The wealth index was also found to be significantly associated with modern contraceptive use at one percent level of significant $(p - value < 0.001)$ across all the methods. Also, the $-2log$ likelihood and AIC estimates for multilevel model were less than that of the standard logistic regression.

Convergence of the Estimation Methods when Fixed Effects for two Level Models are included

Reduction of the three-level to two level model in Table 6 changed the computational time, convergence rate and the estimate of the intercept. PQL convergence time was the shortest followed by NAGQ, also, AGQ (XTMELOGIT) converged faster than AGQ (GLLAMM)

Table 6. Two-level estimates of univariate multilevel quasi likelihood (PQL) and full ML methods (NAGQ, AGQ with XTMELOGIT and AGQ with GLLAMM) from the Survey Data

Variables	Standard Logistic	PQL	NAGQ	AGQ (XTMELOGIT)	AGQ (GLLAMM)
Constant	-4.789(0.107)**	-4.883(0.135)**	-5.092(0.130)**	-5.094(0.130)**	-5.094(0.130)**
Age					
20-24	1.095(0.075)**	1.094(0.085)**	1.202(0.079)**	1.204(0.079)**	1.204(0.080)**
25-29	1.372(0.074)**	1.372(0.084)**	1.512(0.079)**	1.514(0.079)**	1.514(0.079)**
30-39	1.249(0.071)**	1.249(0.081)**	1.406(0.076)**	1.408(0.081)**	1.408(0.076)**
40-49	1.063 (0.077)**	1.063(0.087)**	1.177(0.081)**	1.178(0.081)**	1.178(0.081)**
50-64	0.579 (0.010)**	0.579(0.150)**	0.705(0.105)**	0.706(0.105)**	0.706(0.105)**
Wealth Index					
Poorer	0.408(0.076)**	0.408(0.086)**	0.409(0.084)**	0.409(0.084)**	0.409(0.084)**
Average	0.581(0.074)**	0.584(0.085)**	0.614(0.086)**	0.614(0.087)**	0.614(0.087)**
Wealthier	0.629(0.077)**	0.634(0.097)**	0.738(0.091)**	0.738(0.091)**	0.738(0.091)**
Wealthiest	0.682(0.081)**	0.688(0.091)**	0.882(0.969)**	0.883(0.097)**	0.883(0.097)**
Education					
Quranic	-0.204(0.155)	-0.205(0.175)	-0.108(0.161)	-0.108(0.161)	-0.108(0.161)
Primary	0.945(0.080)**	0.945(0.090)**	0.802(0.084)**	0.802(0.084)**	0.802(0.084)**
Secondary	1.241(0.076)**	1.242(0.086)**	1.139(0.080)**	1.139(0.080)**	1.139(0.080)**
Higher	1.507(0.083)**	1.508(0.093)**	1.464(0.088)**	1.464(0.088)**	1.464(0.088)**
Religion					
Non-Catholic	0.655(0.047)**	0.653(0.067)**	0.307(0.039)**	0.504(0.059)**	0.504(0.059)**
Catholic	0.708(0.060)**	0.705(0.079)**	0.609(0.075)**	0.606(0.075)**	0.606(0.075)**
Traditional	0.472(0.2380)*	0.470(0.288)*	0.409(0.237)	0.407(0.257)	0.407(0.257)
No Religion	0.875(0.275)**	0.875(0.295)**	0.974(0.293)**	0.972(0.293)**	0.972(0.293)**
Others	-0.184(0.405)	- 0.185(0.455)	-0.276(0.425)	-0.279(0.425)	-0.279(0.425)
POR	-0.182(0.042)**	-0.091(1.544)*	-0.166(0.072)*	-0.166(0.072)*	-0.166(0.072)*
$\sigma^2_{\mu o}$(BCV)		0.180(0.427)	0.818(0.033)	0.825(0.034)	0.681(0.056)
Intral CCC		0.0519	0.199	0.201	0.177
$-2LOGL$	20760.79	158912.752	199952.313	19948.83	19948.83
AIC	20772.79	158914.752	19996.313	19962.83	19992.825
BIC	20787.76	158923	20176.4	20176.6	20179.9
Iteration	5	20	3	2	3
Computation	0.6 minutes seconds	0.65 minutes	12.8 minutes	30.2 minutes	203 minutes
N	31235	31235	31235	31235	31235

Additional legend to Table 6 : BCV Between the Cluster Variance, CCC Cluster Correlation Coefficient. **significant at 0.01 *significant at 0.05. Standard error in parenthesis. Reference categories: '15-19' for Age, 'urban' for POR, 'No formal education' for Education, 'Islam' for Religion, and 'Poorest' for WI Standard errors are placed in parentheses.

Variance Component for Two-Level Model

The random effect of the two levels model in Table 6 showed that the variance between the clusters in AGQ (XTMELOGIT) is more than variance obtained from other methods. Comparing variance of AGQ (XTMELOGIT) to variance of AGQ (GLLAMM), the standard error for XTMELOGIT is smaller than that of every other methods which means the estimate obtained using XTMELOGIT method is better than estimate of every other methods when only two level binary logistic regression is considered. The PQL had the smallest random intercept with largest standard error and it also have the largest $-2logL$ which minimized the reliability of the method.

Intra Cluster Correlation Coefficient

AGQ (XTMELOGIT) had the largest ICC (0.201) compared with other methods. This implied that 20% of the total variance was explained by the variance within the cluster. The PQL had an ICC of 5%.

6. Discussion

The multilevel estimation methods for logistic random and fixed effects were evaluated at four performance dimensions: numerical convergence, bias, computation time and model fitting. Numerical convergence was measured by the convergence rate. The information on performance of the estimators were under two different circumstances (that is, when considering the intercept only model and when including the explanatory variable collected from the lower level with random intercept for level two). Large cluster schemes were used in this study. The convergence rate was based on the iteration produced by the macro GENLINMIXED, GLLAMM and XTMELOGIT macro to confirm whether numerical convergence was reached or not. Output from the GENLINMIXED was obtained using PQL, standard available in SPSS version 25, the GLLAMM output was obtained using AGQ which made use of fifteen quadrature point and XTMELOGIT syntax allow estimation via Laplacian approximation (NAGQ) and AGQ. The AGQ (GLLAMM) syntax had the smallest standard error, $-2logl$, AIC and BIC when the three levels were considered while AGQ (XTMELOGIT) produced the smallest standard error, $-2logl$, AIC and BIC when two levels were considered.

Comparison between single level and multilevel models were made and we found that the effect of the primary predictor in the standard logistic regression model was underestimated in comparison with multilevel models. Some covariates were either overestimated or underestimated. This implied that the differences in the estimated β-coefficients from the multilevel models and standard model arose because of the inclusion of the random effects. Therefore, using single level model to predict the future value of modern contraceptive use in cluster survey is inappropriate. This is in agreement with findings of a previous study (Khan & Shaw (2011)). Previous literature on appropriateness of multilevel analysis of clustered data had same conclusions (Leckie & Goldstein (2015); Li *et al.* (2011); Sanago *et al.* (2012); Tendulkar *et al.* (2010))

Considering the two level binary logistic regression, NAGQ (XTMELOGIT), AGQ (XTMELOGIT) and AGQ (GLLAMM) produced the same fixed effect but different random effect in the analysis of the simulated data. However, only AGQ (XTMELOGIT) and AGQ (GLLAMM) with fifteen quadrature points had the same fixed effect in survey data. Furthermore, AGQ produced the smallest $-2logL$, AIC and BIC than other methods in both data. Nevertheless, convergence of PQL was the shortest and its estimate had lowest standard error in both simulated and survey data. This suggested that AGQ was the best method when only two level binary logistic regression is considered. Our finding was corroborated by conclusions of a 2011 study (Khan & Shaw (2011)). On the contrary, it was concluded in another earlier study that Quadrature methods performance was relatively poor compared to PQL, although his conclusion was based on small sample size (Lesaffre & Spiessens (2001)). However, in this research work, AGQ was the best multilevel model compared with PQL and NAGQ.

In the intercept only model for two-level binary logistic regression, there is consistency among the estimates produced by full ML using simulated data but outcomes from the survey data were less consistent. Also, the full likelihood method generated higher random effect than quasi likelihood in the simulated data but it was otherwise with survey data. Based on the result obtained for three level binary logistic regression, AGQ (GLLAMM) had the smallest $-2logL$, AIC and BIC which implies that GLLAMM had the best performance for three-level model. Our study has further demonstrated the tendency of the standard logistic model to seriously bias the parameter estimates of observed covariates when analyzing multilevel data. The differences between estimates obtained using standard logistic and PQL as well as between NAGQ and AGQ were minimal. This is consistent with earlier reports that in the more common case where variances in a multilevel logistic model do not exceed about 0.5, the PQL model can be expected to perform well in term of bias (Goldstein & Rasbash (1996); Maas & Hox (2005)). That is, PQL methods are likely to be adequate for producing nearly unbiased estimates. PQL was also preferred in term of bias in an earlier report (Rodriguez & Goldman (1995)).

AGQ using GLLAMM syntax. In this study geo-political zone cannot stand as a level because it was not randomly selected for the survey and its ICC was approximately zero in most of the procedures considered. Dependence was much higher within the clusters (Level 2) than among the zones (Level 3). However, Level effect investigation is very important in multilevel cluster survey analysis. Researchers, especially the public health researcher, should endeavor to investigate dependency among hierarchies in their data using random effect and Intra class correlation coefficient. All the MLE of modern contraceptive use under different procedures suggested about 10% usage in Nigeria. Interventions aimed at promoting the use of contraception among Nigerians should not only be implemented at the individual level but tailored to the community (that is, cluster) level, as interventions conceived without consideration for cluster context are likely to have limited impact.

7. Declarations

Ethics approval

Ethical approvals for the study was sought and obtained from the National Health Research Ethics Committee assigned number NHREC/01/01/007 as earlier documented (Federal Ministry of Health Nigeria, 2013).

Consent to participate

Written informed consent for participation in the study was obtained from all the participants aged 18-49 years and additional written consent were obtained from parents and guardians of participants who have not attained aged 18 year.

Consent to publish

All the authors agreed on the final version of the manuscript and consented that it should be published.

Availability of data and materials

All materials, data and methods used for this study are readily available and would be shared on demand from the Federal Ministry of Health, Abuja, Nigeria.

Authors' Contributions

AFF partook in study design, data extraction, data simulation, data analysis, results and discussion. BB conceived the study, participated in the study design, data extraction, data analysis, results and discussion. Both authors reviewed and agreed on the final version of the manuscript.

Acknowledgements

We appreciate faculty members of department for their valuable suggestions to improve this work. We acknowledge the Federal Ministry of Health, Nigeria for granting us access to the data used for this study. Also acknowledged is the Consortium for Advanced Research Training in Africa (CARTA) which offered technical support to the first author. CARTA has been funded by the Wellcome Trust (UK) (Grant No: 087547/Z/08/Z), the Department for International Development (DfID) under the Development Partnerships in Higher Education (DelPHE), the Carnegie Corporation of New York (Grant No: B 8606), the Ford Foundation (Grant No: 1100-0399), and the Bill and Melinda Gates Foundation (Grant No: 51228).

Disclosure

The authors report no conflicts of interest in this work.

References

Adebowale, S. A., Fagbamigbe, A. F., & Bamgboye, E. A. (2011). Contraceptive use: implication for completed fertility, parity progression and maternal nutritional status in Nigeria. *African Journal of Reproductive Health*, 15(4), 60–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22571107

Adekunle, A. O., Arowojolu, A. O., Adedimeji, A. A., & Okunlola, M. A. (2000). Emergency Contraception: Survey of knowledge, Attitudes and Practice of Healthcare Professionals in Ibadan. *Journal of Obstetrics and Gynecology*, 20(3), 284–9.

Amin, S., Basu, A. M., & Stephenson, R. (2002). Spatial variation in contraceptive use in Bangladesh: Looking beyond the borders. *Demography*, 39, 251–267.

Andrews, M., Schank, T., & Upward, R. (2006). Practical fixed-effects estimation methods for the three-way error components model. *Stata Journal*, 6, 461–481.

Ankomah, A., Anyanti, J., Adebayo, S., & Giwa, A. (2013). Barriers to Contraceptive Use among Married Young Adult in Nigeria: A Qualitative Study. *International Journal of Tropical Disease and Health*, 3, 267–282.

Austin, P. C. (2005). Bias in Penalized Quasi-Likelihood Estimation in Random Effects Logistic Regression Models When the Random Effects are Not Normally Distributed. *Communications in Statistics: Simulation & Computation*, 34, 549–65.

Bingenheimer, J. B., & Raudenbush, S. W. (2004). Statistical and Substantive Inferences in Public Health: Issues in the Application of Multilevel Models. *Annual Review of Public Health*, 25, 53–77.

Blalock, H. M. (1984). Contextual-Effects Models: Theoretical and Methodological Issues. Annu. Rev. *Sociol*, 10, 353–72.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods* (1st ed.). Newbury Park,CA: Sage.

Bryk, A. S., & Raudenbush, S. W. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods Second Edition* (2nd ed.). Chicago, IL.

Daniel, J., & Sonya, K. (2011). Fitting Multilevel Models with Ordinal Outcomes: Performanceof Alternative Specifications and Methods of Estimation. *NIH Public Access Psychol Methods*, 16(4), 373–90.

DiezRouz, A. V., & Mair, C. (2011). Neighborhood and health. *Annals of the New York Academy of Science*, 1186, 125–45. https://doi.org/10.1111/j.1749-6632,2009.05333.x

Essien, E. J., Monjok, E., Smesny, A., & Ekabua, E. J. (2010). Contraceptive practices in Nigeria: Literature review and recommendation for future policy decisions. *Open Access Journal of Contraception*, 1, 1–14.
https://doi.org/10.2147/OAJC.S9281

Fagbamigbe, A. F., Adebowale, A. S., & Morhason-Bello, I. O. (2015). Survival analysis of time to uptake of modern contraceptives among sexually active women of reproductive age in Nigeria. *BMJ Open*, 5(12), 1–8. https://doi.org/10.1136/bmjopen-2015-008371

Fagbamigbe, A. F., Afolabi, R. F. & Idemudia, E. S. Demand and Unmet Needs of Contraception Among Sexually Active In-Union Women in Nigeria: Distribution, Associated Characteristics, Barriers, and Program Implications. *SAGE Open* (2018),8:215824401775402.

Federal Ministry of Health Nigeria. (2013). *National HIV/AIDS and Reproductive Health and Serological Survey, 2012 (NARHS Plus),*. Federal Ministry of Health, Abuja, Nigeria.

Goldstein, H. (2003). *Multilevel Statistical Models* (3rd ed.). Oxford: Oxford University Pres.

Goldstein, H., & Rasbash, J. (1996). Improved Approximations for Multilevel Models with Binary Responses. *Journal of Royal Statistical Society*, 159, 505–13.

Hox, J. J. (2010). *Multilevel Analysis Technique and Application: Quantitative Methodology Series*. (G. A. Marcoluides, Ed.) (2nd ed.). London: Lawrence Erlbaum Publishers.

Hox, J. J. (2002). *Multilevel Analysis: Techniques and Applications* (1st ed.). Mahwah, N.J.: Lawrence Erlbaum Publishers. Retrieved from joophox.net/publist/pubenjh.htm

John, J. D., DonnaL, C., Stephanie, T. L., & Runze, L. (2012). *Sensitivity and Specificity of Information Criteria* (1st ed.). Pennsylvania: The Methodology Centre and Department of Statistics, Pennsylvania State University,.

Kalam, E. M. N., & Khan, H. T. A. (2002). Modeling contraceptive prevalence in Bangladesh: A hierarchical approach. *Asian Meta Centre, Research Paper Series, 21*(5), 1–15.

Khan, M. H. R., & Shaw, H. E. J. (2011). Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data. *Journal of Data Science, 9*, 93–110.

Klotz, J. H., Milton, R. C., & Zacks, S. (1969). Mean Square Efficiency of Estimators of Variance Components. *Journal of the American Statistical Association, 64*, 1383–94.

Leckie, G., & Goldstein, H. (2015). A multilevel modelling approach to measuring changing patterns of ethnic composition and segregation among London secondary schools, 2001–2010. J. R. *Statist. Soc., 178*(2), 405–424.

Lesaffre, E., & Spiessens, B. (2001). The Effect of the Number of Quadrature Points in a Logistic Random-Effects Model: An Example. *Applied Statistics*, 50, 325–35.

Leyland, A. H., & Goldstein, H. (2001). *Multilevel Modelling of Health Statistics. Series: Wiley series in probability and statistics* (Eds). Chichester, UK: Wiley.

Li, B., Lingsma, H. F., Steyerberg, E. W., & Lesaffre, E. (2011). Logistic random effects regression models: a comparison of statistical packages for binary and ordinal outcomes. *BMC Medical Research Methodology, 11*(774), e1–e11.

Maas, C. J. M., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, 58, 127–37.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *European Journal of Research Methods for the Behavioral and Social Sciences* ,1, 85–91.

Mason, W., Wong, G., & Entwisle, B. (1983). Contextual Analysis Through the Multilevel Linear Model. In Sociological Methodology 1983–1984, ed. S Leinhardt, pp. 72–103., 72–103.

National Population Commission (Nigeria) and ICF International, (2014). *Nigeria Demographic and Health Survey 2013*. Abuja, Nigeria.

Oye-Adeniran, B. A., Adewole, I. F., & Umoh, A. V. (2004). Community-based survey of unwanted pregnancy in Southwestern Nigeria. *African Journal of Reproductive Health, 8(3)*, 103–115.

Portnoy, S. (1971). Formal Bayes Estimation with Application to a Random-Effects Model. *Annals of Mathematical Statistics*, 40, 1379–88.

Rodriguez, G., & Goldman, N. (1995). An Assessment of Estimation Procedures for Multilevel Models with Binary Responses. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 164(2)*, 339–55.

Sampson, R., Raudenbush, S., & Earls, F. (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science, 277*, 918–24.

Sanago, U. M., Wolfe, R., Forbes, A., & Reid, C. M. (2012). Hospital-level associations with 30-day patient mortality after cardiac surgery: a tutorial on the application and interpretation of marginal and multilevel logistic regression. *BMC Medical Research Methodology*, 12(28), 1–13.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models* (1st ed.). Boca Raton, Florida: Chapman & Hall/CRC.

Srikanthan, A., & Reid, R. L. (2008). Religious and Cultural influences on contraception. *Journal of Obstetric and Gynaecol*, 30(2), 129–37.

Tendulkar, S. A., Buka, S., Dunn, E. C., Subramanian, S. V., & Koenen, K. C. (2010). A multilevel investigation of neighborhood effect on parental warmth. *J.Community Psycol*, 38, 557–573.

United Nations. (2015). *Sustainable Development Goals (SDG)*. Washington, DC. Retrieved from http://www.un.org/sustainabledevelopment/sustainable-development-goals/

WHO. (2012). The Millenium Development Goals. Retrieved from http://www.who.int/mediacentre/factsheets/fs290/en/