



Bayesian Generalized Linear Mixed Modeling of Breast Cancer

**Ogunsakin ROPO EBENEZER, Siaka LOUGUE*

Department of Statistics, School of Mathematics, Statistics, and Computer Science, University of KwaZulu Natal, Durban, South Africa

***Corresponding Author:** Email: oreropo@gmail.com

(Received 10 Nov 2017; accepted 21 Feb 2018)

Abstract

Background: Breast cancer is one of the most common cancers among women. Breast cancer treatment strategies in Nigeria need urgent strengthening to reduce mortality rate because of the disease. This study aimed to determine the relationship between the ages at diagnosis and established the prognostic factors of modality of treatment given to breast cancer patient in Nigeria.

Methods: The data was collected for 247 women between years 2011-2015 who had breast cancer in two different hospitals in Ekiti State, Nigeria. Model estimation is based on Bayesian approach via Markov Chain Monte Carlo. A multilevel model based on generalized linear mixed model is used to estimate the random effect.

Results: The mean age of the patients (at the time of diagnosis) was 42.2 yr with 52% of the women aged between 35-49 yr. The results of the two approaches are almost similar but preference is given to Bayesian because the approach is more robust than the frequentist. Significant factors of treatment modality are age, educational level and breast cancer type.

Conclusion: Differences in socio-demographic factors such as educational level and age at diagnosis significantly influence the modality of breast cancer treatment in western Nigeria. The study suggests the use of Bayesian multilevel approach in analyzing breast cancer data for the practicality, flexibility and strength of the method.

Keywords: Bayesian; Breast cancer; Multilevel; Generalized linear mixed modeling; CODA/BOA

Introduction

Breast cancer refers to a malignant tumor developed from cells in the breast. Breast cancer is the most common cause of malignancy among women worldwide (1-3) and is a public health challenge among Nigeria women. Some years ago, breast cancer was not common in African countries especially Nigeria and was thought to be a leading cause of death in the developed countries (4). Thirty Nigerian women die of breast cancer every day in 2008 and this had risen to 40 women in 2012. Over 508000 of Nigerian women died in 2011 as a result of breast cancer (5-7, 2). In a study about prevalence of breast cancer in

western Nigeria found that breast cancer alone accounted for 37% of all the cancer cases present in western Nigeria (8).

With advancement in medical technology, advances have been made in the treatment of breast cancer among Nigeria women. Better treatment for breast cancer patients is difficult to define and older women are sometimes excluded from clinical treatment trials, probably because of their age (9). Since breast cancer biology differs from patient to patient with respect to factors like age, variation in response to treatment, and substantial competing risks of mortality (10-12), the ex-

clusion of some patients might not be valid. This implies that those aged women included in trials are probably not a true representation for the general older population (13, 14). Because of this, treatment of cancer in women with concise strategy is urgently needed. Therefore, there is need for modeling of breast cancer using evidence-based strategy technique.

Prognostic factors for breast cancer especially in Western Nigerian had been well studied; there is paucity of data on population-based research. In addition, few studies have used Bayesian to establish the prognostic factors associated with the disease. Hence, we sought to determine risk factors for treatment given to female breast cancer in western Nigeria using generalized linear mixed models. This research focuses on the justification for the use of the conventional surgery method in the treatment of cancer, based on data from two understudied hospitals in western part of Nigeria, one federally owned, and the other state-owned. Hence, we aimed to provide knowledge on risk factors for breast cancer treatment modality, using both the classical and Bayesian approach via generalized linear mixed model.

Materials and Methods

Ethics

Ethical clearance to conduct the study was sought from the Ethical Review Committee of the Federal Teaching Hospital, Nigeria.

Generalized Linear Mixed Model

Generalized linear mixed models (GLMMs) are an extension of linear mixed models to allow response variables from different distributions, such as binary responses. Alternatively, GLMMs is an extension of generalized linear models (e.g., logistic regression) to include both fixed and random effects (hence mixed models). The general form of the generalized linear mixed model is expressed:

$$E(\varphi_i | \tau_i, B, \mathcal{G}_k) = \tau_i' \lambda + B' \mathcal{G}_k + \delta \quad [1]$$

where φ is an $N_{obs} \times 1$ column vector, the outcome variable; τ_i' is a $N \times p$ matrix of the p pre-

dictor variables; λ is a $p \times 1$ column vector of the fixed-effects regression coefficients; B is an $N_{obs} \times N_r$ design matrix for the random effects (the random complement to the fixed τ); \mathcal{G}_k is a $q \times 1$ vector of the random effects (the random complement to the fixed λ); and is a $N \times 1$ column vector of the errors.

The present study deals with a specific case of generalized linear mixed model where the response variable is binary and two levels of analyses (hospitals) are considered which is the best model to apply when we have this type of situation. Only hospital is considered at two levels in the model, then the B component is removed from the model in equation (1). For a given observation i within a hospital k , the generalized linear mixed model excluding the error term becomes:

$$E(\varphi_i | \tau_i, \mathcal{G}_k) = \tau_i' \lambda + \mathcal{G}_k \quad [2]$$

where \mathcal{G}_k is the random effect representing the influence of hospital k on its within observation. Hence, the model utilized a logistic link function

$$E(\varphi_i) = \text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right)$$

$$\text{logit}(\pi_i) = \log\left(\frac{\text{pr}(\varphi_i = 1 | \tau_i, \mathcal{G}_k)}{1 - \text{pr}(\varphi_i = 1 | \tau_i, \mathcal{G}_k)}\right) \quad [3]$$

From equation (3), it implies that $\pi_i = E(\varphi_i = 1 | \tau_i, \mathcal{G}_k) = \text{pr}(\varphi_i = 1 | \tau_i, \mathcal{G}_k)$. Therefore, the final model is now expressed as:

$$\pi_i = \text{pr}(\varphi_i = 1 | \tau_i, \mathcal{G}_k) = \frac{\exp[\tau_i' \lambda + \mathcal{G}_k]}{1 + \exp[\tau_i' \lambda + \mathcal{G}_k]} \quad [4]$$

where τ_i represents the predictors for an individual i in an hospital k which is categorical. φ_i denote the response variable, λ is the vector of the fixed regression coefficients and \mathcal{G}_k is the random effect component (14). The estimated values of the regression parameters that maximize the probability of obtaining the observed data in classical statistics (15, 17) can be obtained by maximum likelihood. Likelihood of k independent measurements, given vectors of parameters θ

and explanatory (predictors) variables τ_i is represented as (15):

$$p(\varphi|\theta, \tau) = \prod_{i=1}^k p(\varphi|\theta, \tau) \quad [5]$$

In case of binary logistic regression where response (outcome) variable $\varphi_i = 1$ or 0, the likelihood function is estimated as (15):

$$p(\varphi|\theta, \tau) = \begin{cases} \text{logit}^{-1}(\omega_i) & \text{for } \varphi_i = 1 \\ 1 - \text{logit}^{-1}(\omega_i) & \text{for } \varphi_i = 0 \end{cases}$$

The above expression can be expressed as follows

$$\begin{aligned} p(\varphi|\theta, \tau) &= \prod_{i=1}^k (\text{logit}^{-1}(\omega_i))^{\varphi_i} (1 - \text{logit}^{-1}(\omega_i))^{1-\varphi_i} \\ &= \prod_{i=1}^k \pi_i^{\varphi_i} (1 - \pi_i)^{1-\varphi_i} \end{aligned}$$

The estimate of λ that maximizes the likelihood function can be computed using advanced calculus. This is achieved by taking the logarithm of the likelihood function:

$$\log(p(\varphi|\lambda, \vartheta_k, \tau)) = \sum_{i=1}^k [\varphi_i \log(\pi_i) + (1 - \varphi_i) \log(1 - \pi_i)] \quad [6]$$

Hence, the confidence interval of the parameters is then computed using Wald test and is represented as

$$\lambda_j \pm Z_{1-\frac{\alpha}{2}} SE(\lambda_j) \quad [7]$$

where λ_j is fixed effect coefficient associated with j th covariate, SE is the estimate of the standard error while $Z_{1-\frac{\alpha}{2}}$ is the critical value

for a standard normal distribution?

The classical statistics fit the logistic regression by means of an iterative procedure like maximum likelihood. In many situations, as a result of the assumptions underlying this iterative procedure, the estimation in classical statistics may result in non-convergence. These shortcomings as a result of non-convergence can be addressed using Bayesian inference as an alternative approach. Markov chain Monte Carlo algorithm approach can be used to provide a very general recipe for estimating properties of complicated distributions

in Bayesian statistics without any difficulties. There obviously remain, however, some challenges in concluding whether Bayesian methods perform better than classical statistics in modeling breast cancer data. In this paper, we compared the result of classical statistics and Bayesian method. As a requirement of the Bayesian approach, several diagnostics tests were performed to ensure convergence of the Markov chain Monte Carlo and the true reflection of the posterior distribution.

Bayesian Method

The methodology employed in this paper describes the Bayesian inference with emphasis on the posterior distribution, prior distribution as well as likelihood function for the Bayesian logistic regression. In the context of Bayesian statistics, data are regarded as fixed and unknown parameters as random variables. Considering a given parameter θ and a set of observed data, the interest of Bayesian technique lies on the probability of the parameter θ given the set of data π , i.e. $p(\theta|\pi)$. The main object of interest is the posterior distribution of the unknown parameters given the data. Mathematically, this can be written as

$$p(\theta|\pi) = \frac{p(\pi|\theta)p(\theta)}{p(\pi)} \propto p(\pi|\theta)p(\theta) \quad [8]$$

Equation 8 is known as the principle of Bayesian method. Hence, $p(\pi)$ is the prior distribution while the likelihood $p(\theta|\pi)$ is represented as:

$$p(\theta|\pi) = \prod_{i=1}^k [\pi_i]^{P_i} [1 - \pi_i]^{1-P_i}$$

$$\prod_{i=1}^k (\pi) = \frac{e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}{1 + e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}$$

While

$$1 - \prod_{i=1}^k (\pi) = \left[1 - \frac{e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}{1 + e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}} \right]$$

Hence, the likelihood function is expressed:

$$p(\theta|\pi) = \prod_{i=1}^k \left[\frac{e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}{1 + e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}} \right]^{P_i} \left[1 - \frac{e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}{1 + e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}} \right]^{1-P_i} \quad [10]$$

The preferred prior for logistic regression parameters is a multivariate normal distribution and is given as (15, 17, 18, 20):

$$\lambda_k \sim N(\varpi_k, \sigma_k^2) \quad [11]$$

Hence, we have:

$$p(\lambda_k) \propto \exp\left[-\frac{1}{2}(\lambda_k - \varpi_k)' \sigma_k^{-1} (\lambda_k - \varpi_k)\right]$$

From equation [8], replacing θ with λ_k result in $p(\lambda_k | \pi) \propto p(\pi | \lambda_k) p(\lambda_k)$. Therefore, the posterior distribution for logistic regression (fixed effects only) is represented as:

$$p(\lambda_k | \pi) \propto \left[\frac{e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}{1 + e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}} \right]^p \left[1 - \frac{e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}}{1 + e^{\lambda_0 + \lambda_1 \tau_1 + \dots + \lambda_k \tau_k}} \right]^{1-p} \times \exp\left[-\frac{1}{2}(\lambda_k - \varpi_k)' \sigma_k^{-1} (\lambda_k - \varpi_k)\right] \quad [12]$$

Posterior distribution is usually of high dimension and analytically intractable which sometimes required knowledge of powerful integration. In order to overcome this complex nature of posterior, we employed Markov chain Monte Carlo (MCMC) algorithm. MCMC technique is among of the technique employed to generates the estimates of unknown parameters θ and corrects the values generated in order to have a better estimate of the desired posterior distribution (21, 31).

All the analysis in this study was carried out using WinBUGS (21). There were 1500000 iterations, with the first 5000 discarded to cater for the burn-in period. We assessed MCMC convergence of all models parameters by checking Gelman Rubin plot (17, 18). The scale reduction factor, also known as Gelman-Rubin convergence diagnostic (39) was calculated for model parameter to assess convergence and adequate mixing of the chains. The test statistics for the Gelman Rubin diagnostic test can be estimated as

$$W = \frac{1}{M} \sum_{m=1}^M S_m^2$$

$$B = \frac{n}{M-1} \sum_{m=1}^M (\bar{\theta}_m - \bar{\theta})^2$$

$$\hat{v} = \frac{n-1}{n} W + \frac{1}{n} B$$

where k is the number of iterations of the chains.

$$\hat{R} = \frac{\hat{v}}{W}$$

Convergence is monitored when $\hat{R} \rightarrow 1$. \hat{R} is called the estimated potential scale reduction factor.

Results

Socio-demographic characteristics of participants

Out of 237 patients, 192 cases accounting for (81.01%) were malignant breast lesions, while 45 cases (18.99%) were benign giving a ratio of 4.3:1 for malignant to benign breast lesion. The mean age of the respondents was 42.2 ± 16.6 yr with 52% of the women aged between 35-49 yr. 93.67% of those who participated in this study were Christians while 6.33% were Muslims. Various prognostic factors are considered which include: intercept (λ_0), breast cancer types: malignant (λ_1), age: 35-49(λ_2), 50-69(λ_3),70+ (λ_4), level of education: at least high school (λ_5), religion: Christian (λ_6), tribe: Yoruba (λ_7), occupation: site of the cancer (λ_8), hospital (λ_9).

Table 1 gives a view of the sampled data with the following attributes: type of breast cancer, age, educational status, race, ethnicity, hospital and site of the breast cancer. The association between the variables of interest and the treatment of breast cancer patients received was calculated for the entire study (Table 1 model 3). Breast cancer (malignant) type was associated with the type of treatment given to patients, while other socio-economic factors were not significantly related, (model 3). The treatment given to breast cancer patients should be first based on the diagnosis of the breast cancer type. Hospital type was significant when breast cancer type was not included in the model (model 2). Hospital was significant but it's influence passes through the type of breast cancer a patients was diagnosed. We can then say that the hospital a patient attended after being diagnosed with breast cancer played an important role in their chance of survival. Model 1 in Table 1 shows that the age and the hospital type were both significant with re-

spect to the modality of treatment given to patient. The exclusion of type of breast cancer and educational status as a determinant for patients'

possibility of receiving surgical treatment makes age and hospital to be significant (model 1).

Table 1: Classical logit model for independent variables

Parameter	Model M_3		Model M_2		Model M_1	
	Lambda	Exact P-value	Lambda	Exact P-value	Lambda	Exact P-value
λ_0	-0.47[-5.83, 4.88]	1.000	3.96[1.67, 7.97]	0.011	3.95[1.66, 7.95]	1.000
λ_1	-	-	-	-	5.01[3.06, 8.80]	<.001
λ_2	-3.04[-8.17, 0.37]	0.118	-2.13[-6.12, 0.21]	0.088	-2.09[-5.95, 0.02]	0.040
λ_3	-2.33[-7.43, 1.03]	0.375	-1.54[-5.49, 0.75]	0.299	-1.49[-5.40, 0.69]	0.290
λ_4	-2.02[-7.09, 1.28]	0.528	-1.14[-5.05, 1.08]	0.546	-1.14[-5.05, 1.08]	0.546
λ_5	0.36[-0.89, 1.58]	0.698	0.09[-0.94, 1.11]	1.000	-	-
λ_6	-0.50[-2.27, 1.67]	0.817	-0.45[-1.86, 0.91]	0.661	-0.45[-1.86, 0.92]	0.666
λ_7	-0.72[-3.32, 2.66]	0.916	-1.24[-3.42, 0.84]	0.313	-1.25[-3.44, 0.85]	0.315
λ_8	0.39[-1.80, 0.82]	0.694	-0.48[-1.42, 0.43]	0.360	-0.47[-1.42, 0.43]	0.367
λ_9	0.48[-1.53, 0.60]	0.450	-1.99[-2.79, -1.24]	< 0,001	-1.98[-2.76, -1.26]	0.0001

Table 2 is the results for the model with Bayesian non-informative prior. From the 95% posterior intervals of λ_j , we observe that only the posterior distribution of the age and malignant are away from zero, indicating a significant effect of these

variables on the modality of treatment patients received from this part of Nigeria while other predictors considered in this study are not significant.

Table 2: Posterior Summary for Bayesian multilevel approach with non-informative

Variable	Mean	SD	MC error	2.5%	Median	97.50%
λ_0	-0.217	2.314	0.033	-4.183	-0.288	4.127
λ_1	-0.323	0.917	0.003	0.008	-0.367	1.611
λ_2	0.672	0.556	0.002	-0.389	0.660	1.791
λ_3	0.750	0.588	0.002	-0.365	0.736	1.949
λ_4	3.735	1.988	0.007	0.561	3.495	8.237
λ_5	0.680	1.348	0.007	-2.232	0.766	3.103
λ_6	-6.417	1.510	0.005	-10.010	-6.195	-4.123
λ_7	0.280	0.634	0.002	-0.894	0.255	1.590
$\tau.Hosp$	3.375	5.095	0.021	0.012	1.505	17.830

Assessing the performance of Markov Chain Monte Carlo (MCMC) chains in WinBUGS

The performance of a diagnostic test can be examined in several ways. The diagnostics examined in this paper are those of Geweke (15), Heidelberger and Welch (16), and Raftery and Lewis (17), which look at convergence of an individual chain, and that of Gelman and Rubin (18), which bases convergence on analysis of multiple chains. In order to assess whether

a chain has converged or not, we plot the sampled value against its number in the chain. Figure 1 displays the representation of the parameter behavior after 1,500,000 Monte Carlo repetitions. It was found that the kernel densities for shape and scale parameters exhibit approximately symmetric distribution. When the time series centered around a constant mean, it implies that the chain has reached conver-

gence as reported in Fig. 2 and have the chains converged to the same solution.

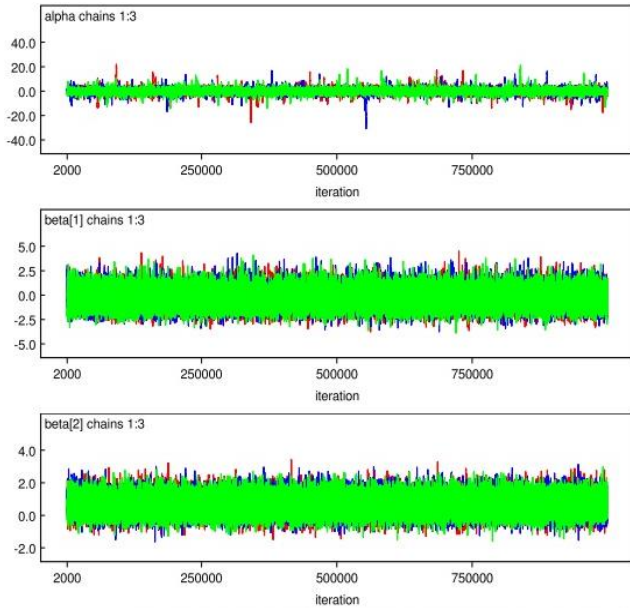


Fig. 1: Convergence history for major independent parameters

Discussion

Our results confirm findings from previous studies conducted in Nigeria on prognostic factors and breast cancer among women in Nigeria (1, 8, 2-13). Age could be a risk factor for breast cancer. Similar studies have been documented in other parts of Africa and the rest of the world (28, 29). Compared to the literature, the main contribution of the present study is the analysis of risk factors associated with modality of breast cancer treatment in western Nigeria in addition to the use of Bayesian approach. Previous studies that focus on the pattern of treatment given to breast cancer patients across time frame (years), age at diagnosis and educational categories are non-existent in western Nigeria. To the best of our knowledge, this could be the first study to address such scenario. For breast cancer treatment in Nigeria, the reason for surgery treatment on cancer patients is not clear (26). The present study showed prognosis factors for breast cancer treatment in western Nigeria.

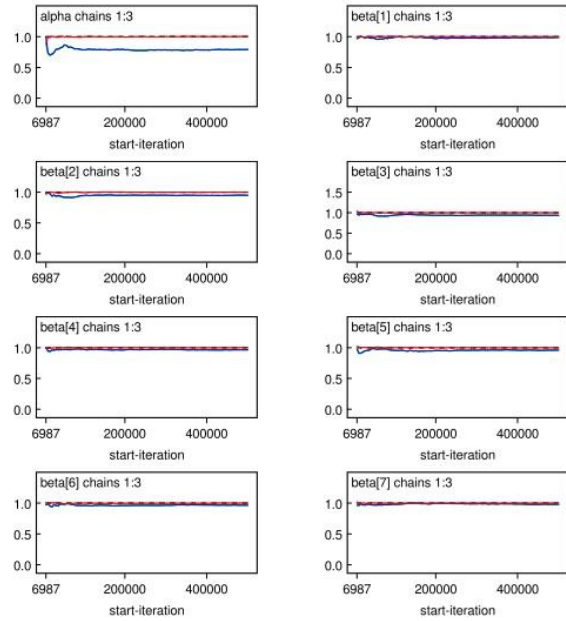


Fig. 2: Gelman-Rubin statistics for major independent parameters

In this study, patient with at least high school education is significant with treatment modality given to breast cancer, which is a new and unexpected finding in this part of Nigeria. From our analysis, the higher proportion of those treated with surgical treatment is found among educated and younger women. One of the possible explanations for these results may be that women with at least high school education presented their breast cancer cases to a medical practitioners than the less educated women. The result of Bayesian also reveals that those women with at least high school education are 2 times more likely to take surgery treatment than others. Other studies have shown similar result but not with treatment modality as investigated in the current finding (28-30, 32). We can infer from our findings that literacy level may influence treatment-seeking behavior. This finding could be used as an avenue to create awareness and advocacy campaigns that target less educated women in western Nigeria. Therefore, in the context of Nigeria, this is a good contribution in the area of breast cancer

modeling as previous studies only found age at diagnosis as the risk factor that influence treatment pattern.

However, our study did not observe the influence of education on breast cancer patients, a recent study by (33) observed the impact of educational level on breast cancer risk using Swedish data. Women with higher education were more likely to be diagnosed with in situ breast cancer. In addition, patients who had low level of education were at higher risk of late presentation of their breast cancer (34, 35). An observation attributed to low level of awareness and knowledge about breast cancer treatment and screening.

Another finding from this study is that age at diagnosis is associated with treatment modality. Majority of the previous studies have indicated that age is an important prognostic factor for breast cancer screening but not many studies have examined age at diagnosis as a factor that influences the modality of breast cancer treatment in Nigeria particularly in western region. Therefore, this study adds to the large body of research indicating that age at diagnosis influence the pattern of treatment modality for breast cancer patients in Nigeria. Considering the effects of age in the treatment modality of breast cancer patient, Bayesian findings highlight that patients aged 35-49 yr are 0.67 times more likely to receive surgical treatment than their counterparts. This suggests that younger patients in western Nigeria were more likely to be treated with surgery, compared to the older patients, an observation supported by previous studies (3, 29,30). Some studies attributed poor prognosis of breast cancer with age and the report indicated further that younger age is affected most due to an increase in invasive breast disease among this age group (3, 30, 32,36). Our result also shows that malignant breast lesion was a predictor of breast cancer treatment and it appeared to have higher distribution among those who had at least high school education, an observation that supports previous studies (3, 37-42) that malignant

breast lesion is predominant in this part of Nigeria.

Conclusion

We have demonstrated the socio-demographic factors associated with treatment modality of breast cancer in western Nigeria. Since understanding the risk factors associated is necessary for curbing the menace of breast cancer prevalence, these results provide a rationale for the need to create more awareness campaigns, strategies and sensitization that target less educated women to enhance patronization of breast cancer screening in Nigeria as a whole.

Ethical considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

This research was supported by the University of KwaZulu Natal, South Africa.

Conflict of interest

The authors declare that there is no conflict of interests.

References

1. Ebughe G, Ekanem I, Omoronyia O et al (2013). Age specific incidence of breast cancer in calabar, nigeria. *Int J Trop Dis*, 16:1-12
2. Chen H, Zhou M-q, Tian W et al (2016). Effect of age on breast cancer patient prognoses: A population-based study using the seer 18 database. *PLoS One*, 11(10):e0165409.
3. Ferlay J, Shin H-R, Bray F et al (2010). Estimates of worldwide burden of cancer in 2008: Globocan 2008. *Int J Cancer*, 127(12):2893-917.

4. Torre LA, Bray F, Siegel RL et al (2015). Global cancer statistics, 2012. *CA Cancer J Clin*, 65(2):87-108.
5. Akarolo-Anthony SN, Ogundiran TO, Adebamowo CA (2010). Emerging breast cancer epidemic: evidence from africa. *Breast Cancer Res*, 12 Supple 4:S8. doi: 10.1186/bcr2737.
6. Popoola AO, Omodele FO, Oludara MA et al (2013). Prevalence and pattern of cancers among adults attending a tertiary health institution in lagos, nigeria. *IOSR J Dental Med Sci*, 6(3):68–73.
7. Olugbenga AM, Olanrewaju MJ, Kayode OM (2012). Profile of cancer patients attending tertiary health institutions in southwestern nigeria. *Asian J Pharm Clin Res*, 5:34–7.
8. Townsley CA, Selby R, Siu LL (2005). Systematic review of barriers to the recruitment of older patients with cancer onto clinical trials. *J Clin Oncol*, 23(13):3112–3124.
9. Biganzoli L, Wildiers H, Oakman C et al (2012). Management of elderly patients with breast cancer: updated recommendations of the international society of geriatric oncology (siog) and european society of breast cancer specialists (eusoma). *Lancet Oncol*, 13(4):e148-60.
10. Mieog JSD, de Kruijf EM, Bastiaannet E et al (2012). Age determines the prognostic role of the cancer stem cell marker aldehyde dehydrogenase-1 in breast cancer. *BMC Cancer*, 12:42
11. Kiderlen M, Walsh PM, Bastiaannet E et al (2015). Treatment strategies and survival of older breast cancer patients—an international comparison between the netherlands and ireland. *PLoS One*, 10(2):e0118074.
12. Jolly TA (2014). External validity of a trial comprised of elderly patients with hormone receptor-positive breast cancer. *Breast Diseases*, 25(4):307–308.
13. Zare N, Haem E, Lankarani KB et al (2013). Breast cancer risk factors in a defined population: weighted logistic regression approach for rare events. *J Breast Cancer*, 16(2):214-9.
14. Gelman A, Hill J (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
15. Ntzoufras I (2011). *Bayesian modeling using WinBUGS*, volume 698. John Wiley & Sons.
16. Ngesa O, Mwambi H, Achia T (2014). Bayesian spatial semiparametric modeling of hiv variation in kenya. *PLoS One*, 9(7):e103299.
17. Gilks, Roberts G (1996). Strategies for improving mcmc. In, markov chain Monte Carlo in practice (gilks, wr, richardson, s., Spiegelhalter, dj eds).
18. Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000). Winbugs—a bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput*, 10:325-337.
19. Cleries R, Ribes J, Buxo M et al (2012). Bayesian approach to predicting cancer incidence for an area without cancer registration by using cancer incidence data from nearby areas. *Statist Med*, 31(10):978-87.
20. Gelman A, Carlin JB, Stern HS, Rubin DB (2014a). *Bayesian data analysis, volume 2*. Chapman & Hall/CRC Boca Raton, FL, USA.
21. Toft N, Innocent GT, Gettinby G, Reid SW (2007). Assessing the convergence of markov chain Monte Carlo methods: an example from evaluation of diagnostic tests in absence of a gold standard. *Prev Vet Med*, 79(2-4):244-56.
22. Odetunmbi O, Adejumo A and Sunday OO (2013). Loglinear modelling of cancer patients cases in nigeria: An exploratory study approach. *Loglinear modelling of cancer patients cases in Nigeria: An exploratory study approach*, pages 1–5.
23. Cluze C, Colonna M, Remontet L et al (2009). Analysis of the effect of age on the prognosis of breast cancer. *Breast Cancer Res Treat*, 117:121–129.
24. Ibrahim A, Salem M, Hassan R (2014). Outcome of young age at diagnosis of breast cancer in south egypt. *Gulf J Oncolog*, 1(15):76-83.
25. Alieldin NH, Abo-Elazm OM, Bilal D et al (2014). Age at diagnosis in women with non-metastatic breast cancer: Is it related to prognosis? *J Egypt Natl Canc Inst*, 26(1):23-30.
26. Brandt J, Garne JP, Tengrup I, Manjer J (2015). Age at diagnosis in relation to survival following breast cancer: a cohort study. *World J Surg Oncol*, 13:33.
27. Adebamowo CA, Ajayi O (2000). Breast cancer in nigeria. *West Afr J Med*, 19(3):179-91.

28. Ikpat O, Ndoma-Egba R, Collan Y (2002). Influence of age and prognosis of breast cancer in nigeria. *East Afr Med J*, 79(12):651-7.
29. Akinyemiju TF, Pisu M, Waterbor JW, Altekruze SF (2015). Socioeconomic status and incidence of breast cancer by hormone receptor subtype. *Springerplus*, 4:508.
30. Schonberg MA, Marcantonio ER, Li D et al (2010). Breast cancer among the oldest old: tumor characteristics, treatment choices, and survival. *J Clin Oncol*, 28(12):2038-45.
31. Jackman S (2000). Estimation and inference via bayesian simulation: An introduction to markov chain Monte Carlo. *Am J Pol Sci*, 44(2): 375-404.
32. Abudu E, Banjo A, Izegebu M et al (2007). Malignant breast lesions at olabisiobanjo university teaching hospital (oouth), sagamu-a histopathological review. *Niger Postgrad Med J*, 14(1):57-59.
33. Arora N, Simmons RM (2009). Malignant breast disease: Diagnosis and assessment. *General Surgery*, pages 1481-1494.
34. Yüksel S, Altun Uğraş G, Çavdar İ et al (2017). A risk assessment comparison of breast cancer and factors affected to risk perception of women in turkey: A cross-sectional study. *Iran J Public Health*, 46(3):308-317.
35. Oladimeji KE, Tsoka-Gwegweni JM, Igboekwe FC et al (2015). Knowledge and beliefs of breast self-examination and breast cancer among market women in ibadan, south west, nigeria. *PLoS One*, 10(11):e0140904.
36. Jebbin N, Adotey J (2004). Attitudes to, knowledge and practice of breast self-examination (BSE) in Port Harcourt. *Niger J Med*, 13(2):166-170.
37. Omaka-Amari LN, Ilo CI, Nwimo IO et al (2015). Demographic differences in the knowledge of breast cancer among women in ebonyi state, nigeria. *International Journal of Nursing, Midwife and Health Related Cases*, 1(3):18-27.
38. Ozturk M, Engin V, Kisioglu A, Yilmazer G (2000). Effects of education on knowledge and attitude of breast self-examination among 25+ years old women. *Eastern J Med*, 5(1):13-17.
39. Popoola AO, Wright KO, Igwilo AI et al (2013). Literacy and Breast Cancer Diagnosis and Treatment among Patients in a Tertiary Health Institution of Lagos Nigeria. *IOSR J Dental Med Sci*, 5(4):49-54.
40. Hussain SK, Altieri A, Sundquist J, Hemminki K (2008). Influence of education level on breast cancer risk and survival in sweden between 1990 and 2004. *Int J Cancer*, 122(1):165-9.
41. Ibrahim N, Oludara M (2012). Socio-demographic factors and reasons associated with delay in breast cancer presentation: a study in Nigerian women. *Breast*, 21(3):416-8.
42. Zare N, Khodarahmi S, Rezaianzadeh A (2015). The role of prognostic factors on the survival of breast cancer patients: Bayesian approach. *Iranian Journal of Epidemiology*, 11(3):23-33.