Scientific
Research
Publishing

Time Series Analysis on Reported Cases of Tuberculosis in Minna Niger State Nigeria

Samuel Olayemi Olanrewaju[1], Emmanuel Oluwatosin Ojo[2], Emmanuel Segun Oguntade[1]

[1]Department of Statistics, University of Abuja, Abuja, Nigeria
[2]Department of Statistics, Federal University of Technology, Minna, Niger State, Nigeria
Email: olanrewaju.samuel@uniabuja.edu.ng

Abstract

Predicting the trend of non-seasonal data is a difficult task in Social Science. In this research work, we used time series analysis of 144 observations on monthly basis for record of reported cases of tuberculosis patients in Minna General Hospital, Niger State from the period of 2007-2018. Exploratory Data Analysis (EDA: Time Plot and Descriptive Statistics), Stationarity Test (ADF), Trend estimation (T_t), Normality Test, and Forecast evaluation were carried out. The Augmented Dickey Fuller test for stationarity was conducted and the result revealed that the series are not stationary but became stationary after first difference. The correlogram established that the ARIMA (2, 1, 3) was the best model this was further confirmed from the result of L-jung Box. Equation for ARIMA (2, 1, 3) was given as $X_t + 0.6867X_{t-1} - 0.8859X_{t-2} = E_t + 1.3077E_{t-1} - 1.2328E_{t-2} + 0.5788E_{t-3}$. Which was used to predict five years likely cases of tuberculosis in Minna for the period of 2019-2023. It was clearly shown from the projection that the reported cases of tuberculosis reduce year by year by 7% over the period under consideration which could be as a result of intervention from government, health worker, and individuals. In line with these findings, we recommend that the management of general hospital to increase awareness campaign to the public on the causes and dangers of tuberculosis.

Keywords

Tuberculosis, Infectious, Vaccine, Stationarity, White Noise Process, Stochastic Process, Gaussian Process

1. Introduction

Tuberculosis is an infectious disease that usually affects the lungs in particular the population in Minna Niger state in Nigeria, where the study takes place [1].

In their study, they determined the prevalence of tuberculosis in Minna and Suleja area of Niger State using standard microscopic technique and sputa culturing procedure. A total of six hundred (600) patients below 80 years of age were screened and enrolled for the study and the percentage prevalence of the physical appearance of the sputa revealed that muco-purulent (20.0%), purulent (10.8%), mucoid (14.2%), and salivary (46.6%). Their study further revealed that 25.5% (153/600) of the studied participants were positive for *Mycobacterium tuberculosis*. The most affected age group was between 11 - 40 years with the percentage prevalence of 81.6% compared with the age group between 41-0 years. Male Patients had a higher prevalence at 60.1%, compared to 39.9% in females. The statistical analysis revealed that there was a significant difference with respect to gender and age ($p < 0.05$). Meanwhile, they only studied the prevalence in 2015 medically but could not predict how the rate of infection would likely be in the nearest future. Meanwhile, this research looks at the present situation of Tuberculosis infection rate and forecast for the future.

Compared with other diseases caused by a single infectious agent, tuberculosis is the second biggest killer globally. In 2015, 1.8 million people died from the disease, with 10.4 million falling ill [2].

In the 18th and 19th centuries, a tuberculosis epidemic rampaged throughout Europe and north America, before the German microbiologist Robert Koch discovered the microbial causes of tuberculosis in 1882. Following Koch's discovery, the development of vaccines and effective drug treatment led to the belief that the disease was almost defeated indeed at one point the United Nation predicted that tuberculosis (TB) would be eliminated worldwide by 2025.

However, in the mid-1980s, Tuberculosis began to risk worldwide, so that in 1993, the World Health Organization (WHO) declared that tuberculosis was a global emergency; the first time that a disease has been labeled as such. The World Health Organization estimates that nine million people a year get sick with tuberculosis, with three million of these "missed" by health system. Tuberculosis is among the top three (3) causes of death for women aged 15 to 44.

Tuberculosis symptoms (cough, fever, night-sweats, weight loss, etc.) may be mild for many months and people ill with tuberculosis can infect up to 10 - 15 other people through close contact over the course of a year. It is an air-born pathogen, meaning that the bacteria that cause tuberculosis can spread through the air from one person to another.

According to [3], people with compromised immune systems are most at risk of developing active tuberculosis. For instance, HIV suppresses the immune system, making it difficult for the body to control tuberculosis bacteria. People with both HIV and TB are around 20 - 30 percent more likely to develop active TB than those who do not have HIV. Tobacco use has also been found to increase the risk of developing active TB. About 8 of T.B cases worldwide are related to smoking. Furthermore, [3] also stated that people with the following conditions have an increased risk: Diabetes - Certain cancers - Malnutrition - Kidney disease.

Aim and Objectives

The main aim of this research is to fit a time series analysis model that will describe the reported cases of tuberculosis in Niger state. The specific objectives are:

➤ To test for stationarity of tuberculosis patient in Minna, Niger State.
➤ Identification and estimation of the model that best describes the data.
➤ To forecast for future occurrences.

2. Literature Review

Tuberculosis (TB) is a major global health problem. Despite progress made in diagnosis and treatment mortality associated with TB remains high. TB was classified by WHO (World Health Organization) in 2016 as the deadliest infection's disease with 5,000 deaths per day ([2]) multi-drug resistant Tuberculosis (MDR-TB) defined as mycobacterium tuberculosis strain resistant to both Isoniazid and rifampicin represents a threat to global TB control. In 2015, among the 10.4 million new cases of TB disease, 48,000 were confirmed to be MDR-TB and 100,000 were found to have rifampicin resistant TB (RR-TB) and treated with second line TB drugs [4]. Also contributing is HIV/AIDs, a well-known risk factor for TB disease, TB drug resistance and TB related death [5]. Among the 1.8million TB death in 2015, 22% were HIV co-infected and 35% of HIV deaths were due to TB [4]. The emergence of MDR-TB has been sustained by the lack of diagnostics, non-adherence to first line treatment, retreatment failure and poor second line therapy success in many countries (Mc Bride *et al.*, 2017). Several risk factors have been associated with MDR-TB. History of prior TB treatment, HIV infection, contact with a known TB patient, receipt of more than two treatment courses, the large burden of bacilli on sputum microscopy, long cavitation and bilateral lung disease [6] [7]. Since 2010, the World Health Organization (WHO) recommended GeneXpert MTB/RIF for rapid detection of rifampicin resistance and diagnosing MDR-TB in more than 90% f those tested [1]. In Mali, the prevalence of MDR-TB was 3.4% and 66.3% respectively in new and previous tested TB patients [8]. The microbiological confirmation of MDR-TD is based on labor-intensive costly and time-consuming culture and drug susceptibility testing (DST) methods that required extensive laboratory infrastructure and are not routinely available in countries with limited resources. These limitations often result in over use of unnecessary second-line TB drugs for individuals with susceptible TB disease and non-tuberculosis Mycobacteria (NTM). We aimed to determine clinical factors associated with microbiologic confirmation of MDR-TB disease among patients suspected of having MDR-TB by clinical criteria in Bamako, Mali. This knowledge is important for clinicians to identify individuals who are most likely to have confirmed MDR-TB in this setting.

Tuberculosis (TB) is a disease caused by germs that are spread from person to person through the air. It usually affects, the lungs, but it can also affect other parts of the body, such as the brain, the kidneys or the spine. A person with TB

can die if they do not get treatment [7]. Tuberculosis is a chronic infectious and communicable granulomatous disease caused by mycobacterium tuberculosis [9]. The tubercle bacilli establish infection in the lungs after they are carried in droplets small enough (5 to 10 microns) to reach the alveolar spaces. If the defense system of the host fails to eliminate the infections, the bacilli proliferate inside alveolar macrophages and eventually kill the cells. The infected macrophages produce cytokines and chemokines that attract other phagocyte cells, including monocytes, other alveolar macrophages and neutrophils, which eventually form a Nodular granulomatous structure called the tubercle. If the bacterial replication is not controlled, the tubercle enlarges and the bacilli enter local draining lymph nodes.

This leads to lymphadenopathy, a characteristic clinical manifestation of primary tuberculosis [5].

There are two sources of TB infection human and bovine (connected with domestic and wild mammals). The most common source of infection is the human cases whose sputum is positive for tubercle bacilli and who has either received no treatment or not been treated fully. Among the members of the mycobacterium tuberculosis a human pathogen, whereas mycobacterium bovid has a broad host range and is the principal agent responsible for tuberculosis (TB) in domestic and wild mammals. Mycobacterium bovid also infect human, causing zoonotic TB through ingestion, inhalation and less frequently, by contact with mucous membranes and broken skin. Zoonotic TB is indistinguishable clinically or pathologically from TB caused by M. tuberculosis [9].

Tuberculosis is one of the most common life-threatening infections among the persons living with HIV/AIDs, but it does not inevitably follow that HIV is common in TB patients. Earlier survey conducted in Bangladesh to evaluate the prevalence of HIV in TB patients has shown insignificant levels [10].

People who have Diabetes mellitus (DM) and live with TB patients, they are more risk for developing TB disease. Body immunity is a key power to protect the body from infections disease. Incidence of tuberculosis is greatest among those with conditions impairing immunity such as human immunodeficiency virus (HIV) infection and diabetes [11].

Currently, the standard short-course chemotherapy for TB treatment comprises a 6-month regimen. There are four drugs for the intensive phase Isoniazid (INH), Rifampin (RIF), Pyrazinamide (PZA), Ethambotol (EMB), and two-drugs for continuation phase Isoniazid and Rifampin. The 4-months continuation phase is used for majority of patient's only two drugs. Although these regimens are broadly applicable, there are modifications that should be made under specified circumstances. Alternative Chemotherapy using most costly and Toxic drugs, often for prolonged duration's generally 18months is required for multidrug resistant and extensively drug resistant tuberculosis. Directly observed treatment, (DOT) as part of a holistic development of drug resistance in tuberculosis [12].

There are two types of prevention clinical and behavioral preventive meas-

ures. Clinical preventions are vaccination with Bacillus calmette Guerin (BCG) and Isoniazid preventive therapy is used to prevent TB infection BCG vaccination significantly reduces the risk of tuberculosis by an average of 50% vaccination with BCG was significantly associated with a reduction in the incidence of Pulmonary tuberculosis and extract Pulmonary disease.

Preventive therapy with Isoniazid reduces the risk of disease among recently children by 60% - 80% and side effects are rare. Preventive treatment among adults with latent tuberculosis infection also has a protective efficacy in the range 60% - 80%, depending on the duration of therapy. Effectiveness is routine practice may be limited by partial uptake and compliance [13].

Effective preventive measures are essential to reduce tuberculosis (TB) transmission. [14] conducted a study to determined knowledge and acceptability of potential patient specific TB infection contract measure in a rural south African community. Their study results showed that most participants (89%) accepted the wearing of face masks in health facilities, but only 42% of TB suspends and 66% of participants accepted separate cohorting in health facilities and avoidance of co-sleeping with uninfected household members.

Another study was conducted by [15] to assess the knowledge of tuberculosis among the general population in North Libya. Their study result shows that 77.4% of respondents knew that a healthy diet can prevent transmission of tuberculosis from one person to another and 43% of respondents know that a well-ventilated home can prevent tuberculosis infection. [16] study found that 56.8% of participants knew that covering the nose and the mouth while coughing or sneezing stops the transmission of tuberculosis. The Practical Approach to Lung Health in Bishkek, Kyrgyzstan by [17] assesses the results of training family doctors in Practical Approach to Lung Health (PAL) techniques. Findings suggest that PAL training has resulted in a decrease by one-third in referrals to hospitals, specialists or diagnostic tests. Data do not show any improvement in tuberculosis case detection. However, in the impact survey the number of drugs prescribed per patient decreased by 13.6% and the average cost of prescription of any drug per patient was reduced by 32.2%. The study suggests that training in standardized PAL guidelines is likely to reduce referral as well as drug prescription costs for respiratory patients.

Summarly, majority of the authors that worked on TB related research in Niger State of Nigeria studied prevalence using percentage analysis which may not be the best to describe the rate of infection and formulation of the best model to predict the future. This study was able to proffer solution to this fault.

3. Research Methodology

The methodology for this research work is time series approach, the method as applied by Box and Jenkins in the year 1970 to 1971 with an iterative six-stages modeling approach: Series stationary, Series transformation, Model identification, Model estimation, Model evaluation, Validation of diagnostic checking,

and Forecasting. These procedures will include the tentative stage which involves the identification of the Autoregressive moving average of the prospective model, estimation of ARMA model and the model diagnostic stage.

Before model formulation in time series analysis, stationarity and normality tests for data to be used is very important as stated in objective 1. We used graphical method and Augmented Dickey-Fuller unit root test for stationarity of the data and computations from skewness and kurtosis were used to determine whether the data is normally distributed or not. Statistical explanation is given below:

4. Unit Root Tests (Testing for Series Stationary)

For a univariate time series, the unit root test is frequently employed for testing stationary. The first test is frequently employed for testing stationary. The first test poses the null hypothesis that the given time series has a unit root, which means that the time series is non-stationary and tests if the null hypothesis is to be statistically rejected in favor of the alternative hypothesis that given time series is stationary. To detect whether a given series is non-stationary, let us assume that the relationship between current (in time t) and last value (intime $t - 1$) in the time series is s follows (Enders, 1995)

$$X_t = \phi X_{t-1} + E_t \qquad (1)$$

where X_t is an observation values at time t E_t is White noise process this model is a first order autoregressive process. The time series $X+$ converges as $t \to$ to a stationary time series if $/\phi/ < 1$. If $/\phi/ = 1$ or > 1, the series $X+$ is not stationary and the variance of X_t is time independent in other words, the series has a unit root. The unit root test subsequently tests the following one-sided hypothesis

H0: $\phi = 1$ (has a unit root)

H1: $\phi < 1$ (has root outside the unit circle)

A well-known test that is valid in large sample is the Augmented Dickey-Fuller test. The optional finite samples test for a unit roof in autoregressive models were developed by John Denis Sargan and Alok Bhargava. Other tests are the Phiolips-person (PP) and kwlathowski-Phillips-Sclimidt-Shin (KPSS) test. This test uses the existence of a unit root as the null hypothesis.

5. Forecasting

One of the most important objectives of time series analysis is to predict its future vales. It is all about making projection into the future from its past values on that basis of a model that effectively describes the evolution of a series.

Forecasting Based on Conditional Expectation

Let y_t be determined by a set of variables. X_t observed at data t are the past values. Suppose a forecast of y_{t+1} is required base on m most recent values X_t would consist of a constant plus X_t y_{t-1}, …

6. Analysis and Interpretation

6.1. Stationarity Test

The time plot shown in **Figure 1** revealed that the trend of reported cases of tuberculosis in Minna General Hospital data is not stationary and it implies that the monthly records of tuberculosis cases do not maintain an identical pattern during corresponding period of successive years significantly from 2007-2018. More so, the unit root test shown in section 4.2 also, revealed the P-value of 0.1620, which means the result is not significant at 1%, 5% and 10% significant levels. We now accept H_0 and conclude that there is unit root (*i.e.* the data is not stationary).

6.2. Unit Root (ADF) Test

Results of Augmented Dickey-fuller Test at Level

Null Hypothesis: TUBERCULOSISPATIENTS has a unit root

Exogenous: Constant

Lag Length: 4 (Automatic—based on AIC, maxlag = 13)

			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.336845	0.1620
Test critical values:		1% level	−3.477835	
		5% level	−2.882279	
		10% level	−2.577908	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D (TUBERCULOSISPATIENTS)

Method: Least Squares

Date: 10/28/19   Time: 13:57

Sample (adjusted): 2007M06 2018M12

Included observations: 139 after adjustments

| Variable	Coefficient	Std. Error	*t*-Statistic	Prob.
TUBERCULOSISPATIENTS(−1)	−0.204282	0.087418	−2.336845	0.0209
D(TUBERCULOSISPATIENTS(−1))	−0.416294	0.105804	−3.934593	0.0001
D(TUBERCULOSISPATIENTS(−2))	−0.260722	0.103704	−2.514092	0.0131
D(TUBERCULOSISPATIENTS(−3))	−0.273103	0.098027	−2.786011	0.0061
D(TUBERCULOSISPATIENTS(−4))	−0.196366	0.085095	−2.307618	0.0226
C	3.936736	1.923825	2.046306	0.0427
R-squared	0.305614	Mean dependent var		0.000000
Adjusted R-squared	0.279509	S.D. dependent var		12.60607
S.E. of regression	10.70025	Akaike info criterion		7.620618
Sum squared resid	15227.89	Schwarz criterion		7.747286
Log likelihood	−523.6329	Hannan-Quinn criter.		7.672092
F-statistic	11.70721	Durbin-Watson stat		2.019307
Prob (F-statistic)	0.000000			
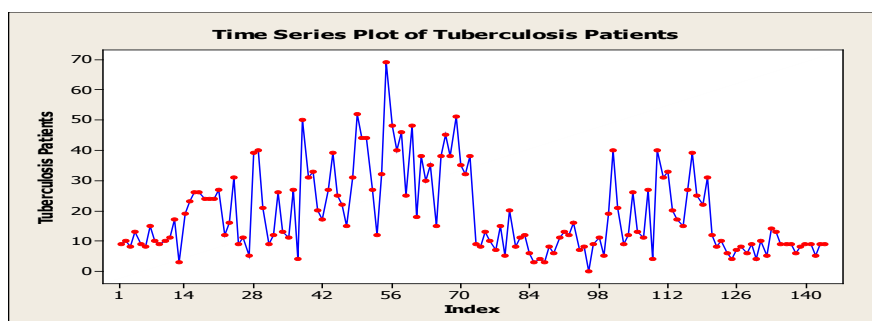
Figure 1. Time plot of reported tuberculosis patients.

6.2.1. Trend Estimation (T_t)

Figure 2 below shows the trend of tuberculosis patients reported in Minna General Hospital. The trend which is the red line in the chart refers to the general direction in which time plot appears to be moving and it is trending downward with fact of negative slope. This is an indication that as the year increases, there is tendency of decrease in the number of people infected by TB.

6.2.2. Descriptive Statistics of Tuberculosis Patients Reported in General Hospital Minna

The summary of the descriptive statistics of this research work is presented below. The skewness is an indicator of the asymmetry or deviation of the variables from a normal distribution with an expected value of zero, the kurtosis defines the degree of flattening or peakedness of a distribution with an expected value of three. More so, It can be shown from **Table 1** and **Figure 3** below that the monthly records of reported cases of tuberculosis patients in Minna General Hospital is not stationary with chart of normal distribution that is skewed to the right with the value of 1.02, mean of 19.03 and S.D of 13.49 which means the variable of tuberculosis record possessed a skewness greater than zero (positively skewed), kurtosis of the variable is peaked with the value of 0.48.

6.3. Stationarity Test at First Difference

Since the result above is not stationary, in principle we have to compute the first difference of the data and rerun the augmented DF test.

6.3.1. Results of Augmented Dickey Fuller Test at First Difference

Null Hypothesis: D (TUBERCULOSISPATIENTS) has a unit root

Exogenous: Constant

Lag Length: 3 (Automatic—based on AIC, maxlag = 13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		−9.986529	0.0000
Test critical values:	1% level	−3.477835	
	5% level	−2.882279	
	10% level	−2.577908	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D (TUBERCULOSISPATIENTS, 2)

Method: Least Squares

Date: 10/28/19 Time: 14:00

Sample (adjusted): 2007M06 2018M12

Included observations: 139 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(TUBERCULOSISPATIENTS(−1))	−2.557428	0.256088	−9.986529	0.0000
D(TUBERCULOSISPATIENTS(−1), 2)	0.985917	0.209341	4.709619	0.0000
D(TUBERCULOSISPATIENTS(−2), 2)	0.607108	0.150650	4.029916	0.0001
D(TUBERCULOSISPATIENTS(−3), 2)	0.246442	0.083712	2.943928	0.0038
C	−0.027217	0.922576	−0.029501	0.9765

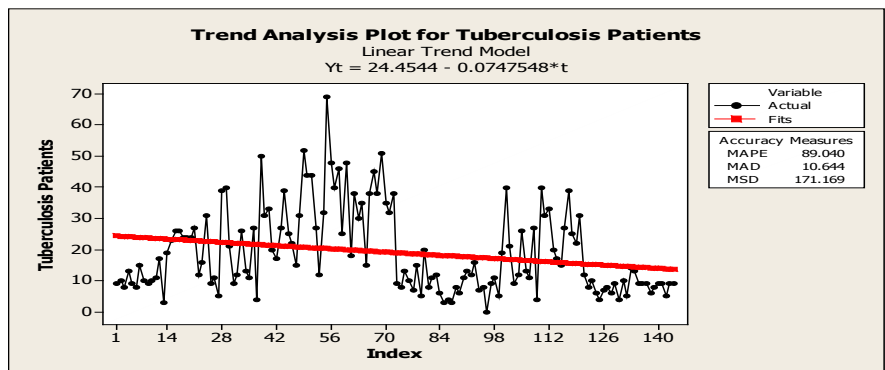|			
R-squared	0.740470	Mean dependent var	0.028777
Adjusted R-squared	0.732722	S.D. dependent var	21.03894
S.E. of regression	10.87690	Akaike info criterion	7.646468
Sum squared resid	15853.13	Schwarz criterion	7.752024
Log likelihood	−526.4295	Hannan-Quinn criter.	7.689363
F-statistic	95.57925	Durbin-Watson stat	2.042465
Prob(F-statistic)	0.000000		

Figure 2. Graph of the trend.
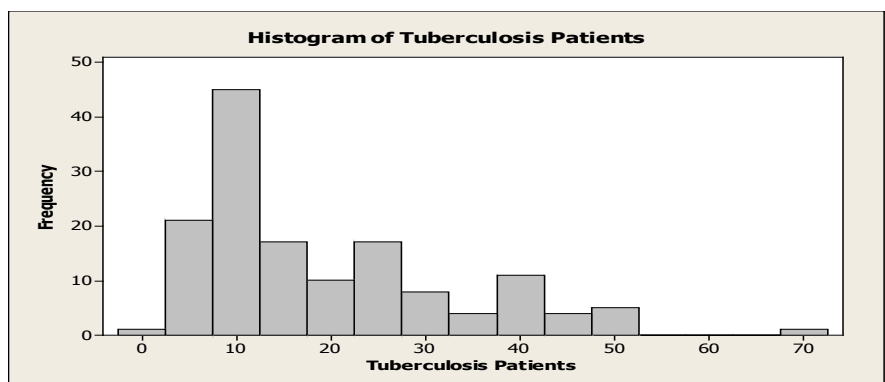
Figure 3. Graph of histogram.

Table 1. Descriptive statistics of tuberculosis of patients' data.

Variable	Mean	SD	VAR	Sum	Min	Max	Median	Skewness	Kurtosis
Values	19.03	13.49	182.09	2741	0.00	69.00	13.00	1.02	0.48

The new result shows that, the *P*-value = 0.000 which is less than significant levels at 1.5 and 10%, we have every reason to reject H0 at zero order levels meaning it is stationary.

6.3.2. Time Plot of Reported Tuberculosis Patients after First Differencing

The time series plot below shows that the first differencing of the data of tuberculosis patients reduces with time. It shows that the data is within the range of linearity from 2007-2018 after the process. This was also confirmed in **Figure 4** and **Figure 5** given below.

It can be affirmed from the above graph and table that the monthly records of reported cases of tuberculosis patients in Minna General Hospital are stationary with the chart of normal distribution that is skew with the value of 0.57, mean of 1.619224E−16 and S.D of 12.44.

6.3.3. Trend Analysis for First Differencing

* NOTE * Zero values of Y_t exist; MAPE calculated only for non-zero Y_t.

Data First Differencing.

Length 144.

Fitted Trend Equation.

$Y_t = 0.432898 − 0.00593010 \times t$.

Accuracy Measures.

MAPE 99.605 MAD 8.875 MSD 153.618.

Model Identification Using Correlogram.

In time series analysis, it is more convenient to use the correlogram table and graph that's Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) for identification of time domain process.

6.4. Correlogram

Date: 10/28/19 Time: 14:54

Sample: 2007M01 2018M12

Included observations: 144

Autocorrelation	Partial Correlation		ACF	PACF	Q-Stat	Prob
.|**** |	.|**** |	1	0.574	0.574	48.465	0.000
.|*** |	.|** |	2	0.480	0.225	82.625	0.000
.|*** |	.|* |	3	0.392	0.078	105.59	0.000
.|*** |	.|* |	4	0.400	0.153	129.66	0.000
.|*** |	.|* |	5	0.439	0.187	158.82	0.000
.|*** |	.|. |	6	0.391	0.036	182.16	0.000

Continued

|						
.|** |	.|. |	7	0.324	−0.029	198.28	0.000
.|** |	.|. |	8	0.289	0.009	211.17	0.000
.|** |	.|* |	9	0.320	0.093	227.08	0.000
.|** |	.|. |	10	0.321	0.037	243.21	0.000
.|** |	.|. |	11	0.317	0.035	259.10	0.000
.|** |	.|. |	12	0.293	0.031	272.80	0.000
.|* |	**|. |	13	0.138	−0.211	275.84	0.000
.|* |	*|. |	14	0.120	−0.070	278.17	0.000
.|* |	.|. |	15	0.087	−0.062	279.41	0.000
.|* |	.|. |	16	0.088	−0.042	280.67	0.000
.|* |	.|. |	17	0.076	−0.020	281.63	0.000
.|. |	*|. |	18	0.002	−0.075	281.63	0.000
.|. |	.|. |	19	−0.008	0.000	281.64	0.000
.|. |	.|. |	20	0.022	0.056	281.72	0.000
.|. |	*|. |	21	−0.036	−0.115	281.95	0.000
.|. |	.|. |	22	−0.028	0.017	282.08	0.000
*|. |	*|. |	23	−0.136	−0.134	285.27	0.000
*|. |	.|. |	24	−0.161	−0.053	289.83	0.000
*|. |	.|. |	25	−0.188	−0.025	296.05	0.000
*|. |	.|* |	26	−0.121	0.089	298.64	0.000
*|. |	.|. |	27	−0.162	−0.037	303.35	0.000
*|. |	.|. |	28	−0.129	0.073	306.36	0.000
*|. |	.|. |	29	−0.134	0.043	309.66	0.000
*|. |	.|. |	30	−0.143	0.023	313.42	0.000
*|. |	.|. |	31	−0.103	0.025	315.39	0.000
*|. |	.|. |	32	−0.128	−0.023	318.48	0.000
*|. |	.|. |	33	−0.123	0.048	321.35	0.000
*|. |	.|* |	34	−0.077	0.106	322.49	0.000
*|. |	*|. |	35	−0.164	−0.098	327.68	0.000
*|. |	.|. |	36	−0.137	−0.006	331.36	0.000
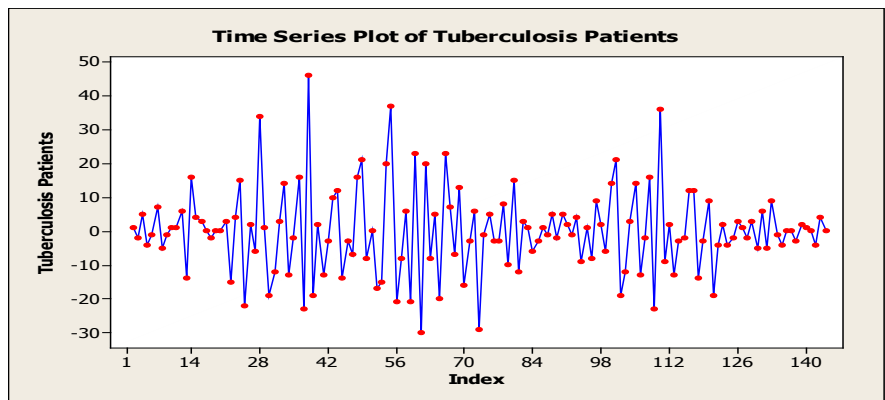
Figure 4. Graph of the first differenced data total population of $N = 143$.
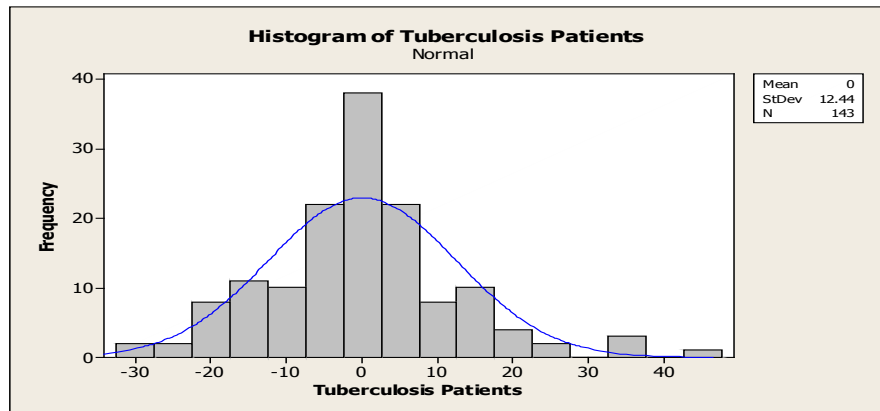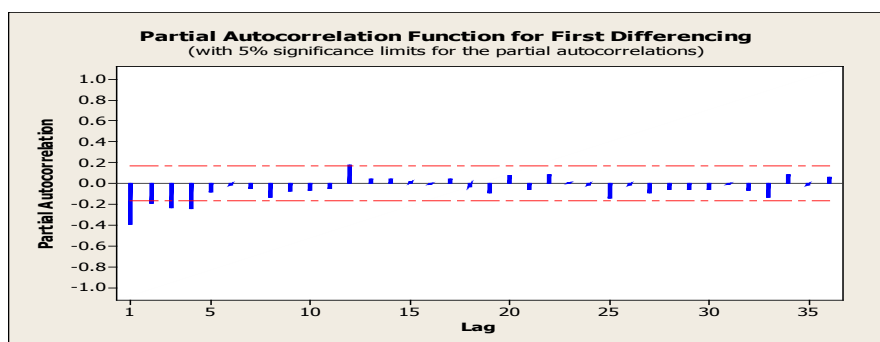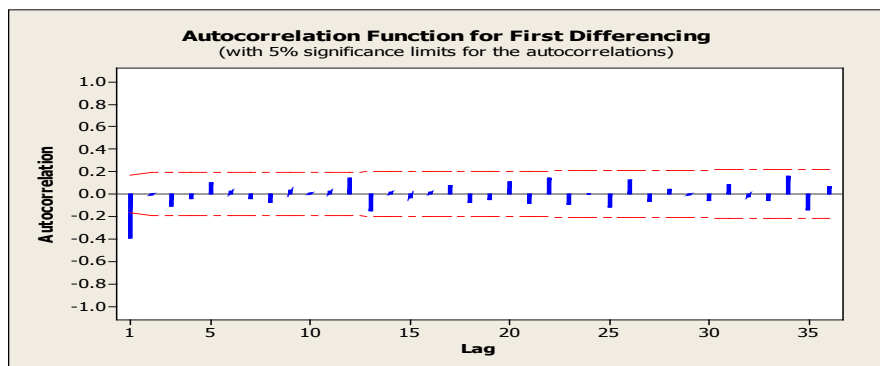
Figure 5. Time plot histogram.

The ACF and PACF from the correlogram above are declining towards zero (0) it shows that there is a trend in the dataset.

The charts below further explained the Autocorrelation and Partial Autocorrelation function of the data at first difference which revealed only one spike at the beginning of plotted data.

7. ARIMA Model Selection

The next stage is now to select the best ARIMA Model (p, d, q) that would be used for forecasting, their goodness of fit has been compared using the following criteria Akaike Information Criteria (AIC) and MAPE. The model with the lowest values would be selected. With careful consideration of all the criteria, ARIMA (2, 1, 3) was found to be the best.

7.1. ARIMA Model: Tuberculosis Patients

Final Estimates of Parameters

Type	Coef SE	Coef	T	P
AR 1	0.6871	0.1227	5.60	0.000
AR 2	−0.8856	0.1602	−5.53	0.000
MA 1	1.3080	0.1486	8.80	0.000
MA 2	−1.2325	0.2599	−4.74	0.000
MA 3	0.5784	0.1438	4.02	0.000
Constant	−0.0140	0.3074	−0.05	0.964

Differencing: 1 regular difference.

Number of observations: Original series 144, after differencing 143.

Residuals: SS = 15456.2 (back forecasts excluded).

MS = 112.8 DF = 137.

Modified Box-Pierce (Ljung-Box) Chi-Square statistic.

Lag	12	24	36	48
Chi-Square	7.0	14.5	26.1	37.4
DF	6	18	30	42
P-Value	0.324	0.694	0.669	0.673

The Observed and Forecasted Values with Its Lowest and Highest Values Obtained by ARIMA (2, 1, 3) Model

Forecasts from period 10

95 Percent Limits

Period	Forecast	Lower	Upper	Actual
11	9.8367	−10.9859	30.6592	11.0000
12	10.1909	−12.0776	32.4593	17.0000
13	10.4246	−12.6993	33.5486	3.0000
14	10.2576	−13.2995	33.8147	19.0000
15	9.9217	−14.1147	33.9582	23.0000
16	9.8250	−15.0761	34.7260	26.0000
17	10.0419	−15.9483	36.0321	26.0000
18	10.2627	−16.5426	37.0680	24.0000
19	10.2083	−17.0815	37.4980	24.0000
20	9.9614	−17.7672	37.6899	24.0000
21	9.8259	−18.5544	38.2063	27.0000
22	9.9376	−19.2956	39.1707	12.0000
23	10.1202	−19.8632	40.1037	16.0000

Continued

|				
24	10.1329	−20.3650	40.6307	31.0000
25	9.9658	−20.9612	40.8928	9.0000
26	9.8258	−21.6392	41.2909	11.0000
27	9.8636	−22.2942	42.0214	5.0000
28	9.9996	−22.8358	42.8350	39.0000
29	10.0455	−23.3131	43.4042	40.0000
30	9.9427	−23.8458	43.7312	21.0000
31	9.8174	−24.4459	44.0807	9.0000
32	9.8084	−25.0390	44.6557	12.0000
33	9.8992	−25.5560	45.3543	26.0000
34	9.9556	−26.0154	45.9265	13.0000
35	9.8999	−26.5024	46.3022	11.0000
36	9.7977	−27.0433	46.6387	27.0000
37	9.7628	−27.5883	47.1140	4.0000
38	9.8154	−28.0823	47.7130	50.0000
39	9.8684	−28.5271	48.2638	31.0000
40	9.8443	−28.9809	48.6695	33.0000
41	9.7668	−29.4754	49.0090	20.0000
42	9.7209	−29.9810	49.4228	17.0000
43	9.7440	−30.4533	49.9413	27.0000
44	9.7865	−30.8850	50.4581	39.0000
45	9.7813	−31.3144	50.8770	25.0000
46	9.7261	−31.7723	51.2245	22.0000
47	9.6788	−32.2443	51.6019	15.0000
48	9.6812	−32.6961	52.0585	31.0000
49	9.7108	−33.1156	52.5372	52.0000
50	9.7150	−33.5262	52.9562	44.0000
51	9.6777	−33.9552	53.3105	44.0000
52	9.6343	−34.3985	53.6671	27.0000
53	9.6236	−34.8308	54.0781	12.0000
54	9.6406	−35.2383	54.5196	32.0000
55	9.6479	−35.6340	54.9298	69.0000
56	9.6237	−36.0400	55.2875	48.0000
57	9.5868	−36.4586	55.6322	40.0000
58	9.5688	−36.8726	56.0102	46.0000
59	9.5751	−37.2683	56.4186	25.0000
60	9.5815	−37.6515	56.8144	48.0000
61	9.5662	−38.0389	57.1714	18.0000

Continued

62	9.5361	−38.4365	57.5088	38.0000
63	9.5150	−38.8334	57.8634	30.0000
64	9.5131	−39.2173	58.2436	35.0000
65	9.5166	−39.5895	58.6226	15.0000
66	9.5066	−39.9619	58.9751	38.0000
67	9.4827	−40.3415	59.3070	45.0000
68	9.4612	−40.7226	59.6449	38.0000
69	9.4535	−41.0949	60.0019	51.0000
70	9.4534	−41.4571	60.3638	35.0000
71	9.4461	−41.8169	60.7090	32.0000
72	9.4272	−42.1813	61.0356	38.0000
73	9.4067	−42.5479	61.3612	9.0000
74	9.3953	−42.9090	61.6997	8.0000
75	9.3917	−43.2617	62.0451	13.0000
76	9.3853	−43.6107	62.3812	10.0000
77	9.3701	−43.9620	62.7022	7.0000
78	9.3514	−44.3156	63.0183	15.0000
79	9.3380	−44.6659	63.3418	5.0000
80	9.3314	−45.0096	63.6723	20.0000
81	9.3247	−45.3491	63.9985	8.0000
82	9.3120	−45.6891	64.3131	11.0000
83	9.2952	−46.0310	64.6213	12.0000
84	9.2809	−46.3711	64.9328	6.0000
85	9.2720	−46.7061	65.2501	3.0000
86	9.2645	−47.0370	65.5660	4.0000
87	9.2533	−47.3672	65.8738	3.0000
88	9.2382	−47.6986	66.1750	8.0000
89	9.2238	−48.0290	66.4766	6.0000
90	9.2133	−48.3558	66.7823	11.0000
91	9.2048	−48.6787	67.0883	13.0000
92	9.1943	−49.0002	67.3889	12.0000
93	9.1806	−49.3222	67.6835	16.0000
94	9.1665	−49.6437	67.9768	7.0000
95	9.1550	−49.9625	68.2725	8.0000
96	9.1455	−50.2780	68.5691	0.0000
97	9.1353	−50.5917	68.8623	9.0000
98	9.1227	−50.9052	69.1505	11.0000
99	9.1090	−51.2184	69.4365	5.0000
100	9.0969	−51.5297	69.7235	19.0000

Looking critically in **Figure 6** below, the black line represents the actual values, the red is the linear trend model and the green line represents the forecasted values.

7.2. Model Evaluation

Model evaluation is an iterative procedure used to access model adequacy by checking whether the model assumptions are satisfied or not in which one of the basic assumptions is that the error term uncorrelated with zero mean and constant variance.

Base on the graph shown below some of the data seems to be out of the line so further test will carryout to ascertain the adequacy of the model.

Normal Probability Plot of the Residuals
(response is Tuberculosis Patients)
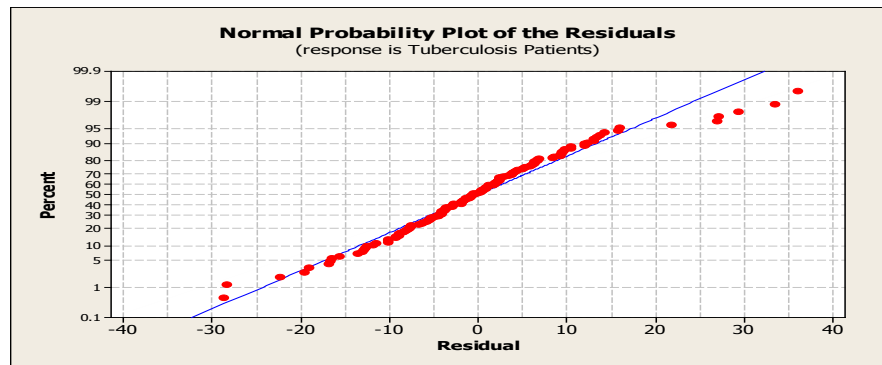
From the graph below it can be seen clearly that the data is normally distributed with pattern less form which shows that the model is a good one. Based on this, there are some degrees of assurance that the data is normally distributed.

8. Summary

This work explicitly explains the concepts of autoregressive integrated moving average (ARIMA). All aim and objectives stated earlier in chapter one was fulfilled. The data were collected from statistics department of general hospital Minna, Niger State based on the monthly records of tuberculosis patients reported in the hospital.

Trend Analysis Plot for C2
Linear Trend Model
$Yt = 24.4544 - 0.0747548*t$

| Variable |
| Actual |
| Fits |
| Forecasts |

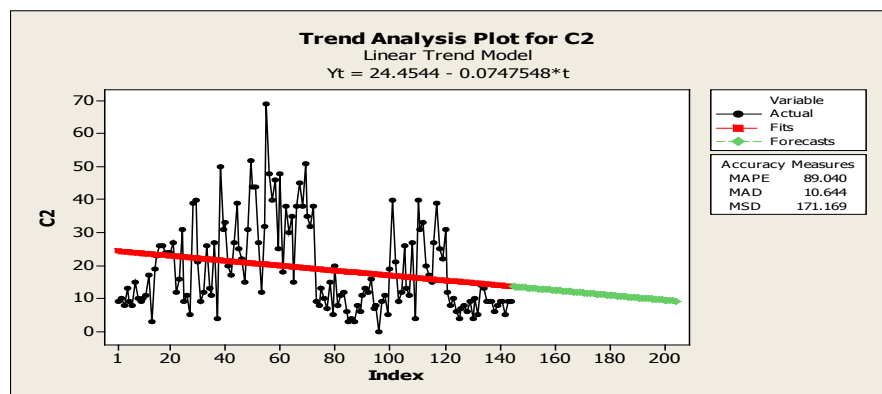| Accuracy Measures	
MAPE	89.040
MAD	10.644
MSD	171.169

Figure 6. Out forecast for 2019-2023.

Exploratory data analysis (EDA) known as descriptive statistics and time plot with the trend of the data were obtained. In the course of conducting an exploratory data analysis it was detected that the data cannot be worked on in it raw form, then a unit root test was conducted which confirmed that the is not stationary with the upward and downward movement which shows the non-stationarity of the data with the help of the histogram and statistics as well. Data estimation were carried out using the least square approach and moving average for the derivation of trend variation (T_t). The unit root test for stationarity was performed using Augmented Dicky Fuller.

(ADF) approach and it was confirmed that the data was non-stationary but later passed through differencing once that is $d = 1$. After conducting data transformation technique likes differencing, the new set of data become stationary in mean confirmed by a unit root test of it differencing shows that the data is stationary at 1%, 5%, and 10% level of significance. Model identification was carryout using the correlogram that is ACF and PACF with graph whereby Akaike Information Criterion was used to determine the order of the ARIMA model.

The model was estimated ARIMA (2, 1, 3) are significant. The model was evaluated using normal probability plot, the residual plot, the histogram, ACF and PACF plot, in which all of them were found to be satisfied at 5% level of significance.

9. Conclusions

From the analysis done in chapter four, the exploratory data analysis (EDA) established that the time plot of the series is fluctuating. This shows that the data is non-stationary with a skew greater than zero and the kurtosis is peak with 0.48 with mean of 19.03 and S.D of 13.49.

The required equation in the trend (T_t) computation for the reported cases of tuberculosis patients in Minna General Hospital by the least square approach was $T_t = 24.4544 + 0.0747548t$ which imply that 24.4544 was the average rate of reported tuberculosis patients between 2007-2018 with 0.07 was the rate of decrease per annum.

It was established that the *p*-value 0.1620 is greater than the alpha level 0.05 that is (0.1620 > 0.05) using the ADF test of stationarity which led to the conclusion of non-stationary data. In this regard, there is a need for differencing which was performed before the stationarity test was conducted again and it was confirmed stationary after differencing once with *p*-value = 0.000 < 0.05 then model estimation was conducted. The correlogram revealed that since the autocorrelation function (ACF) does not decay exponentially zero and the partial autocorrelation function (PACF) did not cut off then ARIMA is suspected with the aid of AIC model. Out of AIC values generated, the minimum value that is stationary and invertible occur at the order (2, 1, 3) of ARIMA which is the best model for the data on monthly basis. It implies that $p = 2$, $d = 1$, $q = 3$ with equation

$$X_t + 0.6867X_{t-1} - 0.8859X_{t-2} = E_t + 1.3077E_{t-1} - 1.2328E_{t-2} + 0.5788E_{t-3}.$$

The model evaluation was carried out to check the time series assumptions in term of violation and satisfaction. Model evaluation shows that none of the time series analysis assumptions is violated at 5% level of significance.

The forecast evaluation or prediction indicated from the graph that if proper measures are put in place by the health workers and individual members of the country, there will be a more decrease in the turn out tuberculosis patients.

Recommendation

Base on the findings and conclusions, it is predicted hypothetically and statistically with the formulated model that there will be a decrease in the number of tuberculosis patients in the next five years (2019-2023). The predicted values are obviously lesser than the values of the years under study (2007-2018). It is therefore recommended that:

- The government through the health provider should increase awareness to the public on the causes and consequence of Tuberculosis.
- The people with compromised immune system should be giving special attention because they are the mostly at risk of contracting the disease.
- Direct observe therapy should be encouraged so the health worker will monitor the administration of drugs.
- More effort should put on education and treatment so it can be totally eradicated.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Sani, R.A., Garba, S.A., Oyeleke, S.B. and Abalaka, M.E. (2015) Prevalence of Pulmonary Tuberculosis (PTB) in Minna and Suleja Niger State, Nigeria. *American Journal of Medicine and Medical Sciences*, **5**, 287-291.

[2] WHO (2016) WHO/HTM/TB/2016.19 Xpert MTB/RIF Assay for the Diagnosis of TB Meeting Report.

[3] World Health Organization (2013) Global Tuberculosis Control: Epidemiology, Planning, Financing. *RIF*, **1**, 218-223.

[4] WHO (2019) WHO Global Tuberculosis Report.
https://www.who.int/tb/publications/global_report/en/

[5] Wani, H. (2013) Prevalence, Pattern, Sensitivity and Resistance to Antibiotics of Different Bacteria Isolated from Port Site Infection in Low Risk Patients after Elective Laparoscopic Cholecystectomy for Symptomatic Cholelithiasis at Tertiary Care Hospital of Kashmir. *International Wound Journal*, **10**, 110-113.
https://doi.org/10.1111/j.1742-481X.2012.00963.x

[6] Abdulhalik, W., Kassahun, W. and Alemseged, F. (2017) Risk Factors for Multidrug-Resistant Tuberculosis among Tuberculosis Patients: A Case-Control Study. *Infection and Drug Resistance*, **10**, 91-96. https://doi.org/10.2147/IDR.S126274

[7] Chuchottaworn, C., Thanachartwet, V., Sangsayunh, P., Than, T.Z.M., Sahassananda, D., Surabotsophon, M., *et al.* (2015) Risk Factors for Multidrug-Resistant Tuberculosis among Patients with Pulmonary Tuberculosis at the Central Chest Institute of Thailand. *PLoS ONE*, **10**, e0139986.
https://doi.org/10.1371/journal.pone.0139986

[8] Diarra, B., Goita, D., Tounkara, S., Sanogo, M., Baya, B., Togo, A.C.G., *et al.* (2016) Tuberculosis Drug Resistance in Bamako, Mali, from 2006 to 2014. *BMC Infectious Diseases*, **16**, Article No. 714. https://doi.org/10.1186/s12879-016-2060-7

[9] Rua-Domenech (2006) Human *Mycobacterium bovis* Infection in the United Kingdom: Incidence, Risks, Control Measures and Review of the Zoonotic Aspects of Bovine Tuberculosis. *Tuberculosis*, **86**, 77-109.
https://doi.org/10.1016/j.tube.2005.05.002

[10] Maiga, M., Siddiqui, S., Diallo, S., Diarra, B., Traore, B. and Shea, Y.R. (2012) Failure to Recognize Nontuberculous Mycobacteria Leads to Misdiagnosis of Chronic Pulmonary Tuberculosis. *PLoS ONE*, **7**, e36902.
https://doi.org/10.1371/journal.pone.0036902

[11] Stevenson, C.R., Forouhi, N.G., Roglic, G., *et al.* (2007) Diabetes and Tuberculosis: The Impact of the Diabetes Epidemic on Tuberculosis Incidence. *BMC Public Health*, **7**, Article No. 234. https://doi.org/10.1186/1471-2458-7-234

[12] Yew, W.W., Van Zyl Smit, R.N., Pai, M., Leung, C.C., Zumla, A., Bateman, E.D. and Dheda, K. (2010) Global Lung Health: The Colliding Epidemics of Tuberculosis, Tobacco Smoking, HIV and COPD. *The European Respiratory Journal*, **35**, 27-33.
https://doi.org/10.1183/09031936.00072909

[13] Borgdorff, M.W., Floyd, K. and Broekmans, J.F. (2002) Interventions to Reduce Tuberculosis Mortality and Transmission in Low- and Middle-Income Countries. *Bulletin of the World Health Organization*, **80**, 217-227.

[14] Gonzalez-Angulo, A.M. and Blumenschein, G.R. (2013) Defining Biomarkers to Predict Sensitivity to PI3K/Akt/mTOR Pathway Inhibitors in Breast Cancer. *Cancer Treatment Reviews*, **39**, 313-320. https://doi.org/10.1016/j.ctrv.2012.11.002

[15] Soliman, H., Gador, A., Lu, Y.H., Lin, G., Bankar, G. and MacLeod, K.M. (2012) Diabetes-Induced Increased Oxidative Stress in Cardiomyocytes Is Sustained by a Positive Feedback Loop Involving Rho Kinase and PKCβ2. *American Journal of Physiology. Heart and Circulatory Physiology*, **303**, H989-H1000.
https://doi.org/10.1152/ajpheart.00416.2012

[16] Hashemi, S.R. and Davoodi, H. (2012) Herbal Plants as New Immuno-Stimulator in Poultry Industry: A Review. *Asian Journal of Animal Veterinary Advances*, **7**, 105-116. https://doi.org/10.3923/ajava.2012.105.116

[17] Brimkulov, N., Ottmani, S.-E., Pio, A., Chubakov, T., Sultanova, A., Davletalieva, N., Kalieva, A., Rittman, J., Erhola, M., Cholurova, R. and Blanc, L. (2009) Feasibility Test Results of the Practical Approach to Lung Health in Bishkek, Kyrgyzstan. *International Journal of Tuberculosis and Lung Disease*, **13**, 533-539.