

Temporal Trends of Malaria Incidence in Southwestern Nigeria: An ARIMA Analysis of Hospital-Based Surveillance Data

Olutunde Michael Ajao¹, Oyindamola John Samson², Felix Azuka Okolie³, Alfred Ayo Ayenigba⁴, Oyindamola Praise Adegbite⁵

^{1,3,4,5}Ajayi Crowther University, Oyo, Nigeria

²Olabisi Onabanjo University, Ago-Iwoye, Nigeria

mo.ajao@acu.edu.ng

Submitted:	Revised:	Accepted:	Published:
Sep 14, 2026	Oct 13, 2026	Oct 24, 2026	Oct 29, 2026

Abstract

Understanding temporal patterns in reported malaria cases is essential for surveillance, healthcare planning, and resource allocation. This study examined trends in malaria case counts and developed a forecasting model using hospital-based surveillance data from Ota, Ogun State, Nigeria. Monthly malaria records from January 2016 to December 2019 were obtained retrospectively from General Hospital, Ota. Descriptive statistics summarized temporal patterns, while Pearson correlation analysis examined the relationship between male and female case counts. Time-series modelling followed the Box–Jenkins autoregressive integrated moving average (ARIMA) approach. Candidate models were compared using the Bayesian Information Criterion, and residual autocorrelation was assessed using the Ljung–Box test. Reported malaria cases fluctuated throughout the study period but exhibited an overall declining trend. The estimated trend equation, $\text{Malaria Cases} = 161,610.958 - 80.025 \times \text{Year}$, indicated a progressive reduction in recorded cases. Male and female case counts followed similar temporal patterns and were strongly positively correlated ($r = .991$, $p < .01$). Among the candidate models, ARIMA(1,1,1) provided the most suitable fit according to the model selection criterion. The Ljung–Box test detected no statistically significant residual autocorrelation, supporting the adequacy of the fitted model for the observed series. Forecasts suggested a continued decline in reported malaria cases over the forecast period. These findings document a reduction in hospital-recorded

cases between 2016 and 2019 and indicate the potential value of ARIMA modelling for surveillance planning. The model may support short-term resource allocation, although predictive accuracy requires evaluation against subsequently observed data. Continued monitoring and assessment of environmental predictors could strengthen future forecasting. Interpretation should remain specific to hospital-reported cases because population incidence cannot be established from case counts alone.

Keywords: Malaria; ARIMA Modelling; Time Series Forecasting; Hospital Surveillance; Nigeria

INTRODUCTION

The name "malaria" derives from medieval Italian words for "bad air" (mala aria). The miasma theory once attributed infectious diseases to foul vapors (Dagen, 2020; Singh et al., 2024). Malaria is a major public health problem in developing countries and is the top-ranked priority tropical disease of the World Health Organization. Plasmodium parasites, transmitted by Anopheles mosquitoes, cause malaria in humans via five species: *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae*, and *P. knowlesi*, with *P. vivax* and *P. falciparum* being the most common and *P. falciparum* the most deadly (Adekola et al., 2023).

Globally, malaria cases have increased in recent years, with Africa bearing the highest burden. In 2022, there were an estimated 249 million cases and 608,000 deaths worldwide (World Health Organization, 2023). Despite control efforts, malaria threatens millions (World Health Organization, 2023; Singh et al., 2024).

In Nigeria, malaria accounts for more than 50% of outpatient visits and 40% of hospital admissions, particularly affecting children under five. It contributes to 30% and 25% of maternal and newborn mortality, respectively (World Health Organization, 2023). Approximately one million deaths occur annually, with 75% in sub-Saharan African children. Nigeria has implemented the Roll Back Malaria initiative to combat this. Malaria poses significant threats to individuals, families, and communities, leading to illness, hospitalization, and death if untreated. Despite efforts by governments and NGOs, it remains a major challenge (Adekola et al., 2023).

However, there is a lack of detailed time-series analysis and forecasting of malaria incidence at the local level in Ota, which is crucial for targeted resource allocation. This study

aims to investigate malaria incidence at General Hospital, Ota, from January 2016 to December 2019, and identify an appropriate statistical model for describing trends and making future projections in Ogun State. This will inform policymakers in developing strategies to reduce malaria incidence.

METHODS

Monthly malaria case records from January 2016 to December 2019 were obtained from General Hospital, Ota. Data were categorized by gender for analysis. Statistical analyses included Chi-square tests for association and ARIMA time-series modeling to identify trends and make forecasts. The Chi-square statistic was computed as:

$$\chi^2 = \sum \frac{(O-E)^2}{E} \dots\dots\dots(1)$$

where O represents the observed frequency and E_i represents the expected frequency.

The Hypothesis to be tested:

Hypotheses:

H_0 : There is no association between gender and malaria.

H_1 : There is an association between gender and malaria.

Correlation Analysis

Correlation analysis was used to assess the relationship between male and female malaria cases:

$$r = \frac{\sum((x-\bar{x})(y-\bar{y}))}{\sqrt{(\sum(x-\bar{x})^2 \sum(y-\bar{y})^2)}} \dots\dots\dots(2)$$

Time Series Analysis: Autoregressive Integrated Moving Average (ARIMA)

ARIMA models were developed to forecast malaria incidence using temporal autocorrelation. The dataset (January 2016–December 2019) was used for model fitting. The ARIMA notation $(p,d,q) \times (P,D,Q)_{12}$ describes autoregressive (p/P), differencing (d/D), and moving average (q/Q) terms. The Box-Jenkins approach was followed: (1) Plot time series and use ACF/PACF to identify terms and achieve stationarity; (2) Fit models and select using BIC to balance fit and complexity; (3) Validate residuals with Ljung-Box test.

Forecast accuracy was assessed via MSE and predictive R^2 . Analyses were conducted using R and SPSS.

RESULTS

This section shows presentation of data, its analysis and the results of our study. Figure 1 is a bar chart of the Total Number of Malaria Cases by Month from 2016 to 2019. Each bar on the chart represents a month, and the height of the bar shows the total number of malaria cases during the month. The bar chart allows monthly comparison malaria cases across four months.

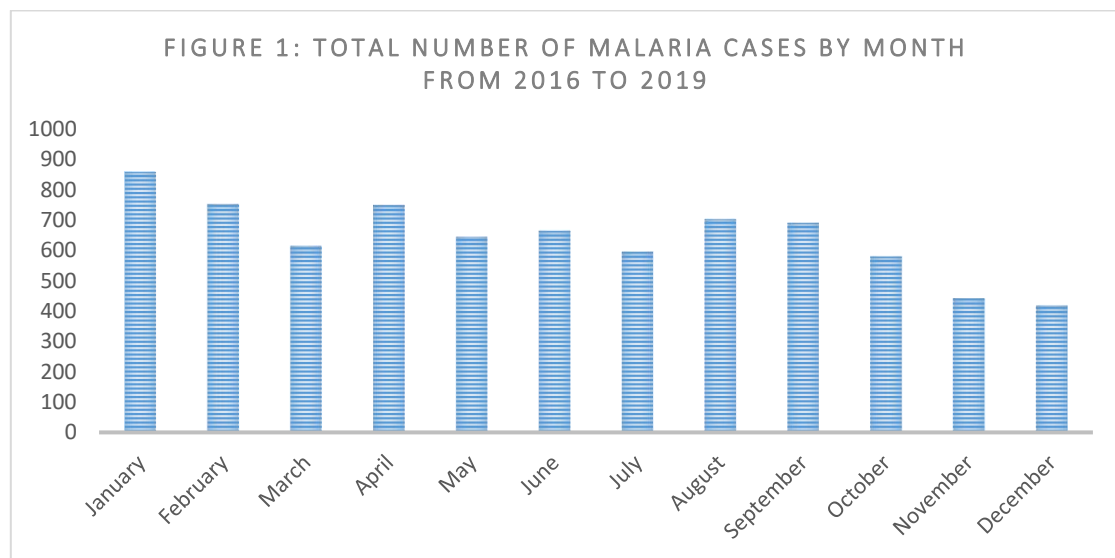


Figure 1: Total Number of Malaria Cases by Month from 2016 to 2019

We went further to perform chi square test to show if there is relationship between malaria based on gender

Table 1: Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1740.000 ^a	1677	.139
Likelihood Ratio	338.364	1677	1.000
N of Valid Cases	48		
<i>a. 1760 cells (100.0%) have expected count less than 5. The minimum expected count is .02.</i>			

The hypothesis tested is:

H₀: There is no significant association between the gender and malaria cases

H₁: There is a significant association between the gender and malaria cases

The p-value of 0.139 is greater than 0.05 indicating a not statistical significance as shown in table 1. we hereby conclude and accept the null hypothesis (H_0) and reject the alternative hypothesis(H_1), indicating there is no relationship between gender and cases of malaria.

Correlation between Malarias in male and female

We want to see the strength of relationship with regards to gender

Table 2: Correlations between Malarias in male and female

		Sex	sex
Sex	Pearson Correlation	1	.991**
	Sig. (2-tailed)		.009
	N	48	48
Sex	Pearson Correlation	.991**	1
	Sig. (2-tailed)	.009	
	N	48	48
**. Correlation is significant at the 0.01 level (2-tailed).			

It was observed that that there is strong relationship between male and female malaria cases that is increase in malaria in male also results to increase in malaria in females as revealed in table 2. Based on the Pearson correlation which is 0.991 A positive relationship indicates two variables moving in the same direction (e.g., as the Male cases of malaria increase, Female cases of female cases also increase). This result indicates that the trends are parallel without gender-specific differences.

Trend Analysis - Malaria Cases

Table 3: Trend Analysis Result for the Malaria Cases

Coefficients ^a					
Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
1 (Constant)	161610.958	26345.543		6.134	.000
Year	-80.025	13.059	-.670	-6.128	.000

a. Dependent Variable: Malaria

$$\text{Malaria Fever} = 161,610.958 - 80.025(\text{Year})$$

Trend Plot of Malaria Cases Between 2016 to 2019

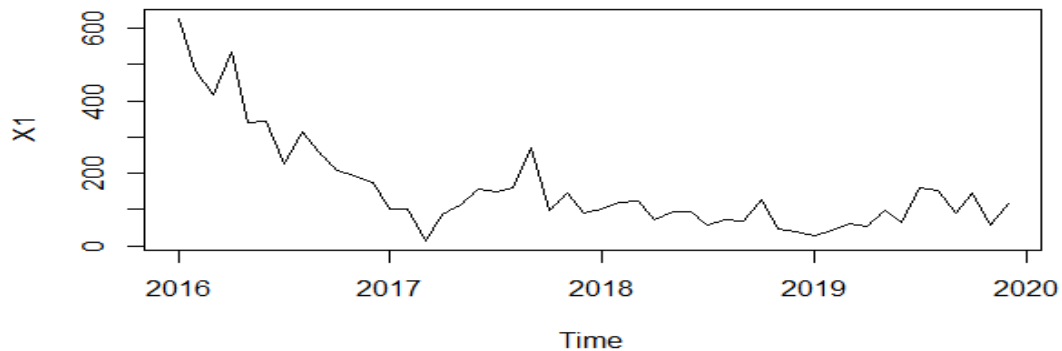


Figure 2: Malaria Cases recorded between 2016 to 2019

This time series helps display the total number of recorded malaria cases from 2016 to 2019 as presented in figure 2. The graph shows how the number of cases changes over time. The horizontal axis represents the years, and the vertical axis represents the total number of cases. Each point on the graph represents the number of cases recorded for a specific year. By looking at the graph, we can easily see any trends or patterns in malaria cases over the four years. It helps us understand how the number of cases fluctuates over time. Singh M. P. et. al (2024)

Estimating the Future Occurrence

Table 4: KPSS Test for Stationarity Check

KPSS Test			
Series	Test Statistics	P-Value	Remark
Malaria	0.80994	0.01	Stationary

The hypothesis to be tested:

H₀: It has a unit root or the data is not stationary

H₁: It is a stationary data

Stationarity is determined by this probability, which is generated with *the* p-value (table 4). In general, the p-value is smaller than 0.05 indicating stationary data for malaria. Since the p-value of 0.01 is less than 0.05, we hereby reject the null hypothesis (H₀) and conclude that the cases of malaria are stationary y.

Table 5: Ljung-Box Test for Model Statistics

Model Statistics						
Model	Number of Predictors	Model Fit Statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
Malaria-Model	0	.626	12.424	15	.647	0

The hypothesis to be tested

H₀: The residuals are independently distributed

H₁: The residuals are dependently distributed, they exhibit serial correlation

The Ljung-Box test, Box G. P. et.al (2016) shows that with significant levels 0.647 (table 5), We accept the null (H₀) hypothesis since our p-values are greater than the given significance level (0.05), which signifies the values are auto-correlated and, the residuals for the time series model are independent for predicting the behavior of spread of malaria over some time.

Table 6: Fitting the ARIMA model for Malaria: six initial models

Malaria Models	BIC conditions	Yes/no
(0,0,0) (0,0,0) ₁₂	11.43205	No
(0,1,0) (0,0,0) ₁₂	12.47895	No
(1,1,1) (0,0,0) ₁₂	11.35443	Yes
(1,1,0) (0,0,0) ₁₂	11.77944	No
(2,0,0) (0,0,0) ₁₂	11.35277	No
(2,1,0) (0,0,0) ₁₂	11.61679	No

The values of pp and qq are then chosen by minimizing the BICs after differencing the data dd times (table 6). Rather than considering every possible combination of pp and qq , the algorithm uses a stepwise search to traverse the model space. Ayele, D. G., & Zewotir, T. T. (2020)

The best model (with the smallest BICs value) is set to be the “current model” for the diabetes model.

An ARIMA (1,1,1) model in figure 3 would have:

$p=1$ $p=1$: One lag of the differenced series in the autoregressive part.

$d=1$ $d=1$: First differencing to make the series stationary.

$q=1$ $q=1$: One lag of the forecast errors in the moving average part.

The best model (with the smallest BICs value) is set to be the “current model” for the malaria model.

an ARIMA (0,1,1) model in figure 4 would have:

$p=0$ $p=0$: No lag of the differenced series in the autoregressive part.

$d=1$ $d=1$: One differencing to make the series stationary.

$q=1$ $q=1$: One lag of the forecast errors in the moving average part.

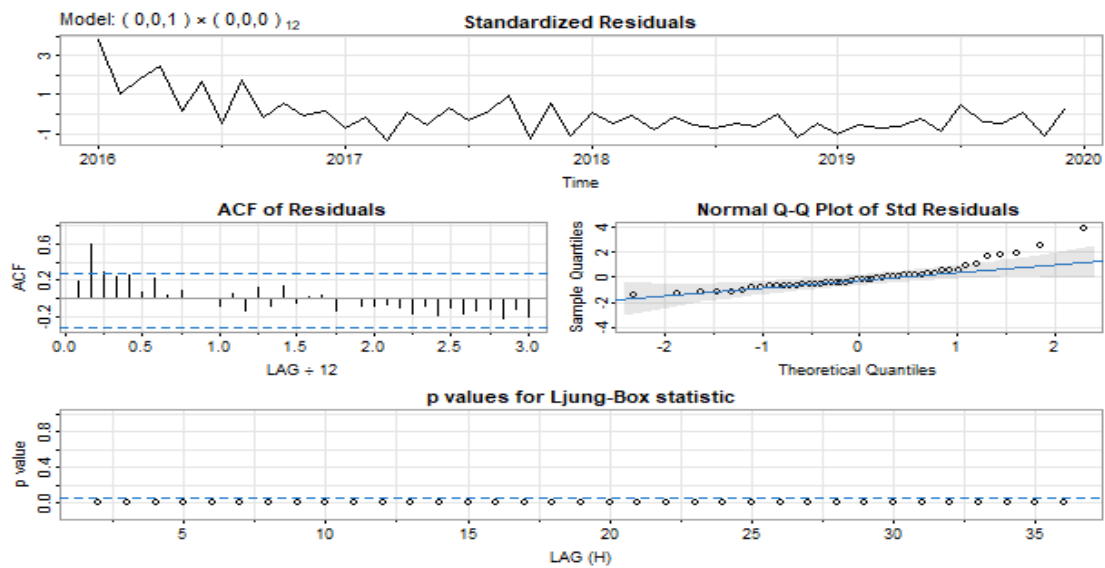


Figure 3: Standardized Residuals for the fitted model

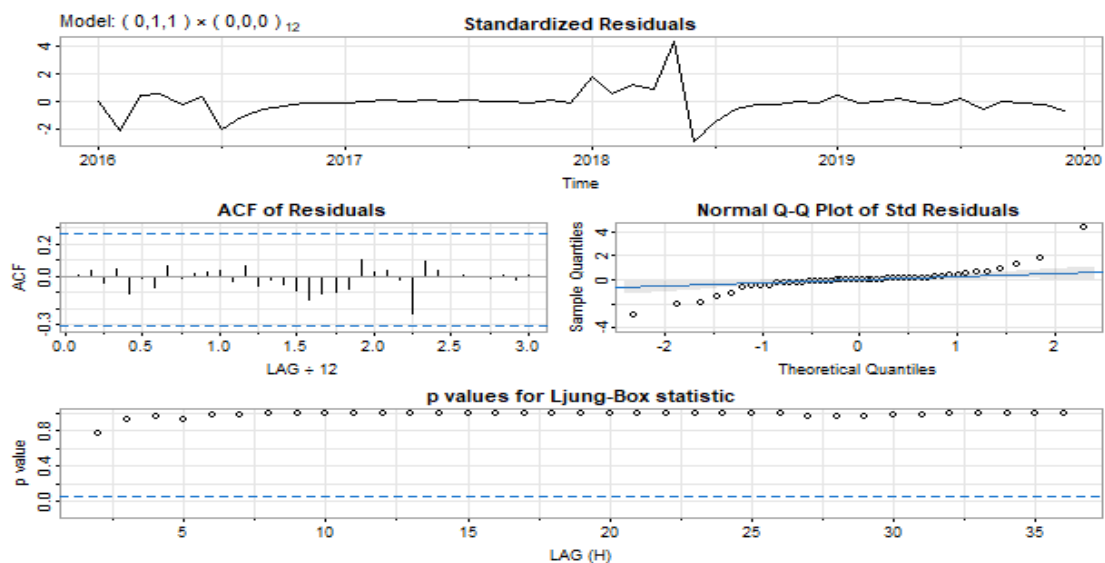


Figure 4: Standardized Residuals for the fitted model (0,1,1) (0,0,0)₁₂

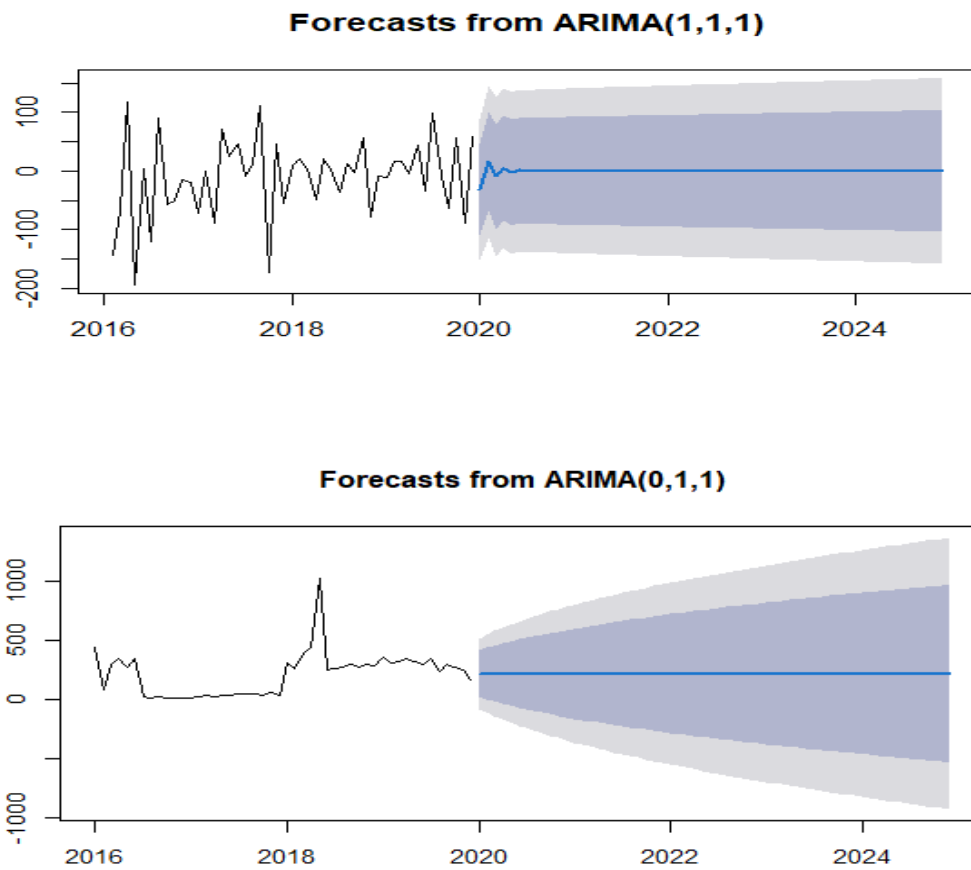


Figure 5: Predictive Analysis of the Incidence Cases of Malaria between 2020 to 2025

The predictive analysis of the incidence cases of malaria between 2020 to 2025 is shown in figure 5. This is a visual representation of the projected number of cases for both diseases during that period. It uses statistical models and historical data to make predictions about future trends. The graph has the years 2020 to 2025 on the horizontal axis and the predicted number of cases on the vertical axis. Each point on the graph would represent a specific year and show the estimated number of malaria and diabetes cases for that year. This type of graph helps us anticipate and understand the potential impact of these diseases in the coming years.

Table 7: Actual (2020) versus Predicted (2021) cases of Malaria

Month/Year	Actual Malaria Cases - 2019	Predicted Malaria Cases-2020
January	28	46.93716
February	44	101.39938
March	60	79.87718
April	55	93.72868

Month/Year	Actual Malaria Cases - 2019	Predicted Malaria Cases-2020
May	99	87.29308
June	64	91.28450
July	162	89.60448
August	154	90.91600
September	90	90.62248
October	147	91.18106
November	58	91.28347
December	117	91.62776
Total	1078	1045.75523

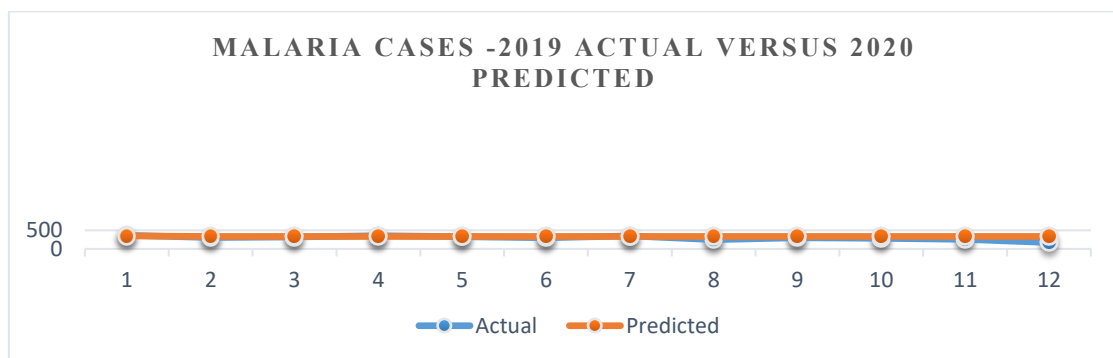


Figure 6: Malaria cases -2019 actual versus 2020 predicted

The graph that shows the actual number of malaria cases in 2019 (figure 6) compared to the predicted number of cases for 2020. This graph shows the difference between the real cases in 2019 and the estimated cases for 2020. The horizontal axis represents the years, with 2019 and 2020 being the specific years we're looking at. The vertical axis represents the number of malaria cases. The graph possesses two lines or bars, one showing the actual cases in 2019 and the other showing the predicted cases for 2020. By comparing these two, we can assess how accurate the predictions were and get an idea of the potential impact of malaria in 2020.

The actual data is 1078 and the predicted data is 1045 as revealed in table 7, which means that the number of malaria cases in 2020 was lower than what was initially predicted. This difference suggests that the situation may have been better than expected, with fewer cases occurring. It's always good news when the actual numbers turn out to be lower than actual statistics, as it means efforts to prevent and control malaria may have been effective.

DISCUSSION

The findings demonstrate that malaria incidence at General Hospital, Ota, decreased between 2016 and 2019. This downward trend may reflect improved preventive interventions, increased community awareness, or seasonal variations in transmission intensity. The lack of gender difference aligns with prior studies indicating that malaria risk is primarily driven by environmental exposure rather than biological sex. The ARIMA(1,1,1) model proved effective in capturing temporal dependencies, supporting its use for short-term forecasting in health surveillance systems. However, limitations include the short duration of data (four years) and reliance on hospital-based records. Future research should incorporate climatic variables such as rainfall and temperature to improve predictive accuracy and extend the forecasting horizon.

CONCLUSION

The conclusion section presents the final synthesis of the results and discussion and should directly address the research objectives. The conclusion should not repeat the results or discussion verbatim; This study concludes that malaria incidence at General Hospital, Ota, showed a declining trend between 2016 and 2019 and was independent of gender. The ARIMA(1,1,1) model provided reliable forecasts and can serve as a practical tool for predicting malaria incidence trends. These insights can assist policymakers in planning proactive malaria prevention and control strategies in Ogun State and similar regions.

Acknowledgement

This study will be impossible without the help of the Management of Ogun State , General Hospital Ota., Ogun state, Nigeria

References

- Adekola, H. A., Egberongbe, H. O., Olanrewaju, M., Onajobi, I. B., Samson, O. J., & Kareem, W. A. (2023). Time-series analysis of malaria cases among suspected febrile patients attending a peri-rural health centre between February 2020–January 2021. *Al-Hayat: Journal of Biology and Applied Biology*, 6(1), 23–30. <https://doi.org/10.21580/ah.v6i1.13915>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Chatfield, C., & Xing, H. (2019). *The analysis of time series: An introduction with R* (7th ed.). CRC Press. <https://doi.org/10.1201/9781351259446>

- Nkumama, I. N., O'Meara, W. P., & Osier, F. H. A. (2017). Changes in malaria epidemiology in Africa and new challenges for elimination. *Trends in Parasitology*, 33(2), 128–140. <https://doi.org/10.1016/j.pt.2016.11.006>
- Patrick, G. L. (Ed.). (2020). *Antimalarial agents: Design and mechanism of action*. Elsevier. <https://doi.org/10.1016/C2015-0-06338-9>
- Segun, O. E., Shohaimi, S., Nallapan, M., Lamidi-Sarumoh, A. A., & Salari, N. (2020). Statistical modelling of the effects of weather factors on malaria occurrence in Abuja, Nigeria. *International Journal of Environmental Research and Public Health*, 17(10), Article 3474. <https://doi.org/10.3390/ijerph17103474>
- Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples* (4th ed.). Springer. <https://doi.org/10.1007/978-3-319-52452-8>
- Singh, M. P., Rajvanshi, H., Bharti, P. K., Anvikar, A. R., & Lal, A. A. (2024). Time series analysis of malaria cases to assess the impact of various interventions over the last three decades and forecasting malaria in India towards the 2030 elimination goals. *Malaria Journal*, 23, Article 50. <https://doi.org/10.1186/s12936-024-04872-8>
- World Health Organization. (2020). *World malaria report 2020: 20 years of global progress and challenges*. <https://www.who.int/publications/i/item/9789240015791>
- World Health Organization. (2021). *World malaria report 2021*. <https://www.who.int/publications/i/item/9789240040496>
- World Health Organization. (2022). *World malaria report 2022*. <https://www.who.int/publications/i/item/9789240064898>
- World Health Organization. (2023). *World malaria report 2023*. <https://www.who.int/publications/i/item/9789240086173>
- World Health Organization. (2024). *World malaria report 2024: Addressing inequity in the global malaria response*. <https://www.who.int/publications/i/item/9789240104440>