Nigerian Journal of Mathematics and Applications
Volume **36**, (2026), 63 – 80. *Printed by Unilorin press*
©*Nig. J. Math. Appl.* *http : //www.njmaman.com*

Monitoring CSM in Northern Nigeria: A Logistic and OLS Regression Approach

Sarah Oluwatosin Elakhe, Rofiyat Patience Braimah, Idemudia Waziri Edokpa and Odunayo Joseph Braimah*

Abstract

Cerebrospinal meningitis (CSM) remains a major public health challenge in sub-Saharan Africa. This study analyzed records of 390 suspected CSM patients (2020–2023) from Aminu Kano Teaching Hospital, Nigeria, using sex, age, and month as explanatory variables. Logistic regression and ordinary least square (OLS) models were fitted to identify epidemic patterns. Findings reveal that children under five are most vulnerable, with infections peaking during Harmattan months. Males are more prone than females. Logistic regression explained 4.3% of variation compared to 3.2% for OLS, indicating better predictive accuracy. Age was not significant, while sex and month showed significant negative effects. The study highlights key demographic and seasonal risk factors for CSM in Northwestern Nigeria.

1. Introduction and Preliminaries

A serious public health issue that still affects tropical nations, especially those in sub-Saharan Africa like Nigeria, is cerebralspinal meningitis (CSM). Large-scale epidemics in numerous nations within the African meningitis belt are attributed

Received: 30/03/2006, Accepted: 05/05/2026, Revised: 08/06/2006. * Corresponding author.
2015 *Mathematics Subject Classification.* 90B06 & 62Jxx.
Keywords and phrases. Meninges; Categorical; Odd Ratio; Cerebrospinal; Response; Explanatory.
Department of Mathematics and Statistics, Ambrose Alli University, Ekpoma, Edo State, Nigeria
E-mail of the corresponding author: braimahjosephodunayo@aauekpoma.edu.ng
ORCID of the corresponding author: https://orcid.org/0000-0001-7817-6338

to Group A and sporadically Group C [4]. In Nigeria, very little research has been done on this serious illness. Logistic regression has not been used in any of the scant literature to estimate the likelihood of developing cerebrospinal meningitis (CSM) subsequent to symptom onset. As seen in Figure 1 below, meningococcal meningitis, commonly referred to as cerebrospinal meningitis (CSM), is a bacterial form of meningitis, a dangerous meningeal infection that damages the brain membrane. If left untreated, it can result in 50% of fatal cases and cause serious brain damage [8].

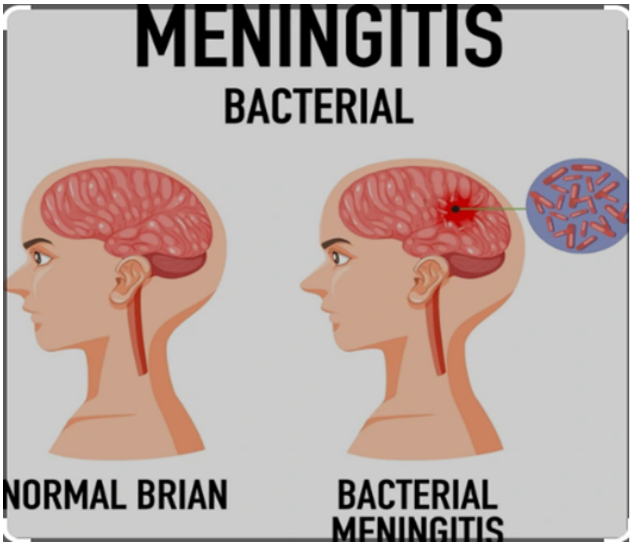


Figure 1: The Anatomy of a Normal and Meningitis Affected Bacterial Brain Source: (https://www.alamy.com/stock-photo/bacterial-meningitis-brain.html?sortBy=relevant).

Meningitis can be brought on by a variety of microorganisms. Large-scale epidemics could be brought on by Neisseria meningitidis. Six serogroups $(A, B, C, W, X$ and $Y)$ of the twelve serogroups of Neisseria meningitidis have been found to be capable of causing epidemics [3]. Droplets of respiratory or throat secretions from carriers are how the bacteria are spread from one person to another. The disease can spread more easily when people come into close contact with an infected individual (a carrier), such as when they kiss, sneeze, or cough on someone, or when they live in close quarters (like a dorm and share eating and drinking utensils). [9] states that the incubation period might vary from 2 to 10 days, with an average of 4 days. Certain environmental conditions, such as dry weather, dusty winds, high daytime temperatures, and low night time temperatures (particularly during the harmattan season from October to

March), enhance the virulence of these otherwise naturally occurring microorganisms, leading to occasional outbreaks of cerebrospinal meningitis. As a result of their increased susceptibility to respiratory diseases, people who live in crowded quarters, inadequately ventilated homes, and unhygienic environments in general promote the growth and survival of germs [1].

In Nigeria, there has been a nationwide outbreak of Cerebrospinal Meningitis (CSM), primarily impacting the northern states that are part of the African Meningitis Belt.

Table 1 displays the state of CSM and the measures implemented in Nigeria as of 1st October 2023, Figure 2 displays the distributions of CSM while Figure 3 depicts the distribution of trend of CSM death for 2020 to 2023 in Nigeria.

In Table 1 below,
A to represent LGAs reporting cases (n)
B LGAs in alert status (n)
C LGAs in epidemic status (n)
D Total cases reported
E Total Lab confirmed (n)
F MenC serotype (n)
G Deaths (n)
H Case fatality rate (%)

Table 1: Reported Cerebrospinal Meningitis Cases by State for 2022/2023 Season

Table 1 – continued from previous page

S/N	State	A	B	C	D	E	F	G	H (%)
1	Zamfara	14	2	12	3392	81	66	361	10.6
2	Sokoto	22	6	9	1610	32	28	95	5.9
3	Katsina	18	3	0	256	36	33	54	21.1
4	Kebbi	16	1	0	84	21	21	15	17.9
5	Niger	6	2	1	113	35	9	34	30.1
6	Yobe	12	1	1	36	7	0	19	52.8
7	Kano	22	0	0	35	3	0	8	22.9
8	FCT	5	0	0	15	0	0	12	80.0
9	Nassarawa	3	0	0	3	0	0	0	0.0
10	Gombe	2	0	0	2	1	0	1	50.0
11	Taraba	1	0	0	4	0	0	0	0.0
12	C/River	2	0	0	3	0	0	3	100.0
13	Osun	1	0	0	1	0	0	0	0.0
14	Lagos	1	0	0	3	0	0	2	66.7
15	Jigawa	13	0	0	5	2	0	3	60.0
16	Plateau	4	0	0	8	1	0	0	0.0
17	Kogi	1	0	0	2	1	0	0	0.0
18	Oyo	4	0	0	11	0	0	0	0.0
19	Delta	2	0	0	2	1	0	0	0.0
20	Adamawa	2	0	0	2	0	0	0	0.0
21	Benue	1	0	0	1	0	0	0	0.0
22	Kaduna	12	0	0	7	0	0	4	57.1
	Total	164	15	23	5595	221	157	611	10.4

Source: NCDC Newsletter as of 1st October 2023

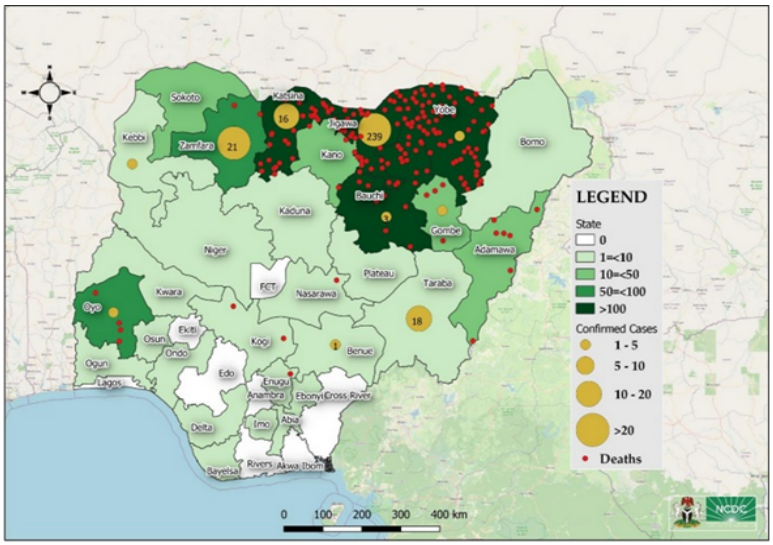


Figure 2: Map of Nigeria showing states with confirmed positive cases and deaths, week 40, 2022 - 39, 2023
Source: (Source: NCDC Newsletter as of 1st October 2023).

This study has set out to fill this gap and provide supportive evidence from Nigeria in the body of literatures and thereby help hospital to prepare better to handle the outburst of CSM in the nearest future. This study will help people in the region as well as health institutions services to reduce mortality and morbidity, thereby promoting economic development. This study will also provide support to show logistic regression as a better predictor than ordinary least square. In order to help hospitals better prepare for the impending outbreak of CSM, this study aims to fill this gap in the literature by including supporting data from Nigeria. This research will also assist the local population and healthcare facilities in lowering mortality and morbidity, which will boost the local economy. Additionally, this investigation will support the notion that logistic regression is a more accurate predictor than ordinary least squares.

1.1. **Methodology.** A method of evaluating situations where one or more independent variables influence a dependent variable (outcome) is called logistic regression. Most of the time, the dependent variable is either in categorical or ordinal order about one or more independent variables, or it is dichotomous (having just two possible outcomes, such as **yes** vs. **no**, **positive** vs. **negative**, **died** vs. **alive**, and so on). Finding the best fitting model to explain the relationship between a set of independent (explanatory) factors and the dichotomous feature of interest (response) is the aim of logistic regression. The coefficients (together with their standard errors and significance levels) of a formula to forecast a logit

transformation of the likelihood of the presence of the feature of interest are produced using logistic regression [2]. When there is a nonlinear relationship between the dependent variable and the independent variables (s), logistic regression is used as an extension of linear regression. Given certain values of X, the likelihood that $Y = 1$ and not 0 is predicted via logistic regression. This suggests that the likelihood of $Y = 1$ increases as the value of X increases if X and Y have a linear relationship. [5]. The odds ratio can be obtained using logistic regression when there are multiple explanatory variables. The process is similar to multiple linear regression, but it uses a binomial response variable. By examining the correlation between all variables collectively, the outcome is the effect of each variable on the odds ratio of the observed event of interest [6]. However, if dichotomous outcomes are employed, there are more significant restrictions on the application of these simple regression techniques. Specifically, as previously mentioned, these circumstances frequently result in violations of the heteroscadacity assumption, which is necessary for regression and states that the variance of the outcome is constant across all observed levels of the predictor. Regression techniques will also result in predictions for certain scenarios that contain erroneous numbers; in other words, the prediction must always indicate whether an outcome will occur (i.e., a zero or a one) [7]. Based on the provided explanation, the analysis of cerebrospinal meningitis (CMS) in some regions of Northern Nigeria was fitted with a logistic and ordinary least square (OLS) regression model in order to predict and monitor the outbreak of the epidemic.

2. Main Result

Models were considered in this section

2.1. **Logistic regression model.** Let Y be a random variable with two possible values. Y can be thought of as a column vector of n Bernoulli random variables y_i given a data collection with a total sample size of n and each observation being independent. Let

$$X_i = (x_{0i}, x_{1i}, \ldots, x_{ki})'$$

be a vector representing the ith subject's factors (explanatory variables), $i = 1, 2, \ldots, n$, where $x_{0i} = 1$. Suppose that y_i has a probability

$$\pi(x_i) = P(Y_i = 1 \mid X_i = x_i)$$

and takes on the value 1 with probability $\pi(x_i)$ and the value 0 with probability $1 - \pi(x_i)$. The response probability $\pi(x_i)$ in logistic regression is calculated as:

$$(1) \qquad \pi(x_i) = P(Y_i = 1 \mid X_i = x_i) = \frac{e^{\beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki}}} = \frac{\exp(\beta' X_i)}{1 + \exp(\beta' X_i)}$$

where the column vector $\beta = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)'$ represents the unknown regression coefficients. The following equation is therefore defined as the odds of success:

$$\frac{\pi(X_i)}{1-\pi(X_i)} = \exp\{\beta' X_i\} \tag{2}$$

The log-odds (logit) are then given by:

$$\log\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right) = \beta' X_i, \quad i = 1, 2, \ldots, n \tag{3}$$

Let Y be the response variable and suppose that it has C ordered categories with probability:

$$P(Y_i = j \mid X_i) = \pi^{(j)}(X_i), \quad j = 1, 2, \ldots, C \tag{4}$$

Therefore, in the multinomial logistic model, we have $(C-1)$ ratios as:

$$\frac{P(Y_i = j \mid X_i)}{P(Y_i = 1 \mid X_i)} = \frac{\pi^{(j)}(X_i)}{\pi^{(1)}(X_i)}, \quad j = 2, 3, \ldots, C;\ i = 1, 2, \ldots, n \tag{5}$$

In contrast to the multinomial logistic model, the $C-1$ cumulative probabilities will be regarded as follows:

$$\gamma^{(j)}(x_i) = P(Y_i \leq j \mid X_i) = \pi^{(1)}(X_i) + \cdots + \pi^{(j)}(X_i), \tag{6}$$

for all $j = 1, 2, \cdots, C-1;\ i = 1, 2, \cdots, n$.
Note that $\gamma^{(C)}(x_i) = P(Y_i \leq C \mid X_i) = 1$ and hence does not require modeling. The following is valid for $\gamma^{(j)}(x_i) = P(Y_i \leq j \mid X_i)$, for each subject $i = 1, 2, \ldots, n$ and for each category $j = 1, 2, \ldots, C-1$. Therefore,

$$\begin{aligned} &\log\left(\frac{\gamma^{(j)}(x_i)}{1-\gamma^{(j)}(x_i)}\right) = \log\left(\frac{P(Y_i \leq j \mid X_i)}{1-P(Y_i \leq j \mid X_i)}\right) \\ &= \alpha^{(j)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki}) = \alpha^{(j)} - X_i'\beta \end{aligned} \tag{7}$$

where $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ and $X_i = (X_{1i}, X_{2i}, \ldots, X_{ki})'$.
In other words, any potential cut-off of the response categories into two sets of "high" and "low" responses is taken into account by the ordinal logistic model. This is only significant if there is an ordering among Y's categories. For each of these cuts, a binary logistic model is then defined.

2.2. **Parameters of the model.** The model for the cumulative probabilities is:

$$\gamma^{(j)}(x_i) = P(Y_i \leq j \mid X_i) = \frac{\exp\left[\alpha^{(j)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}{1 + \exp\left[\alpha^{(j)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}$$
$$= \frac{\exp(\alpha^{(j)} - X_i'\beta)}{1 + \exp(\alpha^{(j)} - X_i'\beta)} \tag{8}$$

The intercepts $\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(C-1)}$ must satisfy the condition $\alpha^{(1)} \leq \alpha^{(2)} \leq \cdots \leq \alpha^{(C-1)}$ to guarantee that $\gamma^{(1)} \leq \gamma^{(2)} \leq \cdots \leq \gamma^{(C-1)}$. The parameters $\beta_1, \beta_2, \ldots, \beta_k$ are the same for each value of j.
The following is a representation of the model that modifies the proportional odds assumption:

$$\text{logit}[\gamma^{(j)}(x_i)] = \alpha^{(j)} - X_i'\beta^{(j)} \tag{9}$$

The probabilities for individual responses become:

$$P(Y_i = 1 \mid X_i) = \gamma^{(1)}(x_i)$$
$$= \frac{\exp\left[\alpha^{(1)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}{1 + \exp\left[\alpha^{(1)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]} \tag{10}$$

$$P(Y_i = j \mid X_i) = \gamma^{(j)}(x_i) - \gamma^{(j-1)}(x_i)$$
$$= \frac{\exp\left[\alpha^{(j)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}{1 + \exp\left[\alpha^{(j)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}$$
$$- \frac{\exp\left[\alpha^{(j-1)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}{1 + \exp\left[\alpha^{(j-1)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]} \tag{11}$$

for $j = 2, 3, \cdots, C-1$, and

$$P(Y_i = C \mid X_i) = 1 - \gamma^{(C-1)}(x_i)$$
$$= 1 - \frac{\exp\left[\alpha^{(C-1)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]}{1 + \exp\left[\alpha^{(C-1)} - (\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})\right]} \tag{12}$$

Note that the logit model's main function is to verify the proportionality assumption, that is, the idea that each category is perfectly homogeneous.

2.3. **Odds Ratios (OR).** The probability of an event happening in one group divided by its probability of happening in another is known as the *odds ratio*. When calculating the odds ratio in a cohort study, one can divide the odds of a risk factor among persons who experienced the event of interest by the odds of the same risk factor among those who did not. An estimated coefficient $\hat{\beta}$, with $\exp(\hat{\beta})$, is the odds ratio in binary logistic regression.
The odds ratio of the event $Y_i \leq j$ at x_1 relative to the same event at x_2 is:

$$OR = \frac{\gamma^{(j)}(x_1)/[1-\gamma^{(j)}(x_1)]}{\gamma^{(j)}(x_2)/[1-\gamma^{(j)}(x_2)]} = \frac{\exp(\alpha^{(j)} - X_1'\beta)}{\exp(\alpha^{(j)} - X_2'\beta)} = \exp\left[(X_2' - X_1')\beta\right] \tag{13}$$

which is independent of j. The distance between X_1' and X_2' is proportionate to the cumulative odds ratio.

2.4. **Goodness of Fit of the Model.** The degree to which a model adequately explains the response variable is referred to as its goodness of fit. Examining how closely the values predicted by the model match the observed values is one way to determine the quality of fit. The likelihood of the full model or saturated model (L_f) and the current model (L_c) can be compared.

2.5. **The Wald Test.** The likelihood ratio (LR), Lagrange multiplier (LM), and Wald tests comprise the "triumphal" of traditional likelihood testing methods. To determine the significance of individual logistic regression coefficients, the Wald test is frequently utilized. The Wald χ^2 statistic is calculated as:

$$Z_j^2 = \left(\frac{\hat{\beta}_j}{se(\hat{\beta}_j)}\right)^2, \quad j = 1, 2, \ldots, k \tag{7}$$

A Chi-square distribution with one degree of freedom is used to compare each Wald statistic.

3. Results and Findings

The Aminu Kano Teaching Hospital's (AKTH 2023) record office provided the data used in this study. The records office provided secondary data for the most recent epidemic year outbreak in the hospital (2021-2023). Data transformation is required to ensure age homogeneity. Records on 390 patients who were suspected of having signs of CSM were collected. Age, sex, and month were used as independent variables, and the data were appropriately presented and transformed into a categorical dependent variable (CSM). Therefore, the outbreak of cerebrospinal meningitis (CSM) has a logistic regression model:

$$\log\left(\frac{Y}{CSM^{+}}\right) = \frac{e^{\beta_0+\beta_1\, Sex+\beta_2\, Age+\beta_3\, Month}}{1+e^{\beta_0+\beta_1\, Sex+\beta_2\, Age+\beta_3\, Month}} \tag{8}$$

while the ordinary least regression model is:

$$P(Y \mid CSM^{+}) = \beta_0 + \beta_1\, Sex + \beta_2\, Age + \beta_3\, Month \tag{9}$$

The logistic and ordinary least square regression models were captured by using statistical software **R**, which made data analysis easier. The tables and figures below show the results of the analysis, which include descriptive statistics, results from logistic regression, and results from analysis of variance.

3.1. **Descriptive Statistics.** The descriptive statistics, as presented in Tables 2 to 5, demonstrate the categorization of CSM by sex, month, and age of infection. Positive infection is represented by a column labeled 1 while negative infection is represented by a column labeled 0.

Table 2: Classification of CSM by Sex of Infection

Sex	CSM Status 0 (-ve)	CSM Status 1 (+ve)	Grand Total
Male	95	100	195
Female	115	80	195
Grand Total	210	180	390

It is evident from Table 2's result that the same number of people (195) were found to have the CSM symptom. Following diagnosis, 80 females and 100 males had positive tests. This demonstrates that men are more prone than women to contract CSM. However, as Table 2 illustrates, the group with negative infection, comprising 210 individuals, is more than the total number of CSM infections, 180.

Table 3: Classification of CSM by Month of Infection

Similarly, the distribution of patients presenting Cerebrospinal Meningitis symptoms by contact month is displayed in Tables 3 and 4. The majority of symptoms (47), 26% of identified cases, can be inferred to have occurred in October. The winter (Harmattan) months of November (17%), December (15%), February (13%) and January (11%) follow next.

Month	CSM Status 0 (-ve)	CSM Status 1 (+ve)	Grand Total
January	23	20	43
February	19	24	43
March	5	17	22
April	8	15	23
October	90	47	137
November	50	30	80
December	15	27	42
Grand Total	210	180	390

Table 4: Classification of CSM by Age of Infection

Age in Years	CSM Status 0 (-ve)	CSM Status 1 (+ve)	Grand Total
0–20 Years	133	120	253
21–40 Years	30	24	54
41–60 Years	23	24	47
61–80 Years	19	9	28
81 Years and Above	5	3	8
Grand Total	210	180	390

Table 4 presents a classification of suspected cases of cerebrospinal meningitis according to age group. 253 suspected cases, 120 positive tests, and 133 negative tests have been reported for patients aged 0 to 20. We may infer that individuals under the age of eighty have the lowest number of cases of Cerebrospinal Meningitis (3%) while children and teenagers are more likely to contract the illness, with 120 (67%) of all cases testing positive. Hence, 24 years is the age at which CSM is majorly classified by infection.

Table 5: Classification of CSM by Age of Infection below 24 Years

Age in Years	CSM Status 0 (-ve)	CSM Status 1 (+ve)	Grand Total
0–5	67	82	149
6–11	47	15	62
12–17	13	15	28
18–23	9	13	22
Grand Total	136	125	261

We were concerned by Table 4's results, so we looked into which age groups; infants (0–5), kids (6–11), teens (12–17), and youths (18–23) are most likely to

develop Cerebrospinal Meningitis. The results in Table 5 reveal that children younger than six are more at danger. Therefore, additional research is needed to determine the best ways to care for, stop this infection, and better safeguard children under the age of six (i.e., 0 to 5 Years). The two states; CSM -ve and CSM +ve respectively are depicted graphically in Figures 10 and 11.

3.2. **Logistic Regression Result.** Using the logistic regression models in equation (8), the following outcomes were obtained, as indicated in Tables 6 through 13. The numbers 0 and 1 denote, respectively, the individuals who tested positively and negatively. To determine the significance of the model coefficient, the Wald Test, as stated in equation (7) is employed.

Table 6: Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-0.154	0.102	2.303	1	0.129	0.857

The first test result for the model, where the coefficients for each independent variable are zero, is shown in Table 6. The result suggests that the model is not significant ($p = 0.129 > 0.05$) at 5% level of significance.

Table 7: Inclusion of Independent Variables to the Equation

	Score	df	Sig.
Step 0 Variables Sex	4.127	1	0.042
Step 0 Variables Age	1.115	1	0.291
Step 0 Variables Month	6.260	1	0.012
Overall Statistics	12.700	3	0.005

If "age" were included in the model, Table 7 demonstrates that it would not be a significant predictor (p = 0.291). If month and sex are included, they are significant predictors (p = 0.012 and 0.042 respectively). This indicates that there is a significant difference between zero and the coefficients of sex and month in the equation. The Chi-square test, as indicated in Table 8, is used to validate the data fitness.

Table 8: Omnibus Tests of Model Coefficients

The step's justification is tested against the null hypothesis using the Chi-square goodness-of-fit test. Here, the transition is from the all-independents model to the constant-only model. In contrast, adding a variable or variables in this case

	Chi-square	df	Sig.
Step 1 Step	12.871	3	0.005
Step 1 Block	12.871	3	0.005
Step 1 Model	12.871	3	0.005

is justified if the step's significance is less than 0.05. Based on Table 8's results, the decision to use logistic model and inclusion of the independent variables is justified because, at a 95% confidence interval, the likelihood of 0.005 is less than 0.05.

3.3. **Result for the Logistics Regression Model.** Table 9 displays the output as expressed in the logistic regression fitting process in order to fit the logistic regression model.

Table 9: Values of the Regression Coefficients for the Independent Variables

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1a Sex	-0.438	0.209	4.402	1	0.036	0.646
Step 1a Age	-0.075	0.044	2.917	1	0.088	0.927
Step 1a Month	-0.067	0.026	6.886	1	0.009	0.935
Step 1a Constant	1.233	0.424	8.457	1	0.004	3.431

The fitted logistic model by using their corresponding value from equation (8) is:

$$\log\left(\frac{Y}{CSM}\right) = \frac{e^{1.233-0.438(Sex)-0.075(Age)-0.067(Month)}}{1+e^{1.233-0.438(Sex)-0.075(Age)-0.067(Month)}} \tag{17}$$

Each covariate and dummy independent in the model is tested for significance using the Wald statistic above and corresponding significance level. It is concluded that age has no significant effect on the CSM test outcome, while sex and month have significant effects.

Table 10: Cox-Snell R^2 and Nagelkerke R^2 for Model Summary

Step	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2
1	525.474a	0.032	0.043

In an effort to provide a logistic equivalent for R2 to OLS regression in Table 10, the Cox-Snell R2 and Nagelkerke R2 were obtained. In OLS, the Nagelkerke

measure and R2 both modify the Cox-Snell measure to vary between 0 and 1. From the above model summary, it is evident that month, age and sex together account for 4.3% of the variation in CSM diagnosis. Std. EE represents Standard Error of the Estimate in table 11 below.

Table 11: Regression Analysis Model Summary

Model	R	R Square	Adjusted R Square	Std. EE	Durbin-Watson
1	0.179a	0.032	0.025	0.493	1.567

Table 11's regression result indicates that the unit change in patients exhibiting symptoms of cerebrospinal meningitis to be tested positive or negative can be accounted for by month, sex and age ($R^2 = 0.032$). On the other hand, the logistic regression yields a unit change in the dependent variable of 4.3%. Thus, logistic regression appears to be a more accurate estimate for determining the presence of CSM meningitis.

In Table 12:
(i) Cons represents Constant
(ii) Unstand Coeff represents Unstandardized Coefficients;
(iii) Stand Coeff represents Standardized Coefficients;
(iv) Colli Statistics represents Collinearity Statistics.

Table 12: OLS Regression Coefficients

Model	Unstand Coeff		Stand Coeff	T	Sig.	Colli Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Cons)	0.812	0.106		7.676	0.000		
Sex	-0.105	0.050	-0.106	-2.095	0.037	0.983	1.018
Age	-0.039	0.024	-0.084	-1.649	0.100	0.967	1.034
Month	-0.016	0.006	-0.132	-2.615	0.009	0.979	1.021

Therefore, the fitted ordinary least square regression model is given as:

$$P(Y \mid CSM) = 0.812 - 0.105(\text{Sex}) - 0.039(\text{Age}) - 0.016(\text{Month}) \tag{18}$$

Our findings here also indicate that while age has no discernible effect on the test's outcome, sex and month have a substantial effect. The outcome of the logistic regression and that of OLS are consistent. However, compared to ordinary least square regression (3.2%), logistic regression (4.3%) yields greater percentage change in the estimates.

3.4. **Performance Measures for the Fitted Models.** For comparison's purpose, the fitted values for the two models are shown in Table 13 and Figure 3. The model with higher fitted values is said to be a better fit for such data.

In Table 13:
(i) LR represents Logistic Regression;
(ii) Dep represents Dependent
(iii) OLSR represents Ordinary Least Square Regression.

Table 13: Performance Measure of the fitted Logistic and Ordinary Least Square Regression

LR (Dep Variable CSM)	OLSR (Dep Variable CSM)
0.5369	0.5369
0.6424	0.6360
0.5035	0.4990
0.6111	0.6040
0.5535	0.5470
0.6577	0.5470
0.5125	0.5080
0.4042	0.4030
0.6269	0.6200
0.5202	0.5150

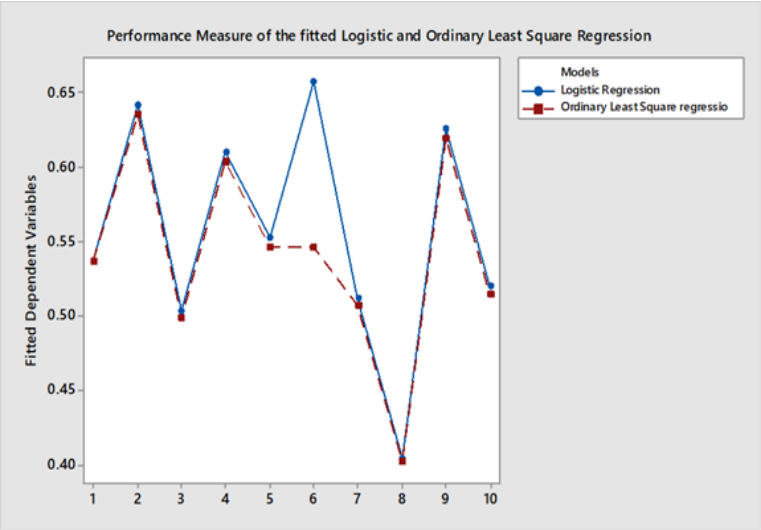


Figure 3: Trend Plot for Logistic and Ordinary Lest Square Regression Model

For comparison, Table 14 and Figure 3 show that the logistic regression model provides a better fit for CSM, as its dependent variable values are consistently higher than those of the ordinary least square regression model using the same predictors (age, sex, and month).

Conclusion. Considering the results of the OLS and the Logistic Regression, we can in general infer that the Logistic Regression is a better estimator of binary outcomes of this type of meningitis data set. Despite the fact that both predictions anticipate the same outcome, logistic regression accounts for 4.3% of the change in the dependent variable with the same data, whilst ordinary least squares account for 3.2% of the change. Further evidence that the logistic regression model is superior to the ordinary least squares regression model comes from the findings that the logistic regression's sum of squared errors is lower than the OLS regression's. Furthermore, the error analysis for both methods reveals that the logistic regression has a lower sum of squared error (83.66889) than the Ordinary Least Squares (OLS) regression (93.81681). This supports the logistic regression model's superior fitness over the Ordinary Least Squares regression model. The study therefore recommended that the outcomes of the logistic regression analysis point in specific paths for developing suggestions to address the Cerebrospinal Meningitis epidemic in the North Western Nigeria. Among the suggestions are:

1. men are more likely than women to have CSM. Males should therefore receive greater attention and education on how to prevent the illness;

2. the outcome unequivocally demonstrates that children under the age of six were most affected by the symptom and more frequently diagnosed as having CSM. In order to stop the sickness from spreading, vaccinations for children under the age of six should be made available, and schools and other public places should implement preventive measures; and

3. according to the research, there are specific months when it is imperative that families, health care providers, and governments should adopt preventive and precautionary actions to avoid CSM.

Acknowledgement: The authors are grateful to Ambrose Alli University, Ekpoma, Edo State, Nigeria for the supports they received during the compilation of this work.

Competing interests: The manuscript was read and approved by all the authors. They therefore declare that there is no conflicts of interest.

Funding: The Authors received no financial support for the research, authorship, and/or publication of this article.

References

[1] Bijlsma M. W., Brouwer M. C., Kasanmoentalib E. S., Kloek A. T., Lucas M. J. & Tanck M. W. (2016). Community-acquired bacterial meningitis in adults in the Netherlands, 2006-14: a prospective cohort study. Lancet Infect Dis. 16:339–47.

[2] Boateng E. Y., Bosson-Amedenu S., Nortey E. N. N. & Abaye D. A. (2019). Non-Medical Determinants of Caesarean Deliveries in Ghana: A Logistic Regression Approach. *Open Journal of Applied Sciences*. **9**, 492-505.

[3] Brouwer M. C. & Van de Beek D. Epidemiology of community-acquired bacterial meningitis. Curr Opin Infect Dis. (2018) 31:78–84. doi: 10.1097/QCO.0000000000000417

[4] Funk A., Uadiale K., Kamau C., Caugant D. A., Ango U. & Greig J. (2014) Sequential outbreaks due to a new strain of Neisseria meningitidis serogroup C in northern Nigeria, 21 – 37

[5] Giancristofaro R. A. & Salmaso L. (2007). Model Performance Analysis and Model Validation in Logistic Regression. *Statistica*. **63**, 375-396.

[6] Kim Y., Kwon S. & Song S. H. (2006). Multiclass Sparse Logistic Regression for Classification of Multiple Cancer Types Using Gene Expression Data. *Computational Statistics and Data Analysis*. **51**, 1643-1655.

[7] Sevvanthi K., Ganapathy S., Penumadu P. & Harichandrakumar K. T. (2023). Comparing the predictive performance of a decision tree with logistic regression for oral cavity cancer mortality: A retrospective study. Cancer Res Stat Treat; 6:103–10.

[8] Soeters H. M., Diallo A. O., Bicaba B. W., Kadadé G., Dembélé A. Y. & Acyl M. A. (2019) Bacterial Meningitis Epidemiology in Five Countries in the Meningitis Belt of Sub-Saharan Africa, 2015-2017. J Infect Dis. 220: pp. 165–S74. doi: 10.1093/infdis/jiz358

[9] Swartz, M. N. (2014) Bacterial Meningitis – A View of the Past 90 Years.The New England Journal of Medicine Boston: 351 (18) 18 – 26

APPENDIX

List of Age, Sex, Month (Season) and presence or absence of cerebrospinal meningitis (CSM) in Kano State University Teaching Hospital, Nigeria from October 2021–April 2022 and October 2022–April 2023.

TABLE 1. Age, Sex, Month (Season), and presence/absence of CSM cases recorded at Kano State University Teaching Hospital.

ID	**Sex**	**Age**	**Month**	**CSM**
1	2	1	2	1
2	1	1	2	1
3	1	1	2	1
4	1	1	2	1
5	1	1	3	1
6	1	1	3	1
7	2	1	4	1
8	2	1	4	1
9	1	1	4	1
10	1	1	4	1
11	1	1	2	1
12	1	3	10	1
13	1	3	10	0
14	1	7	10	0
15	1	2	10	1
16	1	8	10	0
17	1	4	10	0
18	1	3	10	1
19	1	3	10	0
20	1	2	10	1
21	1	4	10	0
22	1	2	10	1
23	1	3	10	0
24	1	4	11	0
25	1	5	11	0
26	1	6	11	0
27	1	2	11	1
28	1	3	11	1
⋮	⋮	⋮	⋮	⋮
390	2	1	4	1