Sains Malaysiana 52(8)(2023): 2395-2406
http://doi.org/10.17576/jsm-2023-5208-16

Recursive Prediction Model: A Preliminary Application to Lassa Fever Outbreak in Nigeria

(Model Ramalan Rekursif: Aplikasi Awal untuk Wabak Demam Lassa di Nigeria)

FRIDAY ZINZENDOFF OKWONU[1], NOR AISHAH AHAD[2,3,*], HASHIBAH HAMID[3] &
OLIMJON SHUKUROVICH SHARIPOV[4]

[1]Department of Mathematics, Faculty of Science, Delta State University, P.M.B.1, Abraka, Nigeria
[2]Institute of Strategic Industrial Decision Modelling, School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
[3]School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
[4]School of Mathematics, Department of Statistics, National University of Uzbekistan

Received: 10 December 2022/Accepted: 24 July 2023

ABSTRACT

Lassa fever (LF) is endemic in West Africa and Nigeria in particular. Since 1969 when the disease was discovered, a seasonal outbreak is often reported in Nigeria. Many researchers have reported inconsistent or varying numbers of suspected, confirmed and death cases since 2012 to date. To enhance this reportage, and due to the high mortality rate associated with LF, it is pertinent to design a suitable and robust model that could predict or estimate the number of LF cases based on the onset data. To achieve these, we proposed a recursive prediction (RP) model that could do predictions with the onset data. The Pearson correlation coefficient (R), and R^2 are applied to determine the performance analysis of the model. The RP model predicted 96.7% confirmed cases and 89.6% death cases for the first three months of 2022 based on the onset data. The model was also applied to predict COVID-19 death cases during the six weeks of the outbreak in India. The result showed a comparable prediction with the regression output for the COVID-19 death cases. This study demonstrated that the proposed model could be applied to perform prediction for any disease of unknown etiology during the onset of the disease outbreak without any treatment similar to the COVID-19 outbreak. The performance analysis of the RP showed that the model is useful to predict the increasing trend of an outbreak of a disease with unknown etiology without prior treatment experience and vaccines.

Keywords: Case fatality ratio; Lassa fever; prediction; recursive

ABSTRAK

Demam Lassa (LF) ialah endemik khusus di Afrika Barat dan Nigeria. Sejak 1969 apabila penyakit itu ditemui, wabak bermusim sering dilaporkan di Nigeria. Ramai penyelidik telah melaporkan bilangan kes yang disyaki, disahkan dan kematian yang tidak tekal atau berberbeza-beza sejak 2012 hingga kini. Untuk menambahbaik pelaporan ini dan disebabkan oleh kadar kematian yang tinggi yang dikaitkan dengan LF, adalah penting untuk memperkenalkan model yang sesuai dan teguh yang boleh meramalkan atau menganggarkan bilangan kes LF berdasarkan data permulaan. Untuk mencapai tujuan ini, kami mencadangkan model ramalan rekursif (RP) yang boleh melakukan ramalan dengan data permulaan. Pekali korelasi Pearson (R) dan R^2 digunakan untuk menentukan analisis prestasi model. Model RP meramalkan 96.7% kes disahkan dan 89.6% kes kematian untuk tiga bulan pertama 2022 berdasarkan data permulaan. Model itu juga digunakan untuk meramalkan kes kematian COVID-19 selama enam minggu wabak berlaku di India. Hasil menunjukkan ramalan yang setanding dengan modelregresi untuk kes kematian COVID-19. Kajian juga mendapati bahawa model yang dicadangkan boleh digunakan untuk melakukan ramalan bagi mana-mana penyakit yang etiologinya tidak diketahui semasa permulaan wabak penyakit tanpa sebarang rawatan, seperti wabak COVID-19. Analisis prestasi RP mendedahkan bahawa model ini berguna untuk meramalkan peningkatan trend wabak penyakit dengan etiologi yang tidak diketahui tanpa pengalaman rawatan dan vaksin.

Kata kunci: Demam Lassa; nisbah kematian kes; ramalan; rekursif

INTRODUCTION

Lassa fever (LF) was first reported in Lassa, a town in Borno State, Nigeria in 1969 due to the death of two international nurses (Frame et al. 1970). It was also reported in Liberia and Sierra Leone in 1972, respectively. LF is endemic in West Africa and some other African countries. Export of LF to the United States, United Kingdom, Germany, and the Netherlands has been reported in recent years (Akpede et al. 2018; NCDC 2019). In Nigeria, the LF outbreak is considered a public health concern in the affected states. Previously, Lassa fever is often described as a seasonal disease but in recent years, the outbreak is all year round (Okoro et al. 2020).

LF is classified into a viral hemorrhagic fever category with single strands of RNA emanating from the Arenaviridae family. LF is a zoonotic disease found in multi- mammate rats belonging to the Mastomys family (Ajayi et al. 2013; Yaro et al. 2021), morbidity and mortality rate is relatively high (NCDC 2019). The disease is transmitted to humans via exposure to food contaminated with excreta or urine of the infected rats (Richmond & Baglole 2003).

LF can be transmitted from human to human via fluid contact with an infected person. Human to human transmission often occurs in healthcare facilities when healthcare workers without proper personal protective equipment (PPE) come in direct contact with the body fluid of an infected patient (Fisher-Hoch et al. 1995). It has been observed that the degree of transmission rate via sexual intercourse may not be quantified (Richmond & Baglole 2003). People of different age categories are susceptible to the LF virus whose main host is multimammate rats found in rural, semi-urban, and urban areas in many West African countries.

Several studies have focused on the epidemiological (Grace, Egoh & Udensi 2021), agricultural, urbanization, and geographical consideration of the disease, and analysis of previous outbreaks are based on reported cases (Redding et al. 2021). To the best of our knowledge, LF suitable prediction model based on onset data is unavailable. In this study, we proposed a prediction model for the annual onset of Lassa fever outbreak before health alert and subsequent treatment to enable the government and health agencies to make adequate budgetary provisions for the disease since LF has transformed from seasonal to an all-year-round disease in some part of the country.

The development of prediction models is diverse with noticeable strengths and weaknesses (Kattan & Gerds 2020). Some prediction models are universal, univariate, bivariate, and multivariate while some are discipline dependent which means it is only applicable to a specific subject area such as health sciences, medical prognosis, marketing, production, politics, sports betting, and purchase predictions (Collins et al. 2015) and sport science (Zhu & Sun 2020). Prediction models may be described as a combination of model development, validation, and some variants.

Several versions of meta-regression prediction have been applied to analyze different applications in governance and business (Heinemann, Moessinger & Yeter 2018). Simple linear regression (Ghosal et al. 2020; Rath, Tripathy & Tripathy 2020), logistic regression and multiple regression (Laicane et. 2015) have been applied to predict different natural and man-made phenomena regarding climate change and electricity consumption (Braun, Altan & Beck 2014; van Smeden et al. 2021). Various prediction models that depend on previous data to predict the future prices of crude oil, exported and imported commodities using regression, artificial neural networks, and genetic algorithms have been studied (Mahdiani & Khamedchi 2016). Variants of machine learning techniques such as linear and multiple regression have been applied to perform price prediction in the real estate sector (Pérez-Rave, Correa-Morales & González-Echavarría 2019).

The general performance of several known models depends on using a large sample size to validate the model. In such cases, outliers often hamper the performance of these models as a result data cleaning, weighting, deleting, winsorizing or truncating techniques may be applied, and as such vital information may be lost. For instance, the meta-regression analysis like other variants of classical prediction models is hampered by outliers. Thus, trimming the left and right tails ranging from one percentile to twenty percentile was adopted by (Heinemann Moessinger & Yeter 2018) to robustify the model performance.

Many of these models require previous data to predict the likelihood of future outcomes. The prediction precision of these models may be hampered by influential observations or outliers. The proposed model is free from the effects of outliers since it requires onset data to predict the likelihood of future occurrence. This study focused on predicting the future outbreak of diseases of unknown etiology such as Lassa fever which may be seasonal or all year round. This article is organized as follows. Materials and Methods are described in the next section, followed by Results and Discussion subsequently. The conclusion is mentioned in the last section.

Materials and Methods

The data set for this study is secondary data based on onset data reported for the Lassa Fever outbreak in 2022. The data set was input into the SAS/IML software code written using an HP core i3 laptop. Different studies have discussed the application of single variables to perform prediction in clinical certain (Lee, Bang & Kim 2016; Moons et al. 2009). Several prediction model procedures have been adduced (Hemingway et al. 2013; Laupacis, Sekar & Stiell 1997; Steyerberg & Vergouwe 2014; Steyerberg et al. 2013). Various prediction models were proposed based on different variables with well-articulated data. Although, a rule of thumb could be applied regarding data usage (Royston et al. 2009). Even at this, it is worthwhile, to search and apply suitable data to analyze the performance of the model. This is valid based on the concept of data dependency theory (Okwonu et al. 2023). In general, there are no excellent predictive models or data (Lee, Bang & Kim 2016). It can be observed that no prediction model can perform excellently universally. This is verifiable from samples collected from different ethnicity and applied to a particular model, it was observed that the prediction analysis showed overestimation or underestimation (Jee et al. 2014; Liu 2004). Sample size requirements are a vital aspect of statistical modeling. Though, a rule of thumb could be applied to determine the minimum and maximum sample size. Many researchers may favor a large sample size due to the concept of normality and central limit theory. The performance of a multivariable model could be determined by training and testing data set if the data set is large by applying data splitting or the use of external data set to validate the model performance. The proposed recursive prediction model may not require large data set to make a prediction. The model is described as follows.

Recursive Prediction Model

The prediction model is described in Equation (1).

$$Q_i = b_0 + \sum \left(b_{i+1}u_{i+1}, (i+1)[t_{ij}, s_i]\right), \quad i = 0,1,2,\ldots,n, j = 1,2,\ldots,n \tag{1}$$

where b_0 denotes the mean value for the collections of all data for the period under consideration, $b_{i+1} = \frac{\alpha}{b_0}$ denotes the fraction of suspected infected cases, $u_{i+1} = \left(\frac{\alpha}{b_0}\right) \times \Delta$, percentage of suspected cases, $t_{ij} = \frac{\alpha}{b_0}(m)$, i, jdenotes the period of investigation, $s_i = \frac{\alpha^m}{b_0^m}$ denotes the fraction of the onset data, α is the prediction constant. i, j indicates the period of the study, say years, months, weeks, or days as the case maybe subject to the time of the outbreak and time the investigation needs to be carried out. Δ denotes constant within the range of [100].

Therefore, the recursive prediction (RP) model is given in Equation 2.

$$Q_{i+1} = Q_i + \sum\left(b_{i+1}u_{i+1}, (i+1)[t_{ij}, s_i]\right) \tag{2}$$

$$w_i = Q_i - b_0. \tag{3}$$

where Q_i denotes the predicted infection rate based on the initial value (b_0). In other words, b_0 is cumulative for the number of confirmed cases at the start of the study period. Where w_i denotes the difference between the predicted infection rate and the initial infection rate. It helps to determine whether the rate of infection is increasing or decreasing for effective decision measures to be initiated for enhance action plans. A positive w_i value indicates increasing infection rate meanwhile a negative w_i value indicates decreasing rate of infection. Whatever sign directions of w_i would help the authorities to initiate plans to ameliorate the situation. The predicted monthly variation can be determined as follows.

$$\nabla_k = w_{i+1} - w_i, \tag{4}$$

where $k = 1,2,3,4,\ldots,n+1$.

Case Fatality Ratio

The case fatality ratio (CFR) and the monthly cumulative infection ratio can be determined as follows.

$$CFR = \frac{\textit{Predicted death cases}}{\textit{predicted confirmed cases}} \times 100$$

$$\emptyset = \frac{w_i}{Q_i} = \frac{Q_i - b_0}{Q_i} \tag{5}$$

where $\emptyset$ is the predicted infection rate (PIR). The PIR simply means the deviation between the predicted infected cases and the actual infected cases over the predicted infection cases. The value of $\emptyset$ is expressed in percentage. This simply tells us at a glance the percentage of infection based on the initial infection rate.

It could be observed that Equation (1) mimics the following Equation (6) discussed in (Anderson et al. 2018; Zaresani & Scott 2021). That is,

$$V_{it} = g_0 + g_1 c_{it} + \sum_l g_l h_{itl} + \varepsilon_{it} \qquad (6)$$

where V_{it} is a measure of infection prediction; g_0 is a user defined constant and g_l slope; c_{it} is the response rate; h_{itl} is the environmental factors that accelerate the infection rate; and ε_{it} denotes the error of prediction. However, Equation (6) is similar to Equation (1) with V_{it}, g_0, $g_1 c_{it}$ and $\sum_l g_l h_{itl}$ in terms of computational presentation. The main difference between the two equations depends on the nature of the data set. For instance, Equation (1) is data independent which means that prediction could be done based on onset data whereas Equation (6) is data dependent which implies that it requires historical data over a long period. Another difference between Equation (1) and Equation (6) is that Equation (1) is not susceptible to outliers whereas Equation (6) maybe susceptible to outliers. Equation (1) depends on two numerical values, the initial value (b_0) and α.

Based on this model, we could predict the outbreak of diseases of unknown etiology using preliminary data from the source at the onset when no treatment options are available. It could also be applied to predict an unmonitored or uncontrolled sudden all-year-round outbreak of disease such as malaria. The all-year-round outbreak of a disease could be described as a cumbersome modeling procedure with respect to the property values in the real estate sector due to inflation and other natural phenomena (Pérez-Rave, Correa-Morales & González-Echavarría 2019). However, there is no benchmark for best model development and validation, but certain assumptions or conditions must be defined (Lee, Bang & Kim 2016).

The basic assumption are (i) this model encapsulates the fact that the system remains constant and also for epidemic data, the model assumed that the onset of the epidemic is unattended to and a vaccine for the disease is unavailable, (ii) We also assumed that the disease's etiology is unknown, and (iii) The disease's natural host is associated with the human habitat and could not be exterminated or eradicated from the host environment.

The prediction accuracy of this model may be affected if any of the assumptions are violated. This is valid because when any of the assumptions is violated, the data or condition(s) need to be retooled to correspond to the assumptions. This is the best procedure to obtain accurate predictions. The basic limitation of the RP model often occurs when the onset data is not accurately determined before it is used. This implies that the model could predict based on the data provided. Another limitation of this model is the choice of the prediction constant (α) which is user define. The choice of α is very critical in determine the prediction accuracy as such, the value of α should not exceed one. The validity of this model depends on the numerical values of the correlation coefficient. A strong correlation value indicates that the rate of prediction is accurate, the contrary is true. maybe injurious to the infected community, humanity, and it may lead to a serious outbreak. Natural, environmental, and economic factors may alter or limit the prediction accuracy of this model which may enhance or alter human behavioral changes.

CORRELATION BETWEEN PREDICTED Q_i AND INCREMENT w_i

The correlation coefficient (Shrestha 2020) has been applied extensively to investigate the strengths of the relationship between variables of interest. For this study, we modify the conventional formula to correspond to the variable definitions of the recursive prediction model which can be described as

$$R_{Q_i} = \frac{\sum_{i=1}^{n}(Q_i - \bar{Q}_i)(w_i - \bar{w}_i)}{\sqrt{\sum_{i=1}^{n}(Q_i - \bar{Q}_i)^2 \sum_{i=1}^{n}(w_i - \bar{w}_i)^2}} \qquad (7)$$

Equation (7) is applied to determine the linear relationship between the predicted value (Q_i) and the predicted residual (w_i) (Okwonu, Laro Asaju & Irimisose Arunaye 2020). The idea is to determine the strength of the predicted outputs. The coefficient of determination (Okwonu et al. 2021) for the RQi can be obtained easily by taking the square of RQi, that is,

$$R^2 = \left(R_{Q_i}\right)^2, 0 \leq R^2 \leq 1 \qquad (8)$$

RESULTS AND DISCUSSION

From previous research (Africa CDC n.d.; Grace, Egoh & Udensi 2021; Okoro et al. 2020; Onyeji 2022; Redding et al. 2021; Uzoma et al. 2020; Yaro et al. 2021), we observed that data on LF suspected, infected, and death cases vary widely. The data set reported for LF from 2012 to date is not consistent as such some researchers often report sparsely monthly and occasionally annually. Since the LF outbreak is all year round in recent times, it is pertinent for the authorized agencies to regularly update the data. There are many unreported cases of this virus in many localities across Nigeria. Based on the lack of accurate data, it is useful to use onset data to predict the intending outbreak so that government agencies could use the predicted figures to make budgetary provisions to cater for the eventual outbreak. Table 1 contains some data on previous outbreaks.

In Table 1, the data reported by Africa CDC (n.d.), Ebuka Onyeji (2022), Grace, Egoh and Udensi (2021), Musa et al. (2022), Okoro et al. (2020), Redding et al. (2021), Uzoma et al. (2020), and Yaro et al. (2021), are conflicting. It is vital to change the narrative of LF data variation by using prediction models to predict the likelihood of an estimated number of infected and death cases. A disease such as LF with a high case fatality ratio (CFR) requires vital and urgent attention and such should not be criticized to have been overestimated for budgetary, resource allocation, and emergency activation of incidence managers. Prediction models based on onset data would help government agencies and relevant authorities to activate and mitigate any outbreak. The COVID-19 experience is very early to forgo. Figure 1 is a pictorial analysis of reported cases in Table 1 and the 2022 predicted cases.

TABLE 1. Previous data on Lassa fever (LF)

Year	Suspected cases	Confirmed cases	Death
2012	1723	219	--
2013	1195	163	--
2014	989	110	--
2015	428	25	--
2016	918	101	--
2017	1022	322	127
2018	3498	633	171
2019	4099	739(810)	154
2020	6791	1189(1136)	156(244)(234)
2021	3158(4654)	382(510)	102(79)
2022	**3818**	**1291**	**353**

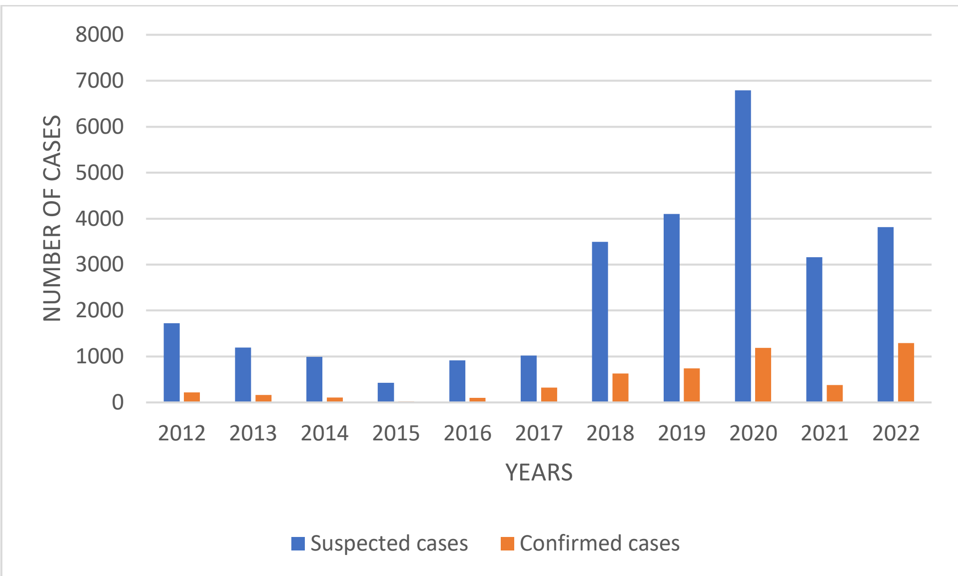

FIGURE 1. Previous outbreak of Lassa fever

CASE I. LASSA FEVER

It was reported that 211 Lasa fever cases were confirmed and 40 reported deaths between 3rd and 30th January 2022 (World Health Organization 2022). In Table 2, we set the prediction probability as 0.95. while in Table 3, the prediction probability was 0.05. This implies that the prediction probability is equal to one for confirmed and dead cases respectively. Before we proceed further, we adopted the conditions and assumptions used in predicting week six death cases as discussed in (Ghosal et al. 2022).

CQ_i denotes delay in diagnosing the virus which may result in a high rate of infection which is the cumulative predicted infected cases. Cw_i denotes early diagnosis which may reduce the rate of virus infection. $C\nabla_k$ denotes early diagnosis and an extremely high rate of control and monitoring would effectively reduce the rate of spread of the virus.

Based on the onset data and confirmed cases in March by (NCDC 2022), CQ_i best described the situation because as of March 2022, about 630 cases have been confirmed to have been infected by the virus. Figure 2 showed the comparison between the cumulative and monthly predicted LF outbreak for 2022.

TABLE 2. Predicted Lassa fever outbreak for 2022

Months	Q_i	CQ_i	w_i	Cw_i	$\emptyset$	∇_k	$C\nabla_k$
0	211.2027	211	0.2027	0.2027	0.0959	4.5069	4.5069
1	215.7096	426	4.7096	4.912	2.1833	9.0137	13.52
2	224.7233	650**(630)**	13.7233	18.64	6.1068	13.5206	27.04
3	238.244	888	27.2439	45.88	11.4353	18.0275	45.07
4	256.2714	1144	45.2714	91.15	17.6654	22.5344	67.60
5	278.8058	1422	67.8058	158.96	24.3201	27.0412	94.64
6	305.847	1727	94.8470	253.80	31.0113	31.5481	126.20
7	337.3951	2064	126.3951	380.20	37.4620	36.0549	162.25
8	373.4501	2437	162.4501	542.65	43.4998	40.5619	202.81
9	414.012	2851	203.012	745.66	49.0353	45.0688	247.88
10	459.081	3310	248.081	993.67	54.0386	49.5756	297.50
11	508.656	3818	297.656	1291	58.5181	--	--

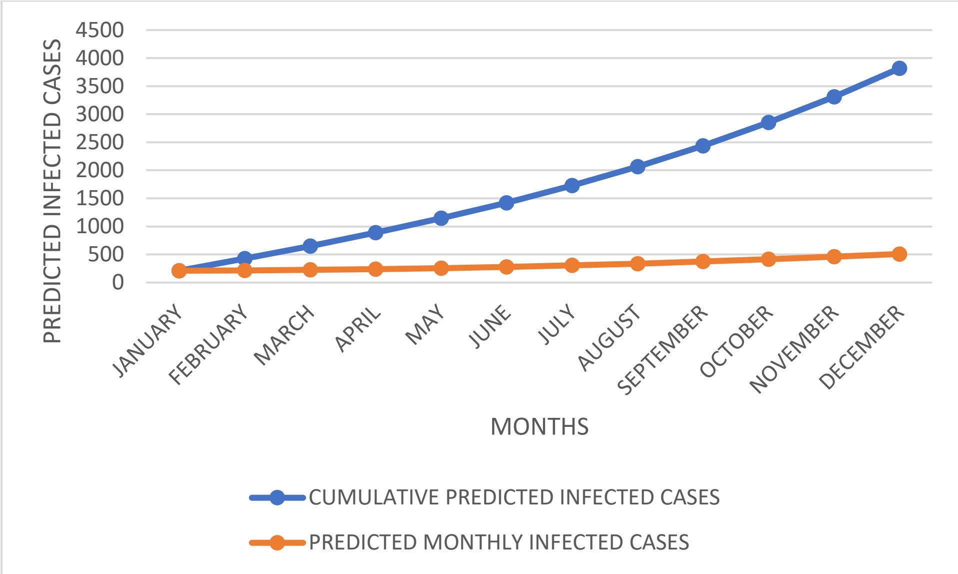

FIGURE 2. Cumulative analysis of predicted LF infected cases for 2022

TABLE 3. Predicted number of Lassa Fever Death 2022

Months	Q_i	$C\hat{Q}_i$	$C\hat{w}_i$	w_i	$\emptyset$	∇_k	$C\hat{\nabla}_k$	CFR (%)
0	40.0156	40	0.0156	-	0.0390	1.2513	1.25	18.9465
1	41.2669	81	1.2669	1.27	3.0699	2.5025	3.75	19.1308
2	43.7694	124**(112)**	3.7694	5.03	8.6119	3.7538	7.51	19.4770
3	47.5231	171	7.5231	12.56	15.8305	5.005	12.51	19.9472
4	52.5281	223	12.5281	25.10	23.8503	6.2563	18.77	20.4971
5	58.7844	281	18.7844	43.88	31.9547	7.5075	26.28	21.0843
6	66.2919	347	26.2919	70.20	39.6608	8.7588	35.00	21.6748
7	75.0506	422	35.0506	105.22	46.7026	10.01	45.00	22.2441
8	85.0606	507	45.0606	150.28	52.9747	11.2613	56.31	22.7769
9	96.3219	603	56.3219	206.60	58.4726	12.5125	68.82	23.2655
10	108.8344	711	68.8344	275.43	63.2469	13.7638	**82.58**	23.7070
11	122.5981	**833**	82.5981	**358.00**	67.3731			24.1024

$C\hat{Q}_i$ denotes delay in diagnosis and treatment which may result in a high mortality rate. $C\widehat{w}_i$ denotes early diagnosis and delay in effective treatment would result in a relatively high mortality rate. $C\widehat{\nabla}_k$ denotes early diagnosis and effective treatment would reduce the mortality rate.

From this study, we observed that the predicted case fatality ratio (CFR) agrees with the previous results reported (Musa et al. 2022; Uzoma et al. 2020; Akpede et al. 2018). Figure 3 showed exactly what may occur if the statement of $C\hat{Q}_i$ persist.

CASE II. COVID-19 DEATH CASES IN INDIA

The data used in **Case II** was the outcome of the study conducted based on death associated with COVID-19 from the onset in India using simple linear regression analysis. The number of death cases was reported from day zero to 14th March 2020. Relying on the predicted outcome from the regression analysis, the average number of week five death cases was 211 (Ghosal et al. 2020). The regression model predicted 467 deaths for week six based on week five death cases. They concluded that the probability of overestimation is assured and is suitable for an epidemic with unknown etiology and no vaccines. In this study, we applied the output from week five to predict the death cases for week six. We also accepted all the study conditions discussed in (Ghosal et al. 2020) regarding the data set until week five.

Based on the average weekly cases reported in Ghosal et al. (2020), the proposed model was able to predict the average number of death cases for week six to be 466.86 which is approximately 467, respectively. Figure 3 shows the number of days after week five death cases reported (Ghosal et al. 2020). From Figure 3, starting for day 1 through day 7 in the horizontal axis, which is week six, we observed that the death cases continue to increase if we rely on the conditions for the study. This showed that at the onset of the epidemic of unknown etiology, the spread of the virus will continue rapidly if we rely on the regression prediction and the RP model. The RP model cumulatively predicted that at the end of week six, an estimated 2460 deaths cases would be expected.

The analysis in Tables 2 to 4 showed that Q_i and w_i are perfectly correlated, that is $R_{Q_i}=1$ and $R^2=1$. This demonstrated that the model's correlation value mimics $r = 0.96$, $r^2 = 0.94$ reported in Ghosal et al. (2020).. The perfect correlation from the proposed model may be due to the near normality of the predicted infected and death cases.

Based on the 211 average cumulative reported Lassa fever confirmed cases by NCDC in January 2022, if the virus natural host is not properly eliminated, monitored, and controlled, the model predicted an increase of infected cases as shown in Tables 2 to 4 and Figures 2 to 4. The correlation analysis showed a strong positive correlation between the predicted value and the predicted monthly infection rate. The analysis

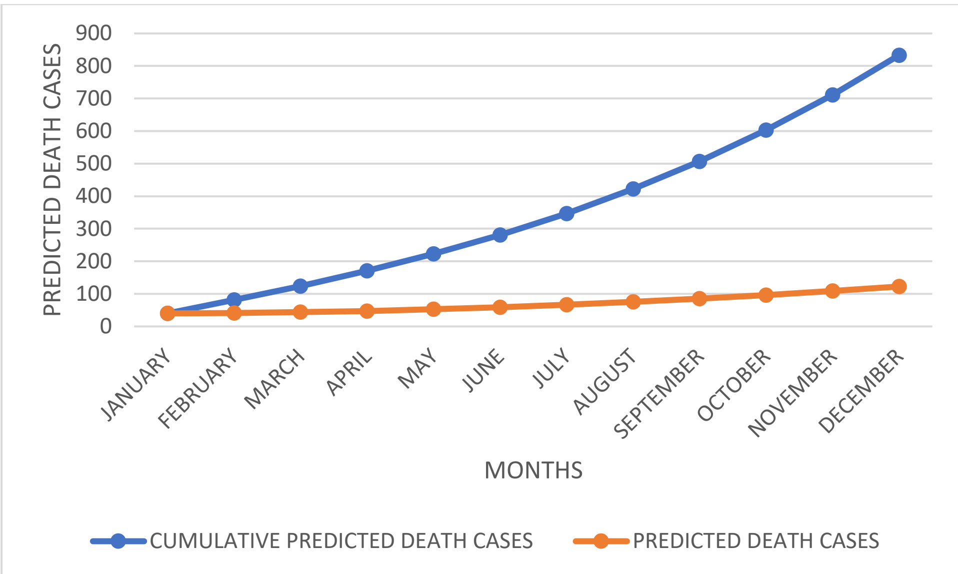

FIGURE 3. Cumulative analysis of predicted LF death cases for 2022

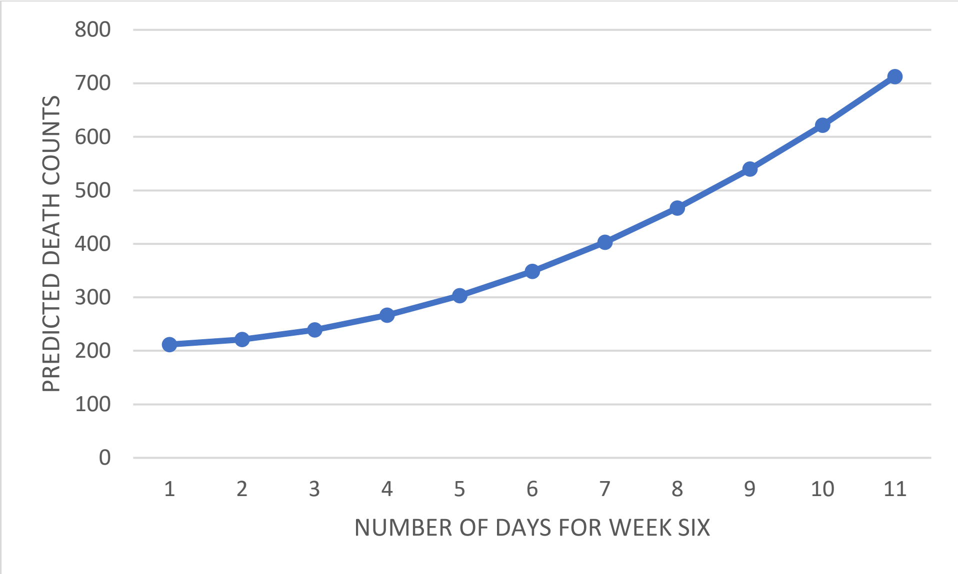

FIGURE 4. Prediction of COVID-19 death count for week six using RP model

TABLE 4. Predicted COVID-19 death count in India 2020 for week 6

Days				
0	211.8280	0.8280	0.39089	9.1086
1	220.9366	9.9366	4.49759	18.2173
2	239.1539	28.1539	11.7723	27.3259
3	266.4798	55.4798	20.8195	36.4345
4	302.9143	91.9142	30.3433	45.5431
5	348.4574	137.4574	39.4474	54.6518
6	403.1092	192.1092	47.6569	63.7604
7	**466.8695**	**255.8695**	54.8054	**72.8690**
8	539.7385	328.7385	60.9069	81.9776
9	621.7162	410.7162	66.0617	91.0863
10	712.8024	501.8024	70.3985	--

demonstrated that the rate of Lassa fever infection may increase continuously if the outbreak during onset was not properly controlled and managed. At present, the rate of infection will outpace previous outbreaks if incidence managers and the respective agencies lag in mitigating the spread of the virus. From recent research, the seasonal outbreak is no longer a holiday task for LF in endemic states rather it should be considered a year-round disease since the natural host coexist with humans, especially in rural and semi-urban areas.

The proposed RP model predicted 650 cumulative infected cases in Table 2 and 124 death cases in Table 3 in March 2022 which relatively correspond to 630 and 112 cumulative infected and death cases reported on 18th March 2022 (NATHNaC 2022; NCDC 2022). This implies that the proposed model predicted 96.9% cumulative infected cases and 89.6% cumulative death cases in March 2022. This study predicted that the cumulative rate of LF infection for 2022 ranges between 1291 and 3818 cases and the cumulative death rate may range between 358 and 833 if proper monitoring and control are not initiated promptly to mitigate the spread of the virus. This study demonstrated that the case fatality ratio (CFR) of LF is higher than the CFR for COVID-19. Based on this observation, it is recommended that proper monitoring, contact tracing, treatment, budgetary allocation, availability of personal protective equipment for healthcare workers, and development of a vaccine for LF be seriously advocated by the endemic LF countries, Africa Union (AU) and World Health Organization (WHO).

Conclusion

The recursive prediction model showed that any disease with unknown etiology during the onset may likely transform into a full-scale outbreak. The finding showed that based on onset data, the model predicted an increasing trend of infection which may if not controlled and monitored may lead to a major outbreak. The outcome indicated that this type of model could be very useful to predict the outbreak of diseases such as LF. The correlation value for the RP predicted infection rate indicates a very strong positive correlation signally the prediction accuracy of the model based on the onset data. The comparative prediction results for the RP model using the LF and the COVID 19 data validated the accuracy of the model. Relying on onset data, this model would aid and enable incident managers and agencies to make proper budgetary allocation and rapid response to curtail the eventual outbreak. From the study, the case fatality ratio for LF is very high as such endemic countries should advocate for the development of LF vaccine to mitigate infection and death rates. Therefore, before generalizing the outcome of the model, the user should try to determine the strength of the relationship between the predicted values, in this case, a very strong linear relationship indicates that the choice of the prediction constant and the predicted values are satisfying the assumptions. This study infers that the degree of prediction and the value of the correlation should be taken seriously to mitigate the outbreak of the disease. This study concurs that a strong positive association which corresponds to increasing infection rate must be taken necessary to avert the disease progression to an epidemic level. Therefore, the rate of infection and recovery should be a vital tool to determine if the outbreak is under control. On the other hand, if the rate of infection is increasing and the rate of recovery is decreasing, the level of association will be weak and as such the incident managers should act swiftly to retool their interventions procedures otherwise the outcome may be injurious to the infected community, humanity, and it may lead to a serious outbreak. Natural, environmental, and economic factors may alter or limit the prediction accuracy of this model which may enhance or alter human behavioral changes. Hence the study concludes that the researcher should comply with the assumptions to obtain accurate prediction values in other to mitigate the escalation of the outbreak at the onset of the disease emergence.

Acknowledgements

The authors want to express our gratitude to Delta State University and Universiti Utara Malaysia for supporting this research work. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Africa CDC. (n.d.). *https://africacdc.org/disease/lassa-fever/*

Ajayi, N.A., Nwigwe, C.G., Azuogu, B.N., Onyire, B.N., Nwonwu, E.U., Ogbonnaya, L.U., Onwe, F.I., Ekaete, T., Günther, S. & Ukwaja, K.N. 2013. Containing a Lassa fever epidemic in a resource-limited setting: outbreak description and lessons learned from Abakaliki, Nigeria (January-March 2012). *International Journal of Infectious Diseases* 17(11). e1011-e1016. https://doi.org/10.1016/j.ijid.2013.05.015

Akpede, G.O., Asogun, D.A., Okogbenin, S.A. & Okokhere, P.O. 2018. Lassa fever outbreaks in Nigeria. *Expert Review of Anti-Infective Therapy* 16(9): 663-666. https://doi.org/10.1080/14787210.2018.1512856

Anderson, E., d'Orey, M.A.J., Duvendack, M. & Esposito, L. 2018. Does government spending affect income poverty? A meta-regression analysis. *World Development* 103: 60-71. https://doi.org/10.1016/j.worlddev.2017.10.006

Braun, M.R., Altan, H. & Beck, S.B.M. 2014. Using regression analysis to predict the future energy consumption of a supermarket in the UK. *Applied Energy* 130: 305-313. https://doi.org/10.1016/j.apenergy.2014.05.062

Collins, G.S., Reitsma, J.B., Altman, D.G. & Moons, K.G.M. 2015. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *European Urology* 350: g7584. https://doi.org/10.1016/j.eururo.2014.11.025

Fisher-Hoch, S.P., Tomori, O., Nasidi, A., Perez-Oronoz, G.I., Fakile, Y., Hutwagner, L. & McCormick, J.B. 1995. Review of cases of nosocomial Lassa fever in Nigeria: The high price of poor medical practice. *BMJ (Clinical Research Ed.)* 311(7009): 857-859. https://doi.org/10.1136/bmj.311.7009.857

Frame, J.D., Baldwin, J.M., Gocke, D.J. & Troup, J.M. 1970. Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings. *The American Journal of Tropical Medicine and Hygiene* 19(4): 670-676. https://doi.org/10.4269/ajtmh.1970.19.670

Ghosal, S., Sengupta, S., Majumder, M. & Sinha, B. 2020. Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020). *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 14(4): 311-315. https://doi.org/10.1016/j.dsx.2020.03.017

Grace, J.U.A., Egoh, I.J. & Udensi, N. 2021. Epidemiological trends of Lassa fever in Nigeria from 2015-2021: A review. *Therapeutic Advances in Infectious Disease* 8: 20499361211058252. https://doi.org/10.1177/20499361211058252

Heinemann, F., Moessinger, M.D. & Yeter, M. 2018. Do fiscal rules constrain fiscal policy? A meta-regression-analysis. *European Journal of Political Economy 51*. https://doi.org/10.1016/j.ejpoleco.2017.03.008

Hemingway, H., Croft, P., Perel, P., Hayden, J.A., Abrams, K., Timmis, A., Briggs, A., Udumyan, R., Moons, K.G.M., Steyerberg, E.W., Roberts, I., Schroter, S., Altman, D.G. & Riley, R.D. 2013. Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ* 346: e5595. https://doi.org/10.1136/bmj.e5595

Jee, S.H., Jang, Y., Oh, D.J., Oh, B.H., Lee, S.H., Park, S.W., Seung, K.B., Mok, Y., Jung, K.J., Kimm, H., Yun, Y.D., Baek, S.J., Lee, D.C., Choi, S.H., Kim, M.J., Sung, J., Cho, B., Kim, E.S., Yu, B.Y., Lee, T-Y., Kim, J.S., Lee, Y-J., Oh, J-K., Kim, S.H., Park, J-K., Koh, S.B., Park, S.B., Lee, S.Y., Yoo, C-I., Kim, M.C., Kim, H-K., Park, J-S., Kim, H.C., Lee, G.J. & Woodward, M. 2014. A coronary heart disease prediction model: the Korean Heart Study. *BMJ Open* 4(5): e005025. https://doi.org/10.1136/bmjopen-2014-005025

Kattan, M.W. & Gerds, T.A. 2020. A framework for the evaluation of statistical prediction models. In *CHEST* 158(1 Supplement): S29-S38. https://doi.org/10.1016/j.chest.2020.03.005

Laicane, I., Blumberga, D., Blumberga, A. & Rosa, M. 2015. Comparative multiple regression analysis of household electricity use in Latvia: Using smart meter data to examine the effect of different household characteristics. *Energy Procedia* 72: 49-56. https://doi.org/10.1016/j.egypro.2015.06.008

Laupacis, A., Sekar, N. & Stiell, I.G. 1997. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA* 277(6): 488-494.

Lee, Y.H., Bang, H. & Kim, D.J. 2016. How to establish clinical prediction models. *Endocrinology and Metabolism* 31(1): 38-44. https://doi.org/10.3803/EnM.2016.31.1.38

Liu, J. 2004. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. *JAMA* 291(21): 2591-2599. https://doi.org/10.1001/jama.291.21.2591

Mahdiani, M.R. & Khamehchi, E. 2016. A modified neural network model for predicting the crude oil price. *Intellectual Economics* 10(2): 71-77. https://doi.org/10.1016/j.intele.2017.02.001

Moons, K.G.M., Royston, P., Vergouwe, Y., Grobbee, D.E. & Altman, D.G. 2009. Prognosis and prognostic research: What, why, and how? *BMJ* 338: b375. https://doi.org/10.1136/bmj.b375

Musa, S.S., Zhao, S., Abdullahi, Z.U., Habib, A.G. & He, D. 2022. COVID-19 and Lassa fever in Nigeria: A deadly alliance? *International Journal of Infectious Diseases* 117: 45-47. https://doi.org/10.1016/j.ijid.2022.01.058

NATHNaC. 2022. March 18. *Lassa Fever in Nigeria*. https://Travelhealthpro.Org.Uk/Updates.Php?Base=1599.

NCDC. 2022, April 1. *An Update of Lassa Fever Outbreak in Nigeria*. https://Ncdc.Gov.Ng/Diseases/Sitreps/?Cat=5&name=An%20update%20of%20Lassa%20fever%20outbreak%20in%20Nigeria

NCDC 2019. *Lassa Fever*. https://www.ncdc.gov.ng/diseases/info/L.

Okoro, O.A., Bamgboye, E., Dan-Nwafor, C., Umeokonkwo, C., Ilori, E., Yashe, R., Balogun, M., Nguku, P. & Ihekweazu, C. 2020. Descriptive epidemiology of Lassa fever in Nigeria, 2012-2017. *Pan African Medical Journal* 37(15). https://doi.org/10.11604/pamj.2020.37.15.21160

Okwonu, F.Z., Ahad, N.A., Hamid, H., Muda, N. & Sharipov, O.S. 2023. Enhanced robust univariate classification methods for solving outliers and overfitting problems. *Journal of Information Communication Technology* 22(1): 1-30.

Okwonu, F.Z., Ahad, N.A., Apanapudor, J.S. & Arunaye, F.I. 2021. Robust multivariate correlation techniques: A confirmation analysis using COVID-19 data set. *Pertanika Journal of Science and Technology* 29(2): 999-1015. https://doi.org/10.47836/pjst.29.2.16

Okwonu, F.Z., Laro Asaju, B. & Irimisose Arunaye, F. 2020. Breakdown analysis of Pearson correlation coefficient and robust correlation methods. *IOP Conference Series: Materials Science and Engineering* 917: 012065. https://doi.org/10.1088/1757-899X/917/1/012065

Onyeji, E. 2022. *In 2021, Nigeria Witnessed Decline in Lassa Fever Infections, Deaths – Official*. 10 January 2022. https://www.premiumtimesng.com/news/more-news/505001-in-2021-nigeria-witnessed-decline-in-lassa-fever-infections-deaths-official.html?tztc=1

Pérez-Rave, J.I., Correa-Morales, J.C. & González-Echavarría, F. 2019. A machine learning approach to big data regression analysis of real estate prices for inferential and predictive purposes. *Journal of Property Research* 36(1): 59-96. https://doi.org/10.1080/09599916.2019.1587489

Rath, S., Tripathy, A. & Tripathy, A.R. 2020. Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews* 14(5): 1467-1474. https://doi.org/10.1016/j.dsx.2020.07.045

Redding, D.W., Gibb, R., Dan-Nwafor, C.C., Ilori, E.A., Yashe, R.U., Oladele, S.H., Amedu, M.O., Iniobong, A., Attfield, L.A., Donnelly, C.A., Abubakar, I., Jones, K.E. & Ihekweazu, C. 2021. Geographical drivers and climate-linked dynamics of Lassa fever in Nigeria. *Nature Communications* 12: 5759. https://doi.org/10.1038/s41467-021-25910-y

Richmond, J.K. & Baglole, D.J. 2003. Lassa fever: Epidemiology, clinical features, and social consequences. *BMJ (Clinical Research)* 327(7426): 1271-1275. https://doi.org/10.1136/bmj.327.7426.1271

Royston, P., Moons, K.G.M., Altman, D.G. & Vergouwe, Y. 2009. Prognosis and prognostic research: Developing a prognostic model. *BMJ* 338: b604. https://doi.org/10.1136/bmj.b604

Shrestha, N. 2020. Detecting multicollinearity in regression analysis. *American Journal of Applied Mathematics and Statistics* 8(2): 39-42. https://doi.org/10.12691/ajams-8-2-1

Steyerberg, E.W. & Vergouwe, Y. 2014. Towards better clinical prediction models: Seven steps for development and an ABCD for validation. *European Heart Journal* 35(29): 1925-1931. https://doi.org/10.1093/eurheartj/ehu207

Steyerberg, E.W., Moons, K.G.M., van der Windt, D.A., Hayden, J.A., Perel, P., Schroter, S., Riley, R.D., Hemingway, H. & Altman, D.G. 2013. Prognosis Research Strategy (PROGRESS) 3: Prognostic model research. *PLoS Medicine* 10(2): e1001381. https://doi.org/10.1371/journal.pmed.1001381

Uzoma, I.R., Dan-Nwafor, C., Ipadeola, O., Rimamdeyati, Y., Iniobong, A., Okoro, B., Namara, G., Uzoma, E., Ilori, E. & Ihekweazu, C. 2020. Increasing trend of Lassa fever outbreak in Nigeria: The more you look, the more you see. *International Journal of Infectious Diseases* 101. https://doi.org/10.1016/j.ijid.2020.09.936

van Smeden, M., Reitsma, J.B., Riley, R.D., Collins, G.S. & Moons, K.G. 2021. Clinical prediction models: Diagnosis versus prognosis. In *Journal of Clinical Epidemiology* 132: 142-145. https://doi.org/10.1016/j.jclinepi.2021.01.009

World Health Organization. 2022. Disease Outbreak News; Lassa Fever-Nigeria. https://www.who.int/emergencies/disease-outbreak-news/item/lassa-fever---nigeria. *Lassa Fever - Nigeria*

Yaro, C.A., Kogi, E., Opara, K.N., Batiha, G.E.S., Baty, R.S., Albrakati, A., Altalbawy, F.M.A., Etuh, I.U. & Oni, J.P. 2021. Infection pattern, case fatality rate and spread of Lassa virus in Nigeria. *BMC Infectious Diseases* 21(1): 149. https://doi.org/10.1186/s12879-021-05837-x

Zaresani, A. & Scott, A. 2021. Is the evidence on the effectiveness of pay for performance schemes in healthcare changing? Evidence from a meta-regression analysis. *BMC Health Services Research* 21: 175. https://doi.org/10.1186/s12913-021-06118-8

Zhu, P. & Sun, F. 2020. Sports athletes' performance prediction model based on machine learning algorithm. *Advances in Intelligent Systems and Computing* 1017: 498-505. https://doi.org/10.1007/978-3-030-25128-4_62

*Corresponding author; email: aishah@uum.edu.my