Annals of Statistical Theory and Applications (ASTA)
Vol. 3, 126 - 141, May, 2020

Original Research Article

Weighted Logistic Regression Modelling of Prevalence and Associated Risk Factors of Malaria in Nigeria

E. E. Ukwajunor[1]; E. E. E. Akarawak[1]; I. O. Abiala[1]; S. B. Adebayo[2]

Received: 19th March 2018 Revised: 4th February 2020 Accepted: 12th March 2020

Abstract

In Nigeria, malaria remains a major public health problem. Assessing the various malaria risk factors have played key roles in policy formulation and implementation directed at preventing and controlling malaria in a population. This study examines the prevalence rate and the associated socio-economic, demographic, geographic factors, as well as, knowledge and attitude of malaria based on microscopy test. A weighted logistic regression model was fitted to the 2015 Nigeria Malaria Indication Survey (NMIS) data. The database for the analysis was based on data augmentation of child recode and mother recodes dataset which involves merging of relevant information of the children with their respective mothers. Blood samples were collected from children under five and tested for malaria using both microscopic and rapid diagnostic test kits. All analyses in this paper were based on malaria results from the microscopic test. The analysis shows that the prevalence of malaria was significantly associated with wealth index and mother educational level. Other significant covariates are age, use of bed net, place of residence, and region. Further, knowledge of malaria symptoms, causes, and treatment was found to be negatively associated with malaria prevalence. However, children whose mothers have correct knowledge of prevention were found to be more at risk of malaria compared with those without this knowledge. Malaria control strategies should be intensified, with more attention on rural arrears and the economically disadvantaged ones. Mother's educational enhancement and malaria awareness should be promoted in the community.

Keywords Malaria, Microscopic test, Nigeria, Odd ratio, weighted logistic regression.

1.0 Introduction

Malaria is a vector-borne infectious disease which poses a major economic burden and public health challenges, taking its greatest toll on children under five and pregnant woman (Centers for Disease Control and Prevention, 2019). Nigeria suffers the world's greatest malaria burden accounting for up to 25% of the global cases and death (WHO, 2018). The disease is responsible for approximately 60% of out-patient visits and 30% of admission. It is also believed to contribute to 11% of maternal mortality, 25% of infant mortality, 30% of under-five mortality (National Malaria Elimination Programme, 2016). The disease overburdened the already weakened health system and exerts severe social and economic burden on the nation, retarding the gross domestic product (GDP) by 40% annually and costing approximately 480 billion nairas in out-of-pocket treatment, prevention costs, and loss of man-hours (National Malaria Elimination Programme, 2016).

Effective control of the disease is key to national development and reduction in associated mortality. Assessing the various malaria risk factors and investigating their interaction at various levels have played key roles in designing and implementing comprehensive measures directed at preventing and controlling malaria in a population (Mutegeki et al., 2016). Statistical models are extensively used in the analysis of

[1]Department of Mathematics, University of Lagos, Lagos, Nigeria. E-mail: eunice.ukwajunor@gmail.com
[2]Planning Research and Statistics Directorate, National Agency for food and Drug Administration and Control, Abuja

infectious diseases (Kleinchmidt, et al., 2001, Holtz, et al., 2002, Zhou, et al., 2004, Chirombo, et al., 2014). The class of Generalized Linear Models (GLMs) for example, permits us to study patterns of systematic variation in disease in relation to associated covariates. These models provide a powerful tool for modeling the relationship between a response and predictor variables.

Literature on generalized linear models with applications to malaria in Nigeria is vast (Gayawan et al., 2014, Onyiri, 2015, Adigun et al., 2015, Dawaski et al., 2016, Tobin-west and Kanu, 2016). However, in all of the approaches, the effect of the survey design which is aimed at accounting for unobserved heterogeneity is rarely considered. Thus, this study aims at assessing socio-economic, demographic, and geographic factors, as well as knowledge and attitude about malaria on the prevalence of malaria in Nigeria, using weighted logistic regression for complex sample design.

The present work is organized as follows. Section 2 describes the method for data collection and statistical analysis. In section 3, we present the results and these are discussed in section 4 with concluding remarks of this work.

2.0 Methods

2.1 *Study Area*

This study is based on the 2015 Nigeria Malaria Indicator Survey (NMIS). Nigeria lies within Sub-Sahara Africa situated between latitude$4^0 16'$and $13^0 53'$ north and longitude$2^0 40'$ and $14^0 41'$ east. The country has the largest population in Africa and the seventh-largest in the world (National Population Commission, 2016). Nigerian has a tropical climate of wet and dry season driven by the movement of the two dominant winds; the rain-bearing southwesterly winds and the cold-dry-dusty northeasterly winds, usually referred to as the harmattan. Nigeria's climate conditions make it suitable for perennial malaria transmission. The most prevalent species of malaria parasites in Nigeria is *plasmodium falciparum*, and it is responsible for the most severe form of the disease (National Malaria Elimination Programme, 2016).

2.2 *Data Collection*

This study relies on the data from the 2015 Nigeria Malaria Indicator Survey. The survey collected data from women within the reproduction age 15 – 49 years and those with children under age 5 years. The database for all analyses in this study was based on data augmentation of both child recode and mother recode datasets from the 2015 NMIS. This augmentation involved merging of relevant information of the children with their respective mothers. A total of 5527 observations were considered in the study.

In addition to the interviews conducted with the respondents, the blood sample was collected from children age 6 – 59 months, after attaining informed consent from the child's parent or guardian. Two methods of biological data collection were involved: the Rapid Diagnostic Test (RDT) and the microscopic test. To use results from RDT, there is a need to access the reliability of RDT compared with microscopic. A previous study on the reliability of RDT using 2015 NMIS revealed that the discriminatory accuracy of RDT was not strong enough (Fagbamigbe, 2019). Hence, for accurate prevalence, the present study uses results from microscopy as the outcome variable.

The independent variables explored in this study are demographic characteristics of the respondents, household wealth index, ownership, and use of bed nets, dwelling sprayed in the last 12 months, age of the child, gender, educational level of the mother, place of residence, and region. The wealth index is a background characteristic used as an indicator of the socio-economic status of households. It is calculated using data on the household's ownership of durable goods, source of drinking water, sanitation facilities, and other characteristics that relate to the household's economic status (National Malaria Elimination Programme, 2016). The independent covariates also include knowledge and attitude about malaria which is measured in terms of knowledge of: symptoms, prevention, causes, and treatment.

2.3 *Statistical Model*

Data were analyzed by fitting a weighted logistic regression model. The population for the 2015 NMIS was broken into 333 clusters. In each cluster, 25 households were selected. The response of thei^{th} person in the j^{th} household and k^{th} cluster can be specified asy_{kji}, $(i = 1,2, \dots, N_{kj}; j = 1,2, \dots, M_k; k = 1,2, \dots, K)$. Let the response variable$y_{kji} = 1$ if a child tested positive of malaria from microscopy diagnosis test inj^{th}household within the k^{th}cluster, and 0 otherwise. Then the fitted weighted logistic model is given as

$$logit\{p(Y_{kji} = 1 \mid X_{kji})\} = ln\left\{\frac{p(Y_{kji}=1 \mid X_{kji})}{1-p(Y_{kji}=1 \mid X_{kji})}\right\} = X'_{kji}\,\beta \tag{1}$$

where X_{kji} is the vector of covariates and β is the vector of the regression coefficient.

To account for the complexities introduced by the sample design and adjustments to the weights, it is necessary to adjust the classical method for parameter estimation. The maximum likelihood solution incorporating the weights is generally known as pseudo or weighted maximum likelihood estimation (Robert et al., 1987; Lumley, 2010). The main idea of this method is to define a function that approximates the likelihood function of the sampled finite population with a likelihood function formed by the observed sample and the known samplings weight (Cassy et al., 2016). The approximated log-likelihood function is given by

$$l_p(\beta) = \sum_{k=1}^{K}\sum_{j=1}^{m_k}\sum_{i=1}^{n_{kj}}[w_{kji} \times y_{kji}] \times ln[\pi(x_{kji})] + [w_{kji} \times (1 - y_{kji})] \times ln[1 - \pi(x_{kji})] \tag{2}$$

Differentiating equation (2) with respect to the unknown regression coefficients and equating it to zero yields the vector of $p + 1$ score equations.

$$\frac{\delta}{\delta\beta} l_p\,(\beta) = X'\,W(y - \pi) = 0 \tag{3}$$

Where X is the $n \times (p + 1)$ matrix of covariate values, the term W is an $n \times n$ diagonal matrix containing the weights, y is the $n \times 1$ vector of observed outcomes and $\pi = (\pi(x_{111}, \dots, \pi(x_{km_k n_{kj}}))$ is the $n \times 1$ vector of logistic probabilities. Numerical search procedures that allow weight could be used to obtain the solutions to equation (3).

Under a complex sampling design, estimation of the standard errors of the parameter estimates is very complicated. The problem arises when obtaining the correct estimator of the covariance matrix of the estimator of the coefficients (Lumley, 2010). The covariance of a standard logistic regression is $(X'DX)^{-1}$ where $D = WV$ is an $n \times n$ diagonal matrix with general element $w_{kji} \times \pi(X_{kji})[1 - \hat{\pi}(X_{kji})]$.For complex sample design, the correct estimator is given by Lumley (2010) as

$$\widehat{Var}(\hat{\beta}) = (X'DX)^{-1}S(X'DX)^{-1} \tag{4}$$

where S is the pooled within-stratum estimator of the covariance matrix of the left-hand-side of equation (3) given by Homer and Lemeshow (2002) as

$$S = \sum_{k=0}^{K}(1 - f_k)\frac{m_k}{m_k - 1}\sum_{j=1}^{m_k}(z_{kj} - \bar{z}_k)(z_{kj} - \bar{z}_k)^{-1} \tag{5}$$

where$z_{kji} = w_{kji} \times \hat{p}(Y_{kji} = 1|X_{kji})[1 - \hat{p}(Y_{kji} = 1|X_{kji})]$, the sum for n_{kj} sample units in the j^{th} primary sampling unit in the K_{th} stratum is $z_{kj} = \sum_{j=1}^{n_{kj}} z_{kji}$ and their stratum specific mean as $\bar{z}_k = \frac{1}{m_k}\sum_{j=1}^{m_k} z_{kj}$. The correction factor is given by$(1 - f_k)$, where $f_k = \frac{m_k}{M_k}$(Homer and Lemeshow, 2002; Cassy et al., 2016).

There are various techniques for obtaining the covariance matrix under a model-based survey. The formula-based Taylor series approximation (linearization) is the most widely used method of variance estimation for complex surveys because it is in most available software.

3.0 Results

Data analyses were conducted using R package survey in which all the design features were accounted for explicitly using the SVY design function. To access the robustness of the fitted model, the model was fitted in stages and variables that were significant at 5% level of significance were retained. In the first step, a model that explores a possible association between the dependent variables and socio-economic and demographic characteristics was fitted. In the second stage, model 1 was extended to account for the influence of geographical locations. The analytical results for the second step of the model are shown in Table 3 in the Appendix.

For the third step, measures on knowledge about malaria were fitted to the model (see Table 4 in the Appendix). The complexity of the model was further increased by fitting all the covariates to the model at the fourth step (Table 5 in the Appendix). To handle this complex model fitting for a valid statistical decision, the function SVYglm in R statistical software which performs logistic regression for sample survey data was used to analyze the data.

Table 1 presents the descriptive analysis of the presence and absence of malaria based on household characteristics of the children. Overall, about 27% (1419/5273) of the children had malaria. Based on the wealth index, children from the richest households suffered the fewest episodes of malaria 5.4% (866/2284). The percentage of children with malaria increased with an increase in the child age. Among the children whose mothers had no education, about 37.9% (866/2284) had malaria, while the fewest episodes were recorded among children whose mothers attained higher education 4.6% (19/409). Among children from rural areas, a greater percentage had malaria 35% (1202/3435) compared with their counterparts in urban areas 11.8% (217/1838). The proportion of children, who had malaria, was higher in the Northwest region compared with other regions. The descriptive analysis of test results based on knowledge and attitude about malaria is summarized in Table 2 in the Appendix.

Results of the fixed effects covariates are presented in Table 3 which shows the odds ratios and corresponding 95% confidence intervals. Wealth index, use of bed net, age of the child, sex of the child, mother educational attainment, place of residence, and region were found to be significant main effects. The effect of these variables can be directly interpreted using the odds ratio (OR). Based on the result, children from the richest household were 90% (OR = 0.981, P-value < 0.0001) less likely to have had malaria compared with children from the poorest household. Children from households in the middle class of wealth quintiles were 50% (OR = 0.5031, p-value, 0.0001) less likely to suffer malaria, while the findings were not significant for children in the poorer wealth quintile.

With regards to the use of bed net the previous night, the survey shows that the odds of having malaria were significantly lower for children who slept under bed net compared with those who did not (OR = 0.7251, p-value = 0.0057). However, findings on dwelling sprayed in the last 12 months were found not to be significant.

Furthermore, the result revealed that the odds of having malaria increase significantly as the child's age increases. Findings on the sex of the child were negatively significant for female children compared with male children. Based on the level of education, the odds of a child having malaria reduces as the educational level of the mother increases. For example, children whose mothers attained higher educational levels were 53% (OR = 0.4718, p-value = 0.0292) to have had malaria compared with children whose mothers had no education. On the other hand, children whose mothers had primary and secondary education were respectively 20% (OR =0.8030, p-value = 0.0858) and 24% (OR = 0.7641, p-value = 0.0044) less likely to have suffered malaria compared with children whose mothers have no education.

Table 1: Frequency (percentage in parenthesis) distribution of socio-economic, demographic, and geographic factors of malaria across the two levels of microscopy test results for 5527 surveyed samples.

Factors	Factor Level	Microscopy		Total
		No	Yes	
Wealth	Poorest (ref)	539 (54.8%)	444 (45.2%)	983 (l8.6%)
Index	Poorer	677 (59.0%)	470 (41.0%)	1147 (21.8%)
	Middle	778 (4.9%)	278 (13.4%)	1056 (20.0%)
	Richer	906 (84.0%)	173 (16.0%)	1079 (20.5%)
	Richest	954 (94.6%)	54 (5.4%)	1008 (19.1%)
Bed net use	No (ref)	1123 (76.7%)	343 (23.3%)	1471 (28.0%)
	Yes	1404 (72.4%)	535 (27.6%)	1939 (36.9%)
	No net	813 (73.2%)	297 (26.8%)	1110 (21.1 %)
Spraying	Yes (ref)	68 (80.0%)	17 (20.0%)	85 (1.6%)
	No	3759 (73.0%)	1392 (27.0%)	5151 (97.7%)
	I don't know	27 (73.0%)	10 (27.0%)	37 (0.7%)
Age	0 – 6 (ref)	69 (90.8%)	7 (9.2%)	76 (1.4%)
	7 – 23	1363 (79.8%)	346 (20.2%)	1709 (32.4%)
	24 – 59	2422 (69.4%)	1066 (30.6%)	3488 (66.1 %)
Sex	Male (ref)	1923 (72.5%)	730 (27.5%)	2653 (50.3%)
	Female	1931(73.7%)	689 (26.3%)	2620 (49.7%)
Mother Educational Level	No Education (ref)	1418 (62.1%)	866 (37.9%)	2284(43.4%)
	Primary	716 (73.8%)	254 (26.2%)	970 (18.4%)
	Secondary	1323 (82.7%)	276 (17.3%)	1599 (0.304%)
	Higher	390 (95.4%)	19 (4.6%)	409 (7.8%)
Place of residence	Urban (ref)	1621 (88.2%)	217 (11.8%)	1838 (34.9%)
	Rural	2233 (65.0%)	1202 (35.0%)	3435 (65.1%)
Region	North Central (ref)	753 (72.4%)	287 (27.6%)	1040 (19.7%)
	North East	777 (73.1%)	286 (26.9%)	1063 (20.2%)
	North West	778 (59.0%)	540 (41.0%)	1318 (25.0%)
	South East	467 (86.5%)	73 (13.5%)	540 (10.2%)
	South South	52 (80.0%)	130 (20.0%)	651 (12.3%)
	South West	558 (84.4%)	I03 (15.6%)	661 (12.5%)

With regards to the place of residence, children who live in rural areas are 87 percent more likely to have malaria compared with their counterparts who live in urban areas. Also, based on region, children from North West were 57% (OR = 1.5679, p-value = 0.0024) more likely to suffer malaria compared with children from north-central. Those from the southeast were twice (OR = 2.1811, p-value = 0.0052) more likely to have malaria compared with children from north-central, while the results were not significant for other regions. The odds ratios plot for the household characteristics of the children is shown in Figure 1.

Figure 1 provides plots of odds ratios of socio-economic, demographic, and geographic factors on microscopy test results.

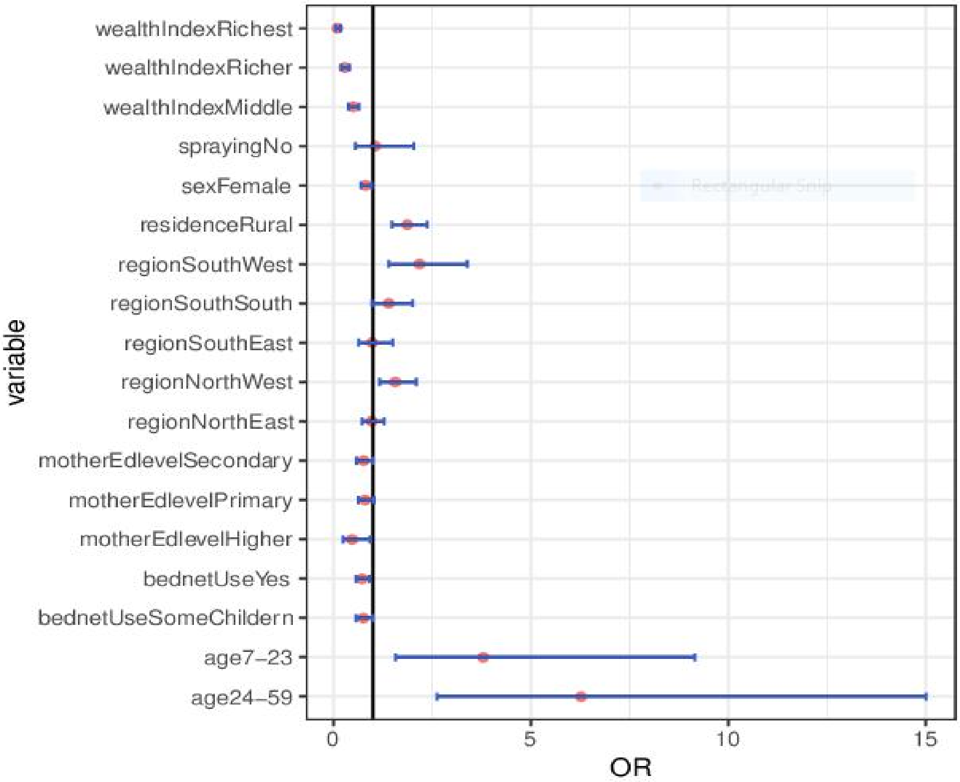

Figure 1: Plot of the odds ratios with error bars for fixed-effects covariates derived from 2015 NMIS.

A summary of the odds ratio of knowledge about malaria is presented in Tables 4. Based on knowledge of symptoms of malaria, it was found that children whose mothers recognized headaches as symptoms of malaria were 20% (OR = 0.7983, P-value = 0.0201) less likely to suffer from malaria compared with children whose mothers lack this knowledge. As regards knowledge of causes of malaria, the prevalence of malaria was significantly lower (OR = 0.7099, P-value = 0.0257) for children whose mothers know mosquitoes to be the cause of malaria than those without this knowledge.

Further, the finding revealed that mothers with correct knowledge of prevention, have a prevalence of malaria to be 29% (OR = 1.292, P-value = 0.0203) higher in their children. A similar result was found with children whose mothers recognized mosquito coil as a preventive measure. However, the odds of having malaria were significantly less likely for children whose mothers used home insecticide spray and keep the surroundings clean. In addition, based on medication for children, artesunat, ACT, and other anti-malaria (unknown component) were negatively associated with the microscopy test result. The result also revealed that mothers who recognized aspirin/panadol/paracetamol as malaria medication for children, have their children be 35% (OR = 1.350, P-value = 0.0397) more likely to suffer from malaria.

An estimate for all the covariates combined is presented in Table 5. Only variables with coefficient significance at 5% level were retained in the model. Wealth index, mother's highest educational level, sex of the child, use of bed net, place of residence, and region were significantly associated with the microscopic test. Also, knowledge of causes of malaria and malaria mediation for children (ACT. Anti-malaria) were found to be significant main effects.

4.0 Discussion

This study was designed to assess the socio-economic, demographic, and geographic factors, as well as knowledge of malaria on the prevalence of malaria among children under 5 in Nigeria. It is essential to explore knowledge and attitude about malaria along with other factors because several research studies have shown that high knowledge about malaria among the community (or household) enables preventive practice and control strategies (Ahmend et al., 2009).

The presence and absence of malaria were considered a binary response variable and a weighted logistic regression model, capable of incorporating the complexities introduced by the survey/sample designs were considered appropriate. Analysis of the model given by equation (1) shown that socio-economic, demographic, and geographic factors and knowledge of malaria have an impact on childhood health. Therefore, the results can serve as a guiding tool for policy formulation and execution.

From this study, the wealth index was found to be a significant main effect. Children from the richest households were 90% less likely to suffer from malaria compared with children from the poorest household. The result is understandable because households living in the lowest wealth quintile will be economically disadvantaged to adopt safer behavior that can lead to reducing episodes of malaria. Belonging to households in the poorest wealth quintile may be associated with poor environmental conditions which may lead to the breeding of mosquitoes which is a major cause of malaria. There are a lot of other challenges associated with living in the households in the poorest wealth quintiles. These negatively affect the socio-economic and well-being of individuals in this quintile.

A higher level of mother education has been associated with improved knowledge and practice with regard to strategies for malaria prevention and control. Children whose mothers attained higher education were found to be less likely to suffer from malaria compared with children whose mothers have little or no education. This finding is similar to those reported by (Gayawan et al., 2014; Ugwu and Zewotir, 2018). Therefore, control strategies should incorporate mother's education enhancement and malaria awareness.

Place of residence was found to play an important role in the spread of malaria among under-five children in Nigeria. Under 5 children who live in rural areas are 85% more likely to have malaria compared with their counterparts in an urban area. This finding is consistent with other studies conducted in Nigeria (Oyewale, 2018; Salwa et al., 2016; Adigun et al., 2015; Carrington, 2001). The high prevalence of the disease may be attributed to poor housing conditions and lack of good sanitation which have implications for malaria transmission and epidemiology. Living in urban areas is associated with better conditions of living which can prevent the breeding of mosquitoes and reduce transmission and epidemiology. For example, the housing in the urban area particularly has structural features that reduce keeping of stagnant water and consequently reduce contact with mosquito vector. Whereas, mud wall houses and grass-thatched roofs mostly found in the rural area are significantly associated with adult mosquito abundance (Zhou et al., 2007).

Findings on knowledge about malaria revealed that mother who recognized headaches as symptoms of malaria have their children 20% less likely to have malaria compared with children whose mothers lack this knowledge. As regard knowledge of prevention, it was worrisome to see that mothers with correct knowledge of prevention have a prevalence of malaria in their children to be 29% higher than those without this knowledge. However, this buttresses the fact that knowledge does not always translate to the adoption of safer bahaviour.

5.0 Conclusion

This study explores the prevalence and associated risk factors of malaria in Nigeria. A weighted logistics regression model which permits the incorporation of a complex sample design to account for any unobserved heterogeneity was fitted to the 2015 NMIS data.

The findings from this study provide insight into socio-economic, mother educational level, and knowledge and attitude about malaria. The mother's educational attainment has been found to influence her children's vulnerability to malaria infections. Children whose mothers attain higher education are less likely to suffer from malaria compared with children whose mothers have little or no education. Similarly, mothers (or caregivers/ guardians) with correct knowledge of malaria symptoms, causes, and treatment have a prevalence of malaria in their children to be lower than those without this knowledge. Hence, malaria control strategies should incorporate mothers' education enhancement and malaria awareness.

In addition, the prevalence of malaria among than five children across regions in Nigeria was found to be significantly higher in the northern part (North West) and southern part (southwest) region. While the southern region is prone to heavy rainfall that could facilitate the breeding of mosquitoes, the higher prevalence of malaria in the northern region could be attributed to several factors, one of which is the high temperature in the northern part of Nigeria. The spread of malaria needs conditions favourable to the survival of the mosquito and the plasmodium parasite. The development of the malaria parasite inside the mosquito is more rapid as the temperature rises (Okuneye and Gumel, 2017; Macdonald, 1957). Thus excessive temperature greatly impacts mosquito proliferation and survival. Therefore, strategies that could reduce people's contact with mosquitoes (use of bed net, good sanitation) should be promoted in the country. This can be achieved through publicity and education.

References

Adigun, A. B., Gajere, E. N. & Oresanya, O. Vounatsou, P (2015). Malaria risk in Nigeria: Bayesian geostatistical modeling of 2010 malaria indicator survey data. *Malaria Journal*, 14:156.

Ahmed, S. M. Haque, R. & Hossain, A. (2009). Knowledge on the transmission, prevention and treatment of malaria among two endemic populations of Bangladesh and their health-seeking behavior. *Malaria Journal,* 8: 1.

Beguine A., Hales S, Rocklov J, Astrom C, Louis V. R. & Sauerborn R (2011). The opposing effect of climate change and socio-economic development on the global distribution of malaria. *Global Environmental Change, Elsevier,* 21: 1209-1214.

Carrington, A. (2001). Malaria: Its Human Impact, Challenges and Control strategies in Nigeria. *HARVARD health policy review,* 2(2).

Cassy, S. R., Natario, I &Martins, M. R. (2016). Logistics Regression Modeling for Complex Survey Data with Application for Bed Net use in Mozambique, *Open Journal of Statistics*, 6: 898 – 907.

Centers for Disease Control and Prevention (2019). Global Malaria Incidence.

Chirombo, J., Lowa R.& Kazamabe L (2014). Using Structured Additive Regression Models to Estimate Risk Factors of Malaria: Analysis of 2010 Malawi Malaria Indicator Survey Data. *Plos ONE* 9 (7): e101116.doi:101371/journal.pone.0101116.

Dawaski S., Al-Mekhlafi H. M., Ithoi I., Ibrahim J., Atroosh W. M., Abdulsalam A. M., Sandy H., Elyana F. N., Adamu A. U., Yelwa S. I., Ahmed A., Al-Areeqi M. A., Subramanian L. R.. Nasir N. A. & Lau Y. L. (2016). Is Nigeria winning the battle against malaria? Prevalence, risk factors and KAP assessment among Hausa communities in Kano State. *Malar J.* doi:10.1186/s/2936-016-1394. PMID: 27397040, PMCID: PCM 4938925

Fabgamigbe A. F. (2019). On the discriminatory and predictive accuracy of the RDT against the microscopy in the diagnosis of malaria among under-five children in Nigeria. *Malaria Journal*, 18:46, doi:10.1186/s12936-019-2678-1.

Gayawan E., Arogundade E. D. & Adebayo S. B. (2014). A Bayesian multinomial modeling of spatial pattern of co-morbidity of malaria and non-malaria febrile illness among young children in Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, doi:10.1093/trstmh/tru068.

Holtz T. M., Marum L. K., Mkandala C., Chzani N., Roberts J. M., Maches O. A.. Parise M. E. & Kachur S. P. (2002). Insecticide-treated bed net use, anemia and malaria parastaemia in Blantyre District, Malawi, *Tropical Medicine and International Health*, 7:220-30.

Homer D. W. & Lemeshow S. (2002). *Applied Logistic Regression*, 2nd Ed, Wiley series in Probability and Statistics, John Wiley and Sons, Inc.

Kleinchmidt T., Sharp B. L., Clarke G. P. Y., Curtis B & Fraser C. (2001). Use of Generalized Linear Mixed Models in Spatial Analysis of Small-Area Malaria Incidence Rates in Kwazulu Natal, South Africa, *American Journal of Epidemiology,* 153 (12).

Lumley T (2010). *Complex Surveys: A Guide to Analysis using* R. Wiley series in survey methodology.

Macdonald G. (1957). *The Epidemiology and Control of malaria*. Oxford University Press, London, UK.

Mutegeki E., Chimbari M. J. & Mukaratirwa S. (2016). Assessment of individual and Household Malaria Risk factors among women in South African village. *Ac Trop*, Elsevier 4140.

National Malaria Elimination Programme (NMEP), National Population Commission (NPopC), National Bureau of Statistics (NBS) & ICF International (2016). *Nigeria Malaria Indicator Survey 2015*. Abuja, Nigeria, and Rockville, Maryland, USA: NMEP, NPopC and ICF International.

Onyiri N. (2015). Estimating malaria burden in Nigeria: a geostatistical modeling approach. *Geospatial Health*; 10 (306).

National Population Commission (2016). Nigeria Population Projections by Age and sex from 2006 to 2007. *National Population Commission,* Abuja, Nigeria.

Okuneye K. & Gumel A. B. (2017). Analysis of a temperature and rainfall-depended model for malaria transmission dynamics. Mathematics Biosciences, 287: 72 – 92.

Oyewale M. M., Folusho M. B. & Adeniyi, F. F. (2018). Housing type and risk of malaria among under-five children in Nigeria: evidence from the malaria indicator survey. *Malaria Journal*, 17(311), available at: https://doi.org/10.1186/s12936-018-2463-6.

Robert G., Rao J. N. K. & Kumar S. (1987). Logistic Regression Analysis of Sample Survey Data. *Biometrical,* 74: 1-12.

Salwa D., Hesham M. A., Init I., Jamaiah I, Wahib M. A., Awatif M. A., Saadatu I. Y., Abdulhami D. A., Mona A. A. & Yee-Ling L (2016). Is Nigerian winning the battle against malaria? Prevalence, risk factors and KAP assessment among Hausa communities in Kano State, *Malaria Journal,* 15:351, available at https//doi.org/10.1186/s12936-016-1394-3.

Tobin-west, C. I., Kanu, E. N. (2016). Factors Influencing the use of Malaria Prevention Methods among Women of Reproduction Age in Peri-Urban Communities of Port Harcourt City, Nigeria. *Niger Posgrad. Med J*, 23: 6-11.

https//www.who.int/malaria/media/world-malaria-report-2018/en/.

Ugwu C. L. J. & ZeWotir T. T. (2018). Using mixed-effect logistic regression models for complex survey data on malaria rapid diagnostic test result. *Malar J*, 17(453)

Zhou G., Minakawa N, Githeko A. K. & Yan G. (2004). Association between climate variability and malaria epidemics in the East Africa highlands. *PNANS*, 101 (8): 2375 – 2380.

Zhou G., Munga S., Minakawa N., Githeko A. K., & Yan G. (2007). Spatial relationship between adult malaria vector abundance and environmental factors in Western Kenya highlands. *A M J Trop. Med. Hyg.* 77:29 – 35.Association between climates

Appendix

Table 2: Frequency (percentage in parenthesis) distribution of knowledge and attitude about malaria across the two level of microscopy test result for 5273 surveyed samples.

Factors (Knowledge of malaria symptoms)	Factor Level	Microscopy No	Microscopy Yes	Total
Fever	No (ref)	1502 (70.7%)	621 (29.3%)	2123 (40.3%)
	Yes	2352 (74.7%)	798 (25.3%)	3150 (59.7%)
Chill	No (ref)	2601 (71.9%)	1018 (28.1%)	3619 (68.6%)
	Yes	1253 (75.8%)	401 (24.2%)	1654 (31.4%)
Headache	No (ref)	2132 (69.9%)	916 (30.1%)	3048 (57.8%)
	Yes	1722 (77.4%)	503 (22.6%)	2225 (42.5%)
Joint pain	No (ref)	2840 (72.4%)	I082 (27.6%)	3922 (74.4%)
	Yes	1014 (75.1%)	337 (24.9%)	1351 (25.6%)
Poor appetite	No (ref)	3204 (72.1%)	1239 (27.9%)	4443 (84.8%)
	Yes	650 (78.3%)	180 (21.7%)	830 (15.7%)
Vomiting	No (ref)	3388 (72.9%)	1258 (27.1%)	4646 (88.1%)
	Yes	466 (74.3%)	161 (25.7%)	627(11.9%)
Convulsion	No (ref)	3775 (73.0%)	1395 (27.0%)	5169 (98%)
	Yes	79 (76.0%)	25 (24.0%)	104(2.0%)
Cough	No (ref)	3673 (72.8%)	1372 (27.2%)	5045(95.7%)
	Yes	181 (79.4%)	47 (20.6%)	228 (4.3%)
Nasal congestion	No (ref)	3726 (72.8%)	1389 (27.2%)	5115 (97.0%)
	Yes	128 (81.0%)	30 (19.0%)	158 (3.0%)
Other systems	No (ref)	3577 (72.9%)	1331 (27.1%)	4908 (93.1%)
	Yes	277 (75.9%)	88 (24.1%)	365 (6.9%)
Knowledge of of malaria causes				
Mosquito	No (ref)	874 (68.4%)	403 (31.6%)	1277 (24.2%)
	Yes	2980 (74.6%)	1016 (25.4%)	3996 (75.8%)
Stagnant water	No (ref)	3181 (72.1%)	1230 (27.9%)	4411 (83.7%)
	Yes	673 (78.1%)	189 (21.9%)	862 (16.3%)
Dirty surroundings	No (ref)	2911 (71.5%)	1161 (28.5%)	4072 (77.2%)
	Yes	943 (78.5%)	258 (21.5%)	1201 (0.228%)
Cartan fued	No (ref)	3643 (72.4%)	1388 (27.6%)	5031 (95.4%)
	Yes	211 (87.2%)	31 (12%)	242 (4.6%)
Other causes	No (ref)	3764 (73.2%)	1376 (26.8%)	5140 (97.5%)
	Yes	90 (67.7%)	43 (32.3%)	133 (2.5%)
Knowledge of malaria prevention				
Mosquito Net	No (ref)	2112 (74.3%)	732 (25.7%)	2844 (53.9%)
	Yes	1742 (71.7%)	687 (28.3%)	2429 (46.1%)
Insecticide Treated Net (ITN)/Long Lasting Insecticide Net (LLIN)	No (ref)	2739 (71.0%)	1119 (29.0%)	3858 (73.2%)
	Yes	1115 (78.8%)	300 (21.2%)	1415 (26.8%)
Spray	No (ref)	3186 (71.8%)	1525 (28.2%)	4438 (84.2%)
	Yes	668 (80.0%)	167 (20.0%)	835 (15.8%)
Mosquito coil	No (ref)	3359 (73.2%)	1232 (26.8%)	4591 (87.1%)
	Yes	495 (72.6%)	187 (27.4%)	682 (12.9%)

Keep surroundings clean	No (ref)	2889 (70.8%)	1189 (29.2%)	4078 (77.3%)
	Yes	965 (80.8%)	230 (19.2%)	1195 (22.7%)
Other prevention	No (ref)	3717 (72.8%)	1390 (27.2%)	5107 (96.9%)
	Yes	137 (82.5%)	29 (17.5%)	166 (3.1%)
Know medication for malaria				
Sp/Fansider	No (ref)	3464 (72.6%)	1305 (27.4%)	4769 (90.4%)
	Yes	390 (77.4%)	114 (22.6%)	504 (9.6%)
Chloroquine	No (ref)	3274 (73.2%)	1196 (26.8%)	4470 (84.8%)
	Yes	580 (72.2%)	223 (27.8%)	803 (15.2%)
Artesunate	No (ref)	3412 (71.7%)	1347 (28.3%)	4759 (90.3%)
	Yes	442 (86.0%)	72 (14.0%)	514 (9.7%)
Quinine	No (ref)	3627 (72.9%)	1350 (27.1%)	4977 (94.4%)
	Yes	227 (76.7%)	69 (23.3%)	296 (5.6%)
Arthemeter Combination Therapy (ACT)	No (ref)	2889 (70.4%)	1217 (29.6%)	4106 (77.9%)
	Yes	965 (82.7%)	202 (17.3%)	1167 (22.1%)
Aspirin/Panadol/paracetamol	No (ref)	2889 (70.4%)	1217 (29.6%)	4106 (77.9%)
	Yes	965 (82.7%)	202 (17.3%)	1167 (22.1%)
Anti Malaria	No (ref)	3204 (72.2%)	1235 (27.8%)	4439 (84.2%)
	Yes	650 (77.9%)	184 (22.1%)	834 (15.8%)
Other treatments	No (ref)	3730 (73.0%)	1379 (27.0%)	5109 (96.9%)
	Yes	I24 (75.6%)	40 (24.4%)	164 (3.1%)

Table 3: Estimate of odds ratios of socio-economic, demographic and geographic factors on microscopy test. The Table shows the odds ratios, 95% confidence interval and p-values for socio-economic, demographic and geographic covariates combined. The parameters that have significance influence on the outcome variable are bolded.

Coefficients	Odd Ratio	95% C.I	Pr(>|t|)
Intercept	0.0798	0.0251 – 0.2540	< 0.0001
Wealth Index (ref. Poorest)			
Poorer	1.0063	0.8123 – 1.2466	0.9541
Middle	**0.5031**	**0.3891 – 0.6504**	**< 0.0001**
Richer	**0.2912**	**0.2094 – 0.4051**	**< 0.0001**
Richest	**0.0987**	**0.0518 – 0.1881**	**< 0.0001**
Use of bed net (ref. No)			
Yes	**0.7251**	**0.5773 – 0.9106**	**0.0057**
Some Children	**0.7601**	**0.5812 – 0.9940**	**0.0450**
Spraying (ref. Yes)			

No	1.0654	0.5572 ─ 2.0370	0.8481
Age (ref. 0 ─ 6)			
7 – 23	**3.7941**	**1.5716 ─ 9.1593**	**0.0030**
24 – 59	**0.2764**	**2.6236 ─ 15.0149**	**<0.0001**
Sex (ref. Male)			
Female	**0.8197**	**0.6946 – 0.9674**	**0.0187**
Mother Educational Level (ref. No Education)			
Primary	0.8030	0.6252 ─ 1.0313	0.0858
Secondary	**0.7641**	**0.5879 ─ 0.9929**	**0.0044**
Higher	**0.4718**	**0.2402 ─ 0.9268**	**0.0292**
Place of residence (ref. Urban)			
Rural	**1.8726**	**1.4806 ─ 2.3684**	**< 0.0001**
Region (ref. North Central)			
North East	0.9677	0.7312 ─ 1.2806	0.8181
North West	**1.5679**	**1.1734 – 2.0950**	**0.0024**
South East	0.9801	0.6393─1.5025	0.9264
South South	1.3992	0.9765 ─ 2.0051	0.0672
South West	**2.1811**	**1.4047 ─ 3..3869**	**0.0005**

Table 4: Estimates of odds ratios of knowledge about malaria on microscopy test. The Table shows the odd ratio, 95% confidence interval and p-values of knowledge about malaria.

Factors	Odd Ratio	95% C.I	Pr(>|t|)
Intercept	0.4486	0.3790 ─ 0.5309	< 0.0001
Knowledge of malaria symptoms			
Fever (ref. No)			
Yes	1.0105	0.8147 ─ 1.2532	0.9246
Chills/Shivering (ref. No)			
Yes	1.1779	0.9693 ─ 1.4314	0.0996
Headache (ref. No)			
Yes	0.7983	0.6602 ─ 0.9652	0.0201
Joint pain (ref. No)			
Yes	1.0280	0.8366 ─ 1.2632	0.7926

Poor appetite (ref. No)			
Yes	0.8819	0.6690 − 1.1624	0.3725
Vomiting (ref. No)			
Yes	1.1583	0.8975 − 1.4948	0.2588
Convulsion (ref. No)			
Yes	1.0170	0.5140 − 2.0122	0.9615
Cough (ref. No)			
Yes	1.0170	0.5140 − 2.0122	0.9015
Nasal congestion (ref. No)			
Yes	1.2417	0.7602 − 2.0283	0.3872
Other symptom (ref. No)			
Yes	2.1853	1.5872 − 3.0088	0.6318
Knowledge of malaria causes			
Mosquitoes (ref. No)			
Yes	0.7099	0.5254 − 1.9594	0.0258
Stagnant water (ref. No)			
Yes	1.0327	0.7902 − 1.3495	0.8138
Dirty surrounding (ref. No)			
Yes	0.8651	0.6723 − 1.1132	0.2601
Certain food (ref. No)			
Yes	0.4115	0.2510 − 0.6748	0.0004
Other cause (ref. No)			
Yes	1.5344	0.9161 − 2.5700	0.7424
Know how to prevent malaria			
Mosquito net (ref. No)			
Yes	1.2921	1.0407 − 1.6042	0.0203
ITN/LLIN (ref. No)			
Yes	1.0361	0.8001 − 1.3415	0.7882
Spray (ref. No)			
Yes	0.5889	0.4442 − 1.7807	0.0002
Mosquito coil (ref. No)			
Yes	1.5406	1.1532 − 2.0580	0.0035
Keep surrounding clean (ref. No)			
Yes	0.7450	0.5725 − 0.9694	0.0285
Eliminate stagnant water (ref. No)			
Yes	0.8812	0.6031 − 1.2874	0.5132
Other prevention (ref. No)			
Yes	0.4917	0.2988 − 0.8091	0.0052
Knowledge of malaria medication for children			
SP/Fansider (ref. No)			
Yes	0.8435	0.6057 − 1.1746	0.3137
Chloroquine (ref. No)			
Yes	1.0695	0.8153 − 1.4028	0.6276
Artesunate (ref. No)			
Yes	0.4164	0.2296 − 0.7552	0.0039
Quinine (ref. No)			
Yes	1.0413	0.6246 − 1.7360	0.8766
ACT (ref. No)			
Yes	0.4815	0.3504 − 0.6617	< 0.0001
Aspirin/Panadol/Paracetamol (ref. No)			
Yes	1.3520	1.0144 − 1.8021	0.0397

Antimalaria (ref. No)			
Yes	0.5654	0.4252 – 0.7519	< 0.0001
Other treatment (ref. No)			
Yes	0.7394	0.4554 – 1.2006	0.2222

Table 5: Estimate of odds ratios of socio-economic, demographic, geographic factors and knowledge about malaria on microscopy test. The table shows the odds ratios, 95% confidence interval and p-values for fixed-effects covariates derived from the 2015 Nigeria Malaria Indicator Survey. Only variables that were statistically significant at 5% level were retained in the model.

Coefficients	Odd Ratio	95% C.I	Pr(>|t|)
Intercept	0.0944	0.0282 – 0.3157	0.0001
Wealth Index (ref. Poorest)			
Poorer	1.0752	0.8626 – 1.3402	0.5188
Middle	0.5238	0.4026 – 0.6815	< 0.0001
Richer	0.3044	0.2166 – 0.4278	< 0.0001
Richest	0.1080	0.0586 – 0.1990	< 0.0001
Use of bed net (ref. No)			
Yes	0.7254	0.5759 – 0.9137	0.0064
Some children	0.7260	0.5515 – 0.9558	0.0225
Age (ref. 0 – 6)			
7 – 23	3.5896	1.4397 – 8.9497	0.0061
24 – 59	6.0844	2.4604 – 15.0462	< 0.0001
Sex (ref. Male)			
Female	0.8062	0.6832 – 0.9514	0.0108
Mother Educational level (ref. No Education)			
Primary	0.8080	0.6289 – 1.0382	0.0956
Secondary	0.7833	0.6028 – 1.0179	0.0677
Higher	0.4712	0.2448 – 0.9071	0.0244
Place of residence (ref. Urban)			
Rural	1.8608	1.4592 – 2.3729	< 0.0001

Region (ref. North Central)			
North East	0.9105	0.6787 – 1.2215	0.5318
North West	1.5229	1.1143 – 2.0813	0.0008
South East	0.9387	0.6186 – 1.4246	0.7664
South South	1.8628	0.5982 – 1.2444	0.4296
South West	1.8548	1.2078 –2.8483	0.0048
Causes: Certain food (ref. No)			
Yes	0.5230	0.3147 – 0.8691	0.0124
Other causes (ref.No)			
Yes	1.9729	1.1363 – 3.4251	0.0158
Medication : ACT (ref. No)			
Yes	0.5607	0.4259 – 0.7381	< 0.0001
Medication : Antimalaria (ref. No)			
Yes	0.6920	0.5122 – 0.9349	0.0165