Transactions of NAMP 25 (2026) 13-22

Transactions of
The Nigerian Association of
Mathematical Physics
Journal homepage: https://nampjournals.org.ng

Structural Breaks and Volatility Shifts in Monthly Malaria Case Counts at FMC Asaba, Nigeria: A Time-Series Analysis, 2010–2022

S. O. Aghamie*, J. C. Ehiwario, R. N. Nwaka, and S. I. Nwakaego
Department of Mathematics/Statistics, University of Delta, Agbor, Delta State, Nigeria
Corresponding author: sunday.aghamie@unidel.edu.ng; jacob.ehiwario@unidel.edu.ng
ORCID: S. O. Aghamie, 0009-0000-0212-3926; J. C. Ehiwario, 0009-0001-7755-9414

ARTICLE INFO

Article history:
Received xxxxx
Revised xxxxx
Accepted xxxxx
Available online xxxxx

Keywords:
malaria,
time series,
Structural break
Volatility
hospital surveillance
Nigeria

ABSTRACT

Facility-level malaria records can reveal operational changes that annual summaries may miss. This study analysed de-identified monthly outpatient malaria counts from Federal Medical Centre (FMC) Asaba, Nigeria, January 2010–December 2022. After auditing the records, a continuous panel of 156 months was analysed; 147 months were observed and nine missing months were linearly imputed. The total series was log-transformed, decomposed by STL, tested for stationarity, and evaluated with CUSUM, BIC-selected segmented regression, Brown-Forsythe variance tests, ARCH-LM diagnostics, and subgroup break searches. Monthly total counts ranged from 107 to 1,829 (mean = 606.55). Three mean-level breaks were retained: October 2013, June 2015, and June 2019. Mean counts declined from 650.42 to 320.52 across the first three regimes, then rose to 962.23 after June 2019. Residual volatility increased after April 2021. These statistical shifts support break-sensitive facility surveillance, while not identifying causal drivers.

1. Introduction

Malaria remains a major public-health burden, especially in sub-Saharan Africa. In 2022, the World Health Organization estimated 249 million malaria cases and 608,000 malaria deaths globally, with the African Region carrying the largest share of the burden [1]. Nigeria continues to contribute substantially to this regional burden; therefore, local surveillance systems and routine facility data remain important complements to national and state-level summaries.

*Corresponding author: S. O. Aghamie, (08063178775)
E-mail address: sunday.aghamie@unidel.edu.ng
https://doi.org/10.60787/tnamp.v25.714

Hospital records do not measure population incidence unless population denominators and catchment boundaries are available. They can, however, describe the monthly service burden faced by a facility and can identify changes in observed case volume that annual summaries may conceal. This is especially relevant for malaria, where case counts may vary with season, reporting practices, diagnostic access, treatment-seeking behaviour, facility attendance, intervention coverage, and other unmeasured local factors. For this reason, facility-level time series should be interpreted as surveillance and preparedness evidence rather than direct causal evidence of transmission change.

Recent malaria time-series studies have used increasingly varied methods. In Nigeria, Amadi and Erandi [2] modelled associations between malaria incidence and meteorological factors using cluster-integrated regression, while Oniyelu et al. [3] applied wavelet and SARIMAX methods to malaria in pregnancy. More recent Nigerian work by Bakare et al. [4,5] further demonstrated the importance of climate-linked seasonality, annual cycles, and heterogeneity across locations. In other African settings, Jalloh et al. [6] compared SARIMA, SARIMAX, and artificial neural network approaches in Sierra Leone, and Armando et al. [7] developed a climate-informed spatio-temporal prediction model for Mozambique. These studies show that malaria time series are often seasonal, heterogeneous, and locally dependent. Interrupted and segmented time-series approaches have also been used to evaluate long-run malaria changes and intervention periods [8].

A less common focus in the Nigerian facility-level literature is the formal detection of structural breaks and volatility shifts in long monthly hospital series. Such diagnostics are important because a malaria series may retain seasonal peaks while changing in mean level, variability, or subgroup composition. This study therefore analysed 13 years of monthly malaria counts from FMC Asaba, Delta State, Nigeria, with emphasis on break-sensitive surveillance, variance changes, and subgroup patterns. The contribution is the transparent audit and modelling of a long hospital series, not an assertion that the observed breaks identify any single external cause.

1.1 Objectives and research questions

The objectives were to: (i) describe the long-term and seasonal behaviour of monthly malaria case counts at FMC Asaba between 2010 and 2022; (ii) test for structural breaks in the total monthly case-count series; (iii) evaluate whether volatility differed across regimes after modelling the mean structure; (iv) compare male, female, and children component-series break patterns; and (v) discuss implications for facility-level surveillance and preparedness.

The corresponding research questions were as follows. Does the FMC Asaba series contain statistically identifiable structural breaks between 2010 and 2022? If breaks are present, when do they occur, and how do average monthly counts differ across regimes? Does volatility change across regimes after accounting for seasonality and mean structure? Do the male, female, and children series show similar or different break patterns? What do the observed statistical shifts imply for facility-level malaria surveillance and preparedness?

2 Methodology

2.1 Study site and dataset

The study used de-identified monthly aggregate outpatient malaria counts from the Federal Medical Centre (FMC), Asaba, a tertiary hospital in Delta State, Nigeria, serving Asaba and surrounding communities. The original records covered January 2010 to December 2022 and listed monthly male, female, children, and total malaria cases. Because the hospital table defined the total

as the sum of the three subgroup counts, the total series was audited against its components before analysis.

The primary analysis file was a continuous monthly panel of 156 months. Of these, 147 months were observed in the source table and nine missing months were added to complete the monthly time axis and linearly imputed. Eleven inconsistent total values were corrected to equal the corresponding subgroup sum, and 11 unlabeled rows after December 2016 were assigned inferred 2017 month labels according to chronological order. An observed-only file was retained for sensitivity analysis. Because no population denominator was available, the outcome was analysed as monthly malaria case counts rather than population incidence rates. Table 1 summarises the audit.

Table 1: Data-audit summary for the primary analysis file

Item	Value
Months in analysis window	156
Observed months in source table	147
Imputed months added to complete panel	9
Imputed months (%)	5.8
Total values corrected to subgroup sum	11
Rows with inferred 2017 labels	11
Study start	2010-01
Study end	2022-12

Note. The 156-month panel consists of 147 observed months plus nine imputed months. The cleaned imputed panel was used for primary break detection because a continuous monthly time axis was required; the observed-only file was used as a sensitivity check.

2.2 Analysis procedures

The total monthly series was the primary outcome. The male, female, and children series were analysed as parallel component series. Descriptive statistics, monthly averages, and yearly totals were generated first to summarise overall burden and subgroup composition. The total count was transformed as log(Total + 1) to reduce skewness and stabilise variance. Seasonal-trend decomposition using Loess (STL) was used to display trend, seasonal, and irregular components [9]. Stationarity was evaluated using the Augmented Dickey-Fuller (ADF) and KPSS tests [10,11].

Structural change was evaluated in two stages. First, a global CUSUM test was applied to a baseline seasonal-trend model to test overall parameter instability. Second, BIC-selected segmented regression was fitted to the seasonally adjusted log total series, allowing zero to four breaks and requiring a minimum regime length of 18 months. Retained break dates were then included in a full segmented regression in which log(Total + 1) was regressed on linear time, month indicators, and breakpoint step and slope terms. This follows the logic of multiple structural-change modelling [12], while remaining descriptive and facility-specific.

Volatility was evaluated after the mean structure had been modelled. Regime-specific dispersion was compared with the Brown-Forsythe test [13]. Residual conditional heteroskedasticity was evaluated using an ARCH-LM test with 12 lags [14]. A secondary variance-break search was applied to log-squared residuals from the segmented mean model to separate mean-level breaks from variance-level shifts. Subgroup-specific break searches were then carried out for the male, female, and children series. Because the total equals the sum of the subgroup counts, subgroup results were interpreted as component-series comparisons rather than causal leading indicators of the total.

2.3 Statistical model specification

The main modelling steps are summarised below. Let C_t denote the total monthly count at month t and y_t its transformed value:

$$y_t = \log(C_t + 1). \quad (1)$$

The STL decomposition expresses the transformed series as

$$y_t = T_t + S_t + R_t, \quad (2)$$

where T_t is the smooth trend, S_t is the seasonal component, and R_t is the remainder.

The CUSUM process for residuals from a baseline model is

$$W_T(r) = \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{i=1}^{\lfloor Tr \rfloor} \hat{e}_i, \qquad 0 \le r \le 1, \quad (3)$$

where $\hat{e}_i$ are residuals and $\hat{\sigma}$ is their estimated standard deviation.

The ADF auxiliary regression is

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{k=1}^{p} \phi_k \Delta y_{t-k} + \varepsilon_t, \quad (4)$$

where γ is the unit-root coefficient of interest.

For retained breakpoints τ_j, the full segmented regression is

$$y_t = \beta_0 + \beta_1 t + \sum_{m=2}^{12} \delta_m I(M_t = m) + \sum_{j=1}^{J} \left[\theta_j D_{jt} + \lambda_j P_{jt}\right] + \varepsilon_t, \quad (5)$$

where $D_{jt} = I(t > \tau_j)$ is a post-break step term and $P_{jt} = \max(0, t - \tau_j)$ is a post-break slope term.

The Brown-Forsythe transformation for comparing regime dispersions is

$$z_{gt} = |x_{gt} - \tilde{x}_g|, \quad (6)$$

where $\tilde{x}_g$ is the median for regime g; a one-way ANOVA is then applied to z_{gt}.

The ARCH-LM auxiliary regression is

$$\hat{\varepsilon}_t^2 = a_0 + \sum_{\ell=1}^{q} a_\ell \hat{\varepsilon}_{t-\ell}^2 + u_t, \quad (7)$$

where $q = 12$ in this monthly analysis.

The variance-break search used the transformed residual-variance proxy

$$v_t = \log(\hat{\varepsilon}_t^2 + \eta) = \alpha_0 + \alpha_1 I(t > \tau_v) + u_t, \quad (8)$$

where η is a small positive constant and τ_v is the BIC-selected variance-break date.

3 Results

3.1 Descriptive patterns

After cleaning and imputation, the analysis contained 156 monthly observations. Monthly total malaria counts ranged from 107 to 1,829, with a mean of 606.55, standard deviation of 348.82, and median of 535.0. Children contributed the highest average share of the monthly total (43.4%), followed by females (35.4%) and males (21.2%). The highest average monthly total occurred in July (736.54), and the lowest occurred in April (486.92). The highest annual total was recorded in 2021 (12,587), and the lowest was recorded in 2016 (3,801).

Table 2: Descriptive statistics of the cleaned imputed monthly series

Series	Mean	SD	Median	Min.	Max.	Mean share
Male	128.56	87.85	95.0	23	462	0.212
Female	214.72	148.07	167.0	42	703	0.354
Children	263.27	141.30	235.5	37	718	0.434
Total	606.55	348.82	535.0	107	1,829	1.000

Note. Mean shares represent average subgroup proportions of the total monthly case count.

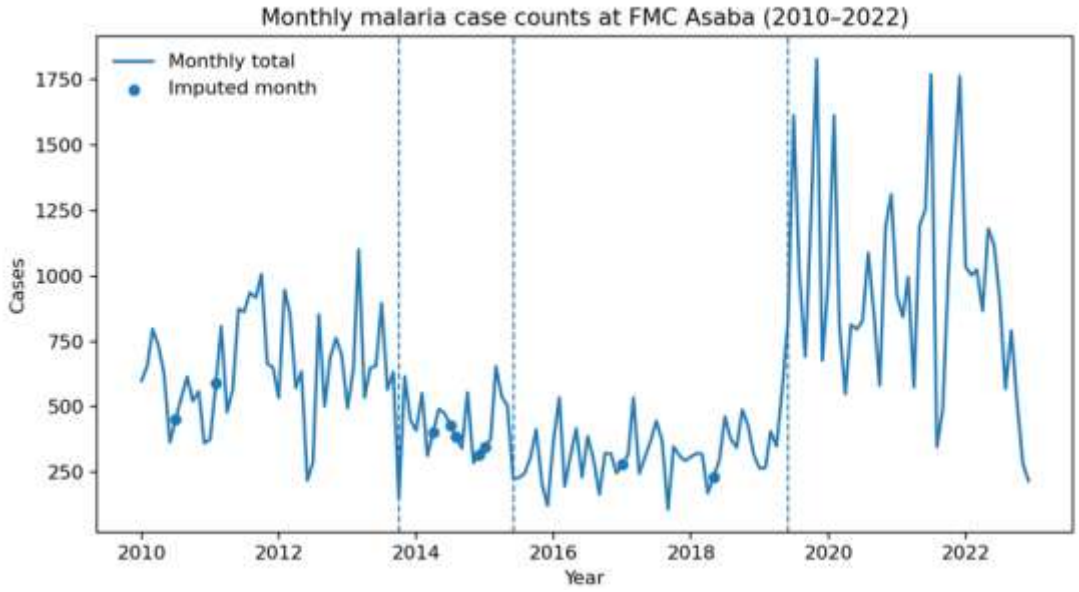

Figure 1: Monthly malaria case counts at FMC Asaba from January 2010 to December 2022. The figure shows the observed monthly total series, identifies imputed months, and marks the BIC-selected structural break dates in October 2013, June 2015, and June 2019.

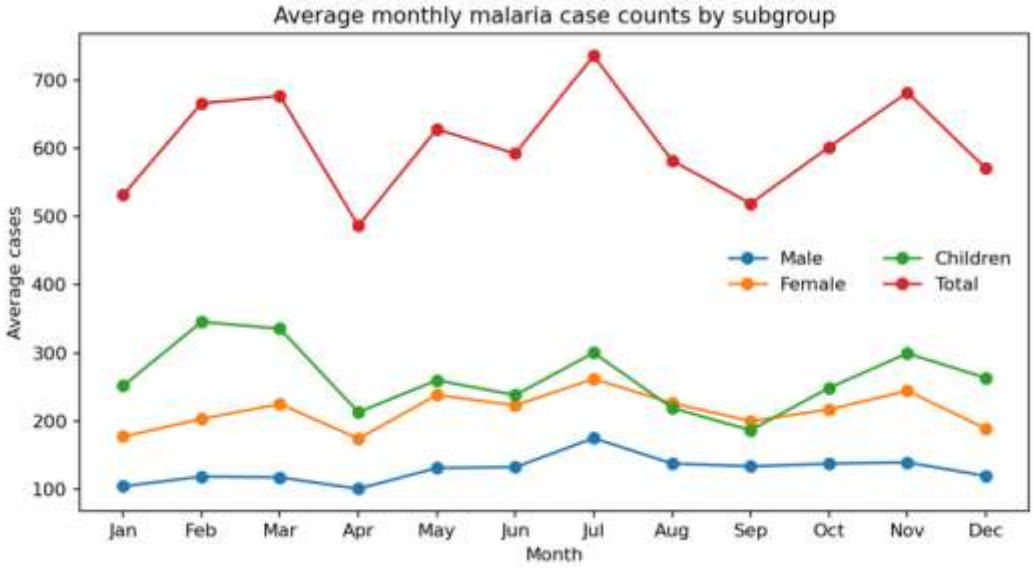

Figure 2: Average monthly malaria case counts by subgroup. The figure displays the male, female, children, and total series by calendar month, highlighting the seasonal pattern and the mid-year peak observed in the total and children series.

3.2 Stationarity and decomposition

The raw log-transformed total series was not stationary by the combined interpretation of the ADF and KPSS tests: ADF = -2.461, $p = 0.125$; KPSS = 0.349, $p = 0.010$. After first differencing, stationarity improved substantially: ADF = -9.264, $p < 0.001$; KPSS = 0.144, $p > 0.10$. The STL decomposition showed a declining long-run trend through 2018, followed by a sharp rise from 2019 onward, while seasonal variation remained present across the study period.

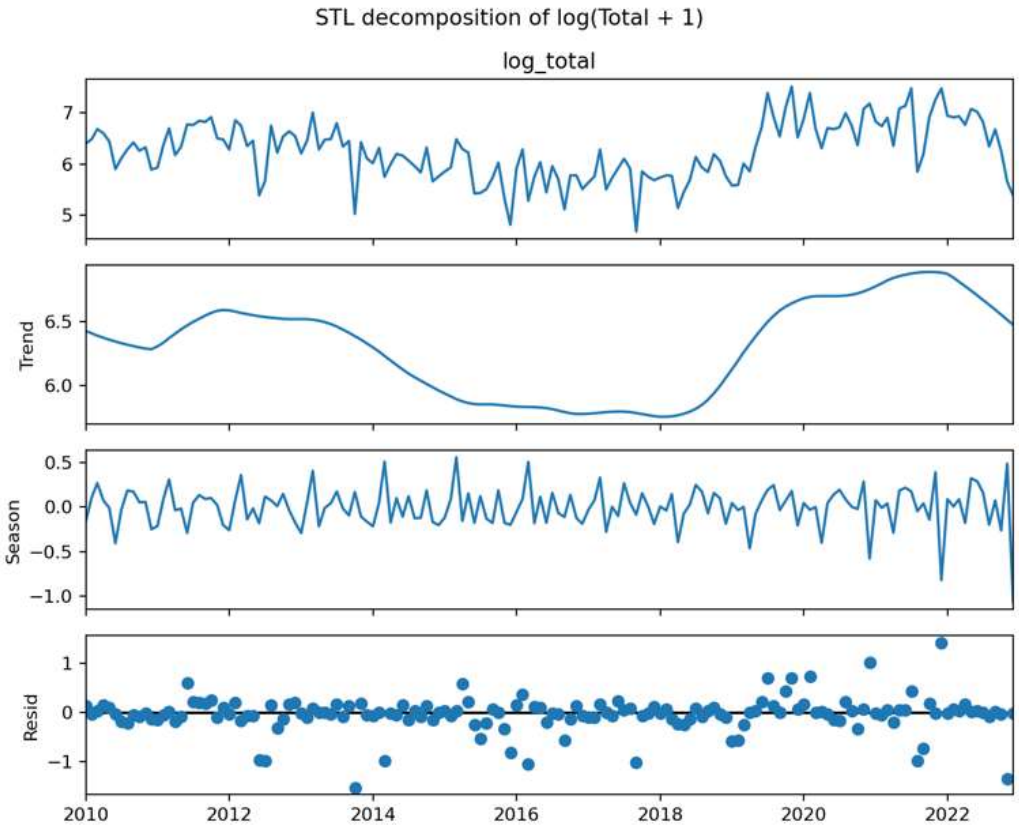

Figure 3: STL decomposition of log(Total+1). The figure displays the transformed series, trend, seasonal component, and remainder, highlighting the gradual decline through 2018 and the pronounced increase beginning in 2019.

3.3 Structural breaks

The global CUSUM test indicated structural instability in the baseline seasonal-trend model (S = 2.445, $p < 0.001$). The segmented search favoured a three-break solution on the seasonally adjusted log series, with retained break dates in October 2013, June 2015, and June 2019 (BIC = -297.32). The full segmented regression with month indicators fit substantially better than the no-break model, $F(6,137) = 36.12$, $p < 0.001$, adjusted $R^2 = 0.594$. These results support the interpretation that the FMC Asaba series is better described as a sequence of regimes than as one stable long-run pattern.

Table 3: Model comparison for segmented regressions on the seasonally adjusted log total series

Number of breaks	Break dates	BIC
0	None	-180.34
1	2019-06	-285.63
2	2015-06, 2019-06	-294.57
3	2013-10, 2015-06, 2019-06	-297.32
4	2013-10, 2015-06, 2019-06, 2021-05	-296.68

Note. The three-break model achieved the lowest BIC and was retained for the main segmented analysis.

Table 4: Regime summaries for the BIC-selected three-break total-count model

Regime	Start	End	Months	Mean	SD	Median
1	2010-01	2013-09	45	650.42	193.58	635.0
2	2013-10	2015-05	20	429.60	122.40	419.5
3	2015-06	2019-05	48	320.52	102.04	315.0
4	2019-06	2022-12	43	962.23	389.05	923.0

Note. Mean monthly counts declined across the first three regimes and increased after June 2019. Relative to the previous regime, the mean changed by -33.95% in Regime 2, -25.39% in Regime 3, and +200.21% in Regime 4.

The observed-only sensitivity analysis retained the June 2015 and June 2019 breaks but not the October 2013 break. This indicates that the mid-2015 and mid-2019 changes are the most robust features of the series, while the October 2013 break should be interpreted more cautiously.

3.4 Volatility shifts

Volatility differed across regimes. The Brown-Forsythe test showed that dispersion in monthly values varied significantly across the four structural regimes, $F(3,152) = 17.33,\ p < 0.001$. Standard deviations declined from 193.58 in Regime 1 to 122.40 and 102.04 in Regimes 2 and 3, then increased to 389.05 after June 2019. This means that the post-2019 period was both higher in average case volume and more variable. Residual conditional heteroskedasticity also remained after modelling the mean structure, ARCH-LM(12) = 23.45, $p = 0.024$. A secondary variance-break search on log-squared residuals identified April 2021 as an additional volatility shift; residual standard deviation increased from 0.303 before the variance break to 0.521 after it.

Table 5: Volatility and stability diagnostics for the total series

Diagnostic	Value
CUSUM statistic	2.445
CUSUM p-value	<0.001
Segmented vs. no-break F	36.12
Segmented vs. no-break p-value	<0.001
Brown-Forsythe F	17.33
Brown-Forsythe p-value	<0.001
ARCH-LM statistic, 12 lags	23.45
ARCH-LM p-value	0.024
Variance break date	2021-04
Pre-break residual SD	0.303
Post-break residual SD	0.521
Observed-only sensitivity breaks	2015-06, 2019-06

Note. The variance-break analysis was applied to log-squared residuals from the segmented mean model.

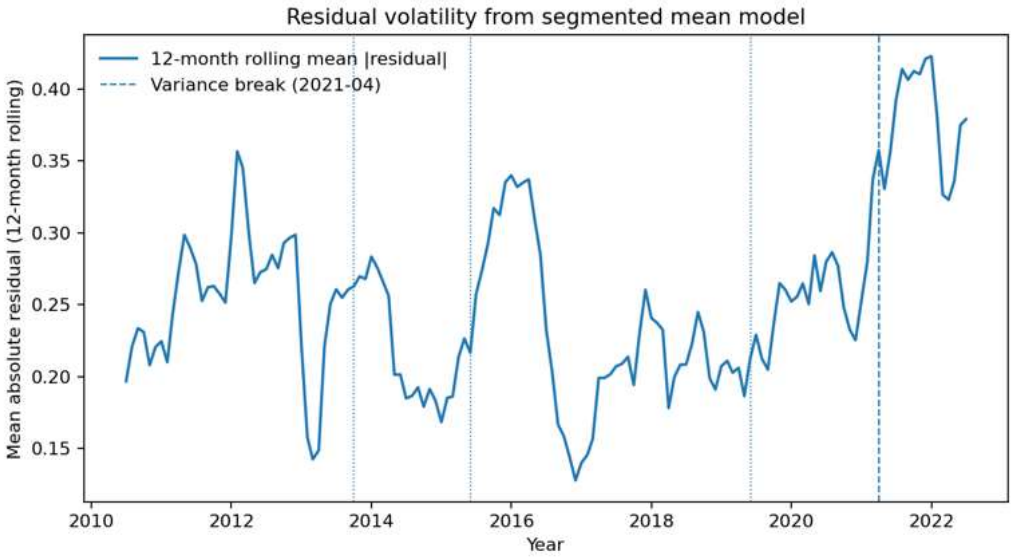

Figure 4: Residual volatility from the segmented mean model. The figure displays the 12-month rolling mean of absolute residuals, with the dashed vertical line indicating the BIC-selected variance break in April 2021. The post-2021 period exhibits increased residual variability relative to earlier regimes.

3.5 Subgroup patterns

Subgroup-specific break searches showed that structural change did not occur uniformly across component series. The male series showed earlier breaks in April 2012 and October 2013, the female series showed an early break in November 2011, and the children series showed a later break in May 2021. All component series showed a post-2019 break, indicating that the post-2019 increase was broadly reflected across groups. Earlier disruptions, however, were more subgroup-specific.

Table 6: Subgroup break dates selected by BIC

Series	Number of breaks	Break dates	BIC
Male	3	2012-04, 2013-10, 2019-05	-242.16
Female	3	2011-11, 2015-06, 2019-05	-256.55
Children	3	2015-06, 2019-06, 2021-05	-275.59
Total	3	2013-10, 2015-06, 2019-06	-297.32

Note. Break dates are reported for the best BIC solution in each series-specific search.

4 Discussion

The main finding is that monthly malaria case counts at FMC Asaba were not adequately described by a single stable trend. The series showed statistically identifiable changes in both mean level and volatility. Counts declined from 2010 to the middle of the decade, remained comparatively low through mid-2019, and then shifted into a higher and more volatile period after June 2019. This pattern supports the use of break-sensitive hospital surveillance because a fixed historical average would underestimate the more recent service burden.

The post-2019 regime is operationally important. In facility planning, a higher mean increases expected demand for rapid diagnostic tests, antimalarial medicines, clinical staff time, and outpatient workflow capacity. A higher variance increases uncertainty around that demand. The April 2021 variance break further suggests that the recent period was not simply a higher-level extension of

earlier patterns, but also a period with larger month-to-month swings. For facility management, volatility is relevant because sudden deviations from expected case volume can create stock pressure, crowding, and workload instability.

These findings should not be interpreted as causal evidence. The data did not include rainfall, temperature, intervention schedules, diagnostic-policy changes, stock-out histories, catchment-population changes, health-service disruptions, or care-seeking indicators. Therefore, the detected breaks identify statistical changes in the observed FMC Asaba case-count series, not the reasons those changes occurred. They may be consistent with several possible epidemiological, environmental, administrative, or behavioural explanations, but the present design cannot distinguish among them.

The subgroup results also require cautious interpretation. Children contributed the largest average share of the monthly total, and the children series showed a later additional break in 2021. However, because the total is mechanically formed from the subgroup counts, no subgroup can be treated as an independent causal leading indicator of the total. The subgroup results are best read as evidence that aggregate hospital burden can mask differences in component-series timing and dispersion.

The study has several strengths. It uses a 13-year monthly hospital panel from southern Nigeria, applies a transparent data-audit process, distinguishes observed and imputed months, and combines descriptive, structural-break, and volatility diagnostics in a single surveillance-oriented analysis. The main limitations are that the data come from one facility, the outcome is case count rather than population incidence, nine months were imputed, year labels for 2017 were inferred from chronology, and no external explanatory covariates were available. Future work should link facility records to rainfall, temperature, intervention calendars, diagnostic practices, stock records, and multi-facility data from Delta State.

5 Conclusion

Monthly malaria case counts at FMC Asaba from 2010 to 2022 showed statistically identifiable structural breaks in October 2013, June 2015, and June 2019, with the most robust sensitivity-supported changes occurring in June 2015 and June 2019. The post-2019 period had both a higher mean case volume and greater volatility, with an additional residual variance shift around April 2021. Children accounted for the largest average share of monthly counts. These findings support facility-level malaria monitoring that goes beyond annual averages by incorporating break-sensitive and volatility-sensitive surveillance. The detected changes are best interpreted as operational warning signals in the hospital record, not as proof of specific causal drivers.

References

1. World Health Organization. *World malaria report 2023*. Geneva: World Health Organization; 2023. Available from: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023

2. Amadi M, Erandi KKWH. Assessing the relationship between malaria incidence levels and meteorological factors using cluster-integrated regression. *BMC Infectious Diseases*. 2024;24:664. doi: 10.1186/s12879-024-09570-z

3. Oniyelu DO, Folorunsho O, Adewole L, Bakare EA, Okoronkwo C, Eze N, et al. Time series analysis of malaria in pregnancy, using wavelet and SARIMAX models. *PLOS ONE*. 2025;20(8):e0328888. doi: 10.1371/journal.pone.0328888

4. Bakare EA, Dukundane D, Salako KV, Kakaï RG, Okoronkwo C, Nelson E. Modelling the association of rainfall and temperature with malaria incidence in Adamawa State, Nigeria. *Scientific Reports*. 2026;16:8761. doi: 10.1038/s41598-026-38705-2

5. Bakare EA, Mogbojuri OA, Bakare DA, Odewusi OJ. Exploring the past and forecasting the future of malaria in selected Nigerian states: A time series modelling approach using wavelet and SARIMA. *PLOS ONE*. 2026;21(4):e0337791. doi: 10.1371/journal.pone.0337791

6. Jalloh SW, Malenje B, Imboga H, Hodges MH, et al. Forecasting malaria cases using climate variability in Sierra Leone. *Malaria Journal*. 2025;24:158. doi: 10.1186/s12936-025-05389-4

7. Armando CJ, Rocklöv J, Sidat M, Tozan Y, Mavume AF, Sewe MO. Spatio-temporal modelling and prediction of malaria incidence in Mozambique using climatic indicators from 2001 to 2018. *Scientific Reports*. 2025;15:11971. doi: 10.1038/s41598-025-97072-6

8. Singh MP, Rajvanshi H, Bharti PK, Anvikar AR, Lal AA. Time series analysis of malaria cases to assess the impact of various interventions over the last three decades and forecast malaria in India towards the 2030 elimination goals. *Malaria Journal*. 2024;23:50. doi: 10.1186/s12936-024-04872-8

9. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics*. 1990;6(1):3–73.

10. Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*. 1979;74(366a):427–431. doi: 10.2307/2286348

11. Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*. 1992;54(1–3):159–178. doi: 10.1016/0304-4076(92)90104-Y

12. Bai J, Perron P. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*. 2003;18(1):1–22. doi: 10.1002/jae.659

13. Brown MB, Forsythe AB. Robust tests for the equality of variances. *Journal of the American Statistical Association*. 1974;69(346):364–367. doi: 10.1080/01621459.1974.10482955

14. Engle RF. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*. 1982;50(4):987–1007. doi: 10.2307/1912773