

# **A STATISTICAL STUDY OF LEAD-LAG RELATIONSHIP BETWEEN BCG AND DPT VACCINATIONS IN ANAMBRA STATE: A CROSS SPECTRUM ANALYSIS**

## **Abstract**

This study examined the lead-lag relationship between Bacille Calmette-Guerin (BCG) and Diphtheria, Pertussis and Tetanus (DPT) vaccinations in Anambra State using cross spectrum analysis with BCG as the input series. The monthly data on the number of vaccinations for BCG and DPT were collected from Expanded Programmes on Immunization office in Awka, Anambra State for the period of January, 2012 to December, 2020. The Fisher's Kappa and Bartlett Komolgorov-Smirnov white noise tests showed that both series were white noise. The series were differenced stationary using Augmented Dickey-fuller test. The coherence squared presented a strong relationship between BCG and DPT vaccinations, though it did not show consistent pattern of relationship, however identified pattern of relationship existed at both lower and higher frequencies. In addition, the phase showed that at lower frequencies BCG led DPT, while at higher frequencies BCG lagged DPT. The analysis showed that there is a high awareness for both vaccinations in the State and it should be sustained by the state government.

**Keywords :** Vaccination, Periodogram, Cross spectrum, Coherence

## **1.0 Introduction**

Vaccination can be defined as an act to introduce specific antigenic substances into the body to produce immunity to a specific disease. It constitutes an important public health management function globally as this strategically reduces morbidity and mortality that maybe associated with identified infectious diseases. Therefore, among these vaccines are BCG (Bacille Calmette Guerin) vaccine which has shown to be very effective against tuberculosis, and DPT vaccine used against three infectious diseases in human: diphtheria, pertussis (whooping cough), and tetanus [1]. Consequently, the administration of these vaccines is carried out with BCG given at birth and DPT administered at 6, 10 and 14 weeks of age [2].

In Nigeria, Anambra State is one of the states in south eastern part of the country, with strong interest to promoting vaccinations of BCG and DPT with a vision to improve the health of

children in the state by eradicating diseases preventable by these vaccines. Nevertheless, there are still challenges to achieving the vision such as immunization fatigue, political problems, poor storage system and poor accessibility to rural dwellers [3]. Therefore, it becomes imperative, to assess the comovement relationship between BCG and DPT vaccination in Anambra State since BCG is administered first at birth and it is expected that other vaccinations should follow up. Thus, does administration of BCG at birth encourage a follow up for DPT vaccination in the state?

Hence, to answer the above question, a frequency domain of time series analysis was adopted in this study. This method is cross spectrum analysis using the Fourier sinusoidal transformation. Consequently, the cross spectrum analysis of lead-lag relationship between BCG and DPT vaccinations in Anambra State was evaluated with BCG as the input series.

## **2.0 Method**

The method that was employed to establish the lead-lag relationship between BCG (Bacille Calmette-Guerin) and DPT (Diphtheria, Pertussis and Tetanus), with BCG as the input series is cross spectrum analysis. Nevertheless, before estimating the cross spectrum, the properties of the variables were substantiated in terms stationarity and white noise of the data series. The statistical tools that were used for these verifications were the Augmented Dickey-Fuller test for stationarity and, Fisher's Kappa and Bartlett Kolmogorov-Smirnov tests for white noise. In addition, the degree of relationship among these variables was ascertained using the computed squared coherence. The negative and positive signs of the phase showed the lead-lag between BCG and DPT.

The data for the variables included in this work were obtained from the Expanded Programmes on Immunization office in Awka, Anambra State. The scope is a monthly data from January, 2012 to December, 2020, while the variables are:

BCG (Bacille Calmette-Guerin) as the input series (q)

DPT (Diphtheria, Pertussis and Tetanus) as the output series (k)

## 2.1 Model Specifications

In order to establish the relationship between BCG and DPT comovement, the frequency spectrum models for the two series were developed using the Fourier transformation. Therefore, BCG ( $q$ ) model is:

$$q_t = \hat{\mu} + \sum_{j=1}^M \left\{ \hat{\alpha}_j \cdot \cos[\omega_j(t-1)] + \hat{\delta}_j \cdot \sin[\omega_j(t-1)] \right\} \text{ with } \hat{\mu} = \bar{q} \text{ (sample mean BCG)} \quad (1)$$

$$\hat{\alpha}_j = (2/T) \sum_{t=1}^T q_t \cdot \cos[\omega_j(t-1)] \text{ for } j = 1, 2 \dots M \quad (2)$$

$$\hat{\delta}_j = (2/T) \sum_{t=1}^T q_t \cdot \sin[\omega_j(t-1)] \text{ for } j = 1, 2 \dots M \quad (3)$$

The sample variance of  $q_t$  can be expressed as

$$\left( \frac{1}{T} \right) \sum_{t=1}^T (q_t - \bar{q})^2 = \left( \frac{1}{2} \right) \sum_{j=1}^M (\hat{\alpha}_j^2 + \hat{\delta}_j^2) \quad (4)$$

And the portion of the sample variance of  $q$  that can be attributed to cycles of frequency  $\omega_j$  is given by

$$\frac{1}{2} (\hat{\alpha}_j^2 + \hat{\delta}_j^2) \quad (5)$$

The portion of the sample variance of  $q$  that can be attributed to cycles of frequency  $\omega_j$  can equivalently be expressed as

$$\frac{1}{2} (\hat{\alpha}_j^2 + \hat{\delta}_j^2) = \left( \frac{4\pi}{T} \right) \cdot \hat{s}_q(\omega_j) \quad (6)$$

where  $\hat{s}_q(\omega_j)$  is the periodogram at frequency  $\omega_j$

Also, the DPT ( $k$ ) model can be expressed as:

$$k_t = \hat{\mu} + \sum_{j=1}^M \left( \hat{a}_j \cdot \cos(\omega_j(t-1)) + \hat{d}_j \cdot \sin(\omega_j(t-1)) \right) \text{ with } \hat{\mu} = \bar{k} \text{ (sample mean DPT)} \quad (7)$$

$$\hat{a}_j = \left( \frac{2}{T} \right) \sum_{t=1}^T k_t \cdot \cos(\omega_j(t-1)), \text{ for } j = 1, 2 \dots M \quad (8)$$

$$\hat{d}_j = \left( \frac{2}{T} \right) \sum_{t=1}^T k_t \cdot \sin(\omega_j(t-1)), \text{ for } j = 1, 2 \dots M \quad (9)$$

The sample variance of  $k_t$  can be expressed as

$$\left( \frac{1}{T} \right) \sum_{t=1}^T (k_t - \bar{k})^2 = \left( \frac{1}{2} \right) \sum_{j=1}^M (\hat{a}_j^2 + \hat{d}_j^2) \quad (10)$$

And the portion of the sample variance of  $k$  that can be attributed to cycles of frequency  $\omega_j$  is given by

$$\frac{1}{2} (\hat{a}_j^2 + \hat{d}_j^2) \quad (11)$$

The portion of the sample variance of  $k$  that can be attributed to cycles of frequency  $\omega_j$  can equivalently be expressed as

$$\frac{1}{2} (\hat{a}_j^2 + \hat{d}_j^2) = \left( \frac{4\pi}{T} \right) \cdot \hat{s}_k(\omega_j) \quad (12)$$

where  $\hat{s}_k(\omega_j)$  is the periodogram at frequency  $\omega_j$

## 2.2 The Cross Spectrum Model

The cross spectrum  $q$  and  $k$  can be expressed through their covariance as:

$$\left( \frac{1}{T} \right) \sum_{t=1}^T (q_t - \bar{q})(k_t - \bar{k}) = \left( \frac{1}{2} \right) \sum_{j=1}^M (\hat{\alpha}_j \hat{a}_j - \hat{\delta}_j \hat{d}_j) \quad (13)$$

Hence the portion of the sample covariance between  $q$  and  $k$  that is due to common dependence on cycles of frequency  $\omega_j$  is given by:

$$\frac{1}{2} (\hat{\alpha}_j \hat{a}_j - \hat{\delta}_j \hat{d}_j) \quad (14)$$

and this is called the cospectrum.

The sample cross periodogram from  $q$  to  $k$  at frequency  $\omega_j$  can be expressed as:

$$\hat{s}_{kq}(\omega_j) = \left( \frac{T}{8\pi} \right) (\hat{a}_j + \hat{d}_j)(\hat{\alpha}_j - i\hat{\delta}_j) = \left( \frac{T}{8\pi} \right) (\hat{a}_j \hat{\alpha}_j + \hat{d}_j \hat{\delta}_j) + i \left( \frac{T}{8\pi} \right) (\hat{d}_j \hat{\alpha}_j - \hat{a}_j \hat{\delta}_j) \quad (15)$$

The real part of (15) is called the cospectrum, while the imaginary part is called the quadrature spectrum and represented in (16) as:

$$\hat{s}_{kq}(\omega_j) = \hat{g}_{kq}(\omega_j) + i\hat{h}_{kq}(\omega_j) \quad (16)$$

Where

$$\hat{g}_{kq}(\omega_j) = \left(\frac{T}{8\pi}\right) (\hat{a}_j \hat{\alpha}_j + \hat{d}_j \hat{\delta}_j)$$

$$\hat{h}_{kq}(\omega_j) = \left(\frac{T}{8\pi}\right) (\hat{d}_j \hat{\alpha}_j - \hat{a}_j \hat{\delta}_j)$$

However in all cases above,  $T$  = total number of observations,  $M \equiv \frac{T-1}{2}$  when  $T$  is odd,

$\hat{\alpha}_j, \hat{a}_j, \hat{\delta}_j, \hat{d}_j$  are coefficients.

### 2.3 The Squared Coherence Model

The coherence squared is defined below as:

$$z_{kq}(\omega_j) = \frac{[g_{kq}(\omega_j)]^2 + [h_{kq}(\omega_j)]^2}{s_{qq}(\omega_j)s_{kk}(\omega_j)} \quad (17)$$

where  $s_{qq}(\omega_j)$  and  $s_{kk}(\omega_j)$  are nonzero. If  $s_{qq}(\omega_j)$  or  $s_{kk}(\omega_j)$  is zero, the coherence squared is defined as zero.  $0 \leq z_{kq}(\omega_j) \leq 1$  for all  $\omega_j$ . If  $z_{kq}(\omega_j)$  is large, this indicates that  $q$  and  $k$  have important cycles of frequency [4].

### 2.3 The Squared Coherence Significance Test

$$C^2 = 1 - \alpha^{\frac{1}{n-1}} \quad [5] \quad (18)$$

Where  $\alpha = 1 - p$

$P$  = significance level

$N$  = degrees of freedom

The null hypothesis of zero coherency is rejected if the value of the computed coherence squared exceeds the tabulated value at a chosen probability level and degrees of freedom [6]. In this study  $p = 0.95$ .

## 2.4 The Phase Model

The phase gives phase difference between two series that yields the greatest correlation for the given frequency and signs (positive or negative) of the phase determine the lead-lag. Positive sign shows that the input series is leading while; negative sign shows that the input series is lagging.

$$\frac{\sin[\theta(\omega_j)]}{\cos[\theta(\omega_j)]} = \frac{h_{kq}(\omega_j)}{g_{kq}(\omega_j)} \quad (19)$$

[4]

## 2.5 The Stationarity Test

A lot of tests are available to test for the stationarity of these variables, but in this work, the Augmented Dickey-Fuller test was applied and is given by the equation

$$\Delta Y_t = \beta_0 + \beta_1 t + \phi Y_{t-1} + \sum_{j=1}^p \alpha_j \Delta Y_{t-j} + \varepsilon_t, \text{ where } t = p+1, \dots, T \quad (20)$$

Where  $p$  lags of  $\Delta Y_{t-j}$  are added to remove serial correlation

Hypothesis

$H_0: \phi = 0$  (there is a unit root in the series)

$H_1: \phi < 0$  (there is no unit root in the series)

The hypothesis is tested on the basis of t-statistic of the coefficient of  $\phi$ , with  $\alpha = 0.05$

Decision rule: Reject  $H_0$  if test statistic is less than critical values, otherwise do not reject [7].

## 2.6 The Fisher's Kappa White Noise Test

This tests the null hypothesis that the values in the series are white noise against the alternative hypothesis that the series has a component. Kappa is the ratio of the maximum value of the periodogram,  $I(f)$ , and its average value

$$\Pr(K > k) = 1 - \sum_{j=0}^q (-1)^j \binom{q}{j} \left[ \max\left(1 - \frac{jk}{q}, 0\right) \right]^{q-1} \quad (21)$$

$q = \frac{N}{2}$  if N is even,  $q = \frac{N-1}{2}$  if N is odd. K is the observed value of Kappa. The null

hypothesis is rejected if this probability is less than the significance level.

## 2.6 Data

The number of monthly vaccinations of BCG and DPT in Anambra State were detrended using monthly growth rates of these vaccinations as:

$$\begin{aligned} q_t &= 100 * [\ln(BCG_t) - \ln(BCG_{t-1})] \\ k_t &= 100 * [\ln(DPT_t) - \ln(DPT_{t-1})] \end{aligned} \quad (22)$$

## 3.0 Results

### 3.1 Descriptive Statistics

**Table 1: Descriptive Statistics on Number of Monthly BCG and DPT Vaccinations in Anambra State**

Statistic	B.C.G	D.P.T
No. of observations	108	108
Minimum	0.00	0.00
Maximum	8900.00	29320.00
Mean	2550.94	6640.94
Variance	22048.74	292298.84
Standard deviation	148.49	540.65

Table 1 shows the descriptive statistics of BCG and DPT monthly vaccinations of Anambra State within the period under consideration in this study.

### 3.2 Stationarity Test

<b>Table 2.0 Dickey-Fuller Test (ADF) for q Series</b>	
Tau (Observed value)	-5.2268
Tau (Critical value)	-0.7888
p-value (one-tailed)	0.0001
alpha	0.05
<b>Table 2.1 Dickey-Fuller test (ADF) for k Series</b>	
Tau (Observed value)	-5.9298
Tau (Critical value)	-0.7888

p-value (one-tailed)	< 0.0001
alpha	0.05

Tables 2.0 and 2.1 show that the computed p-values are less than the alpha level of 0.05, therefore there is no unit root in both series q and k. This means that both series are stationary.

### 3.3 White Noise Test

**Table 3.0 Fisher's Kappa Test for q Series**

White noise tests (q):		
Statistic	Value	p-value
Fisher's kappa	3.7644	0.7545

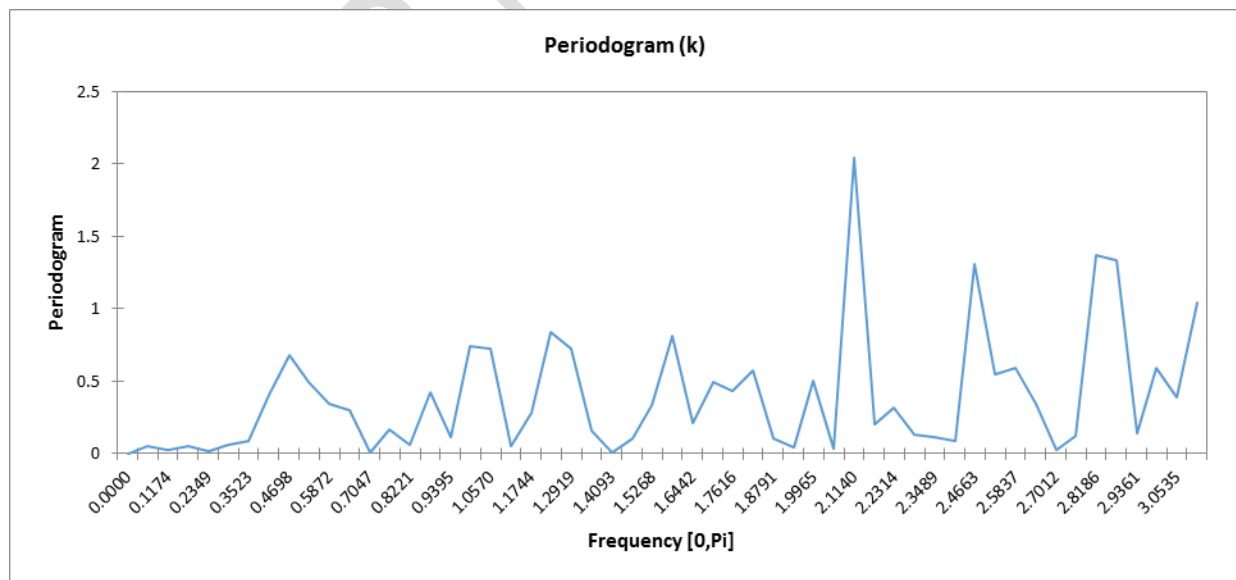
**Table 3.1 Fisher's Kappa Test for k Series**

White noise tests (k):		
Statistic	Value	p-value
Fisher's kappa	5.1141	0.2515

Source: Authors Computation

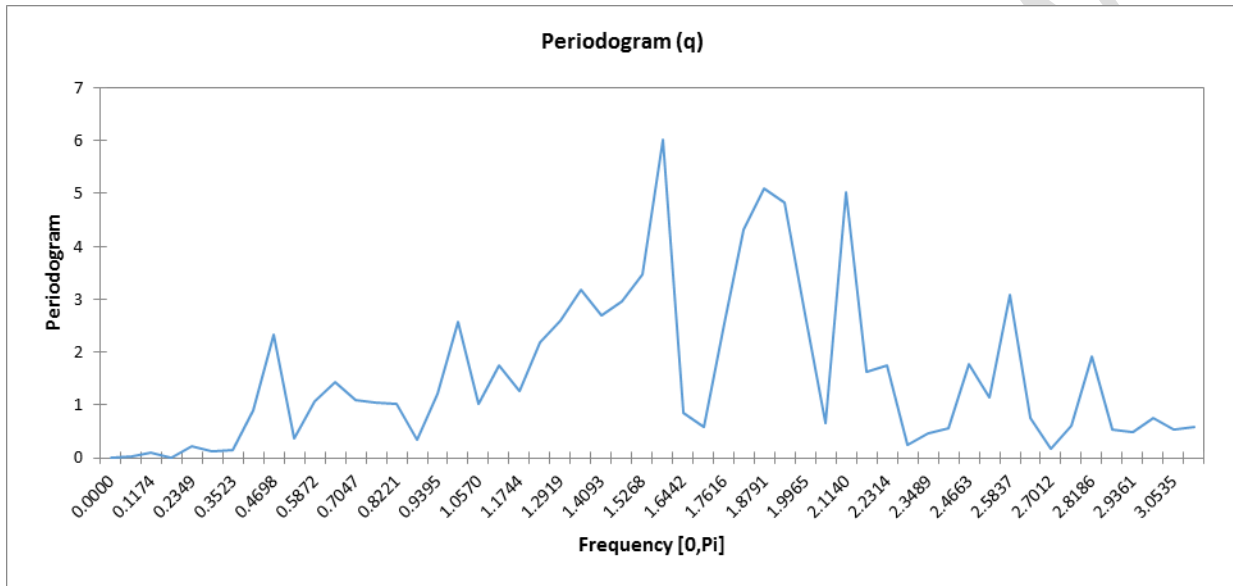
Tables 3.0 and 3.1 show that the computed p-values are greater than the alpha level of 0.05, this indicates that both series q and k are white noise and do not contain any component.

### 3.4 Periodogram of monthly growth rate of DPT and BCG Vaccination series



**Fig. 1: The Periodogram of Monthly Growth Rate of DPT Vaccination in Anambra State**

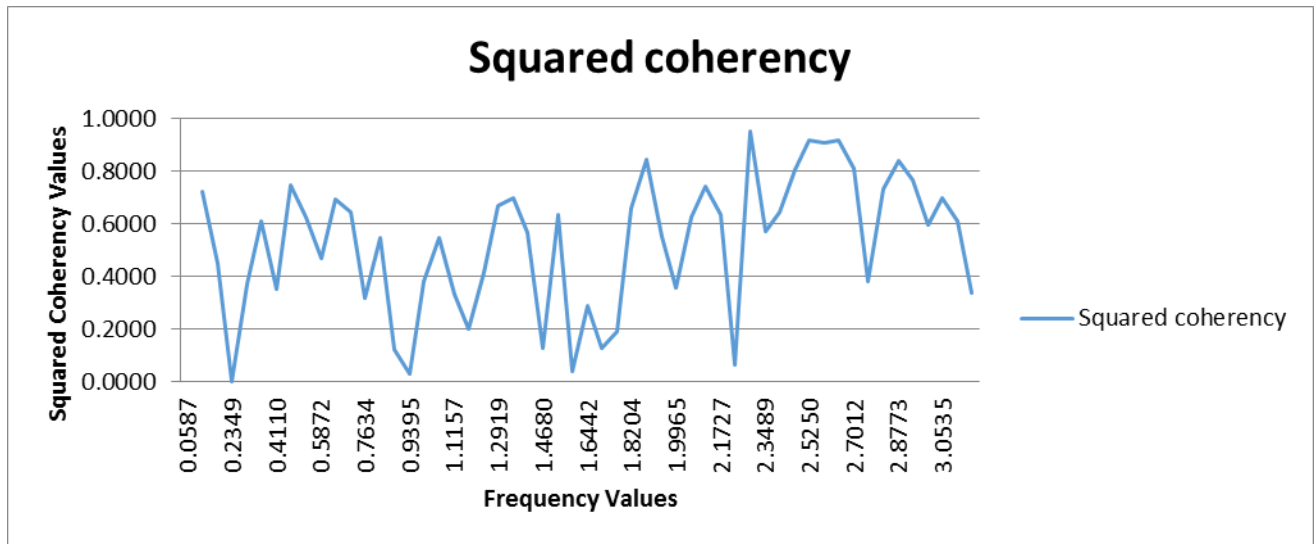
Fig.1 shows that within the power spectrum of the monthly growth rate of DPT vaccinations in Anambra State, frequency 2.1140 contributed the highest proportion of 9.65% to the total variance of DPT, followed by frequency 2.8186 which add 6.50% proportion to the total variation. However, frequency 1.4093 contributed the lowest proportion of 0.03% of the total variance.



**Fig. 2: The Periodogram of Monthly Growth Rate of BCG Vaccination in Anambra State**

In Fig.2, the power spectrum of the monthly growth rate of BCG vaccinations in Anambra State shows that frequency 1.5855 contributed the highest proportion of 7.10% to the total variance of BCG series, while frequency 1.8791 followed with a contribution of 6.03%. Frequency 0.1762 added the lowest proportion of 0.01%.

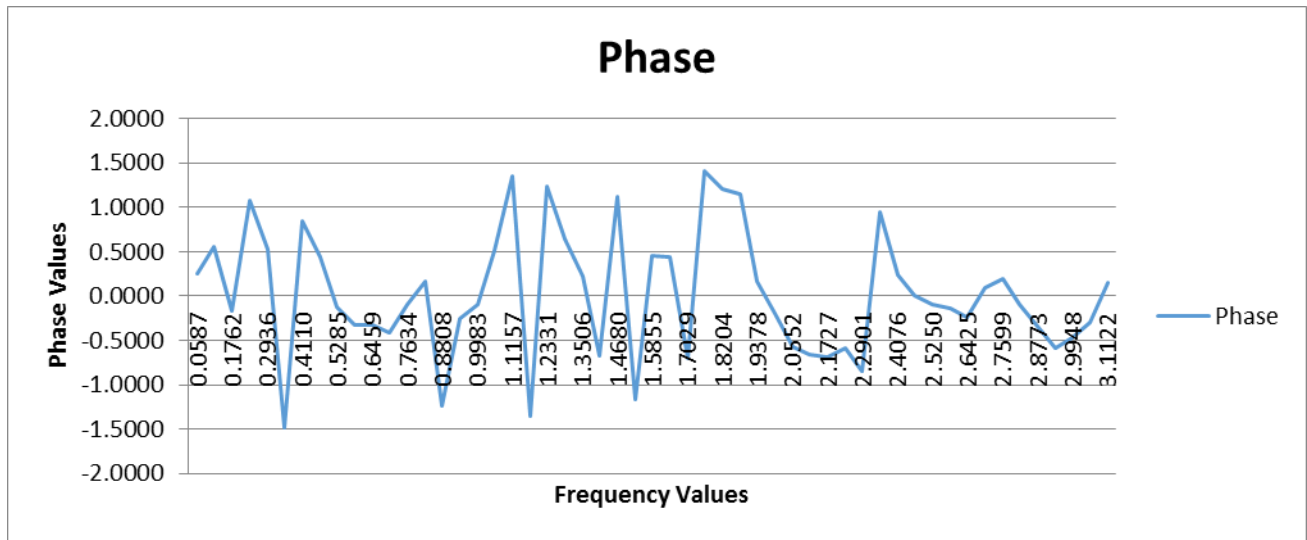
### 3.5 The Squared Coherence



**Fig. 3: The Squared Coherence of Monthly Growth Rate of BCG and DPT Vaccinations in Anambra State**

The tabulated squared coherency value at 106 degrees of freedom and probability of 0.95 is 0.0279. Therefore, Fig.3 shows that majority of the square coherency values of monthly growth rate of BCG and DPT vaccinations are significant at different frequencies since these values are greater than the tabulated squared coherency value. The highest coherency value of 0.9520 occurred at frequency 2.2314 and it is significant, while the least coherency value of 0.0041 happened at frequency 0.1762.

### 3.6 The Phase



**Fig. 4: The Phase of Monthly Growth Rate of BCG and DPT Vaccinations in Anambra State**

Fig. 4 shows that that the phase between monthly growth rate of BCG and DPT vaccinations in Anambra State did not show consistent pattern.

#### 4.0 Discussion

A cross-spectrum analysis was conducted with the number of BCG vaccinations as the input series and number of DPT vaccinations as the output series. Fig.3 and Fig. 4 showed respectively the coherence and phase plots of monthly growth rate of BCG and DPT vaccinations in Anambra State. The horizontal axis of both Fig.3 and Fig.4 shows the frequency plots. Also, the vertical axis of both Fig.3 and Fig.4 shows the coherence and phase levels respectively. A positive phase value means that the input series is leading the output series, while negative value of phase means that the input series is lagging behind the output series. However in Fig.3, majority of the coherence values between monthly growth rate of BCG and DPT in the State were all significant at probability value of 0.95 which indicates that there is a strong relationship and comovement between the two series. The highest coherency value of 0.9520 occurred at frequency 2.2314; this shows that the two series were highly linear at that frequency.

In addition, in Fig.4 the phase did not show a consistent pattern, nevertheless in terms of lead-lag relationship between the two series with BCG as the input series, at lower frequencies the phase showed more of positive sign indicating that BCG vaccinations lead DPT vaccinations confirming that BCG is administered to children first before DPT in Anambra State, therefore suggesting high awareness of these vaccination in the State. Also, at higher frequencies BCG vaccinations lagged behind DPT vaccinations showing that more DPT vaccinations were administered to children in Anambra State after taking BCG vaccine.

## **5.0 Conclusion**

This study used cross-spectrum method to analyze the lead-lag relationship between monthly BCG and DPT vaccination in Anambra State, Nigeria. The analysis of the coherence showed a significant strong relationship between BCG and DPT at almost all the frequencies. The phase showed that monthly BCG vaccinations led DPT vaccination at lower frequencies, in the middle frequencies; there was no clear pattern of leads and lags observed since the diagram crossed the horizontal line severally. However, at higher frequencies BCG vaccination lagged behind DPT vaccinations. This showed that there is a high awareness and administration of these vaccines in Anambra State, Nigeria. Invariably, the results of this study disproved that there is poor administration of these vaccines in Anambra State, Nigeria as mentioned in [3]. Therefore, the government of Anambra State should continue to sustain the current tempo of these vaccines administration in the state.

## References

- [1] World Health Organization (WHO) (2000). Report on Research related to Measles Control and Elimination, Geneva. The Department of Vaccines and Biologicals. Geneva.
- [2] National Programme on Immunization (NPI) (2009). Assuring Vaccine Security in Nigeria. Report of NPI Vaccine Security Mission.
- [3] Brieger W. R., Salami K. K., and Ogunlade B. P. (2004). Catchment Area Planning and Action: Documentation of the Community-based Approach in Nigeria. Arlington: Va.: BASICS II for USAID, 2004.
- [4] Hamilton J. D. (1994). Time Series Analysis. Princeton University Press. Princeton, New Jersey.
- [5] Amos D. E. and Koopmans L. H. (1963). Tables of the Distribution of the Coefficient of Coherence for Stationary Bivariate Gaussian Process.. Published 1963. Mathematics. Journal of the American Statistical Association.
- [6] Goodman N.R. (1957). On the Joint Estimation of the Spectra, Cospectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process. Engineering Statistics Laboratory – Scientific paper 10 (AD 134 919) New York University, New York, 1957)
- [7] Abdullahi F.Z., Ladan, S. and Bakari, Haruna R. (2011). Foreign Private Investment and Economic Growth in Nigeria: A Cointegrated VAR and Granger Causality Analysis”. CBN Journal of Applied Statistics, vol.2. No.2. Pp. 15-28