Alexandria Engineering Journal (2021) 60, 4547–4554

HOSTED BY

ELSEVIER

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Estimation of COVID-19 under-ascertainment in Kano, Nigeria during the early phase of the epidemics

Salihu S. Musa [a,b], Shi Zhao [c,d], Nafiu Hussaini [e], Zian Zhuang [a], Yushan Wu [c], Abdurrahman Abdulhamid [f], Maggie H. Wang [c,d], Daihai He [a,*]

[a] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China*
[b] *Department of Mathematics, Kano University of Science and Technology, Wudil, Nigeria*
[c] *JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China*
[d] *Shenzhen Research Institute of the Chinese University of Hong Kong, Shenzhen, China*
[e] *Department of Mathematical Sciences, Bayero University, Kano, Nigeria*
[f] *Department of Statistics, Kano State Polytechnic, Nigeria*

Received 20 January 2021; revised 25 February 2021; accepted 3 March 2021
Available online 16 March 2021

KEYWORDS

SARS-CoV-2;
COVID-19;
Statistical modelling;
Reproduction number;
Under-ascertainment;
Epidemic

Abstract This study aims to estimate the number of COVID-19 cases under- ascertained (η), and the basic reproduction number (R_0) during the early stage of epidemic in Kano, Nigeria. We adopt a simple exponential growth model to capture the patterns of COVID-19 early epidemic curve in Kano. The R_0 is estimated at 2.7 (95%CI: 2.5, 3.0). We find that the number of COVID-19 cases under-ascertained likely occurred during the fourth week of April 2020, and should be considered for future epidemiological investigations and mitigation plan.

1. Introduction

Coronaviruses are a group of related Ribonucleic acid (RNA viruses) from the family of *Coronaviridae* and the order *Nidovirales* that cause diseases in humans and animals (da Costa, 2020). These viruses belong to the genus betacoronaviruses which largely affected human beings in the past two decades [1,2]. They include the Severe Acute Respiratory Syndrome

Abbreviations: R_0, Basic reproduction number; SI, Serial interval; CI, Confidence interval; η, Number of COVID-19 cases under-ascertained; C_i, Cumulative number of cases; τ_i, Cumulative unreported cases; γ, Intrinsic growth rate; α_i, Summation of the cumulative reported cases; ℓ, log-likelihood estimation; $h(\phi)$, Probability distribution for the serial interval; $h(k)$, Lognormal distribution; β_i, daily number of cases; p, Autoregressive order; d, Degree of differencing; q, Moving average order; AIC, Akaike information criterion; ACF, Auto-correlation function; PACF, Partial autocorrelation function.

* Corresponding author.
E-mail address: daihai.he@polyu.edu.hk (D. He).

Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2021.03.003
1110-0168

Coronavirus (SARS-CoV) which appeared in 2002 and has affected over 8000 people with 774 associated deaths; the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) which emerged in 2012 and has affected over 2500 people with 886 associated deaths; and of recent, the coronavirus disease 2019 (COVID-19) which started in late 2019 and (by 20th February 2021) has affected over 110 million people with more than 2.4 million associated deaths (and still ongoing) worldwide [1,3]. Previous studies revealed that most people infected with coronaviruses show no symptoms (asymptomatically infected) or mild symptoms (symptomatically infected) [2,4]. The global spread of COVID-19 (which is caused by SARS-CoV-2) started in Wuhan, Hubei province of China, and has rapidly disperse to over 210 countries and territories, continues to induce catastrophic public health and socio-economic affliction globally. On 30th January 2020, the World Health Organization (WHO) declared COVID-19 as a Public Health Emergency of International Concern [5], and by 11th March 2020, it has been declared the first pandemic caused by the coronavirus [6].

Most clinical manifestations of COVID-19 infection are similar to that of the other two coronaviruses of the same family. The symptoms include fever, headache, myalgia, diarrhoea, dry cough, nausea, chest pain, fatigue, and dyspnea [2,4]. And in severe cases, symptomatically infected people usually present moderate-to-severe respiratory symptoms that often progress to severe pneumonia [7]. Unlike SARS-CoV and MERS-CoV infections, few COVID-19 infected patients show prominent upper respiratory tract signs and symptoms such as sneezing or sore throat, indicating that SARS-CoV-2 might have a greater preference for infecting the lower respiratory tract [2]. Also, severe complications such as arrhythmia, hypoxemia, acute cardiac injury, acute ARDS, acute kidney injury, and shock have been reported among COVID-19 patients [2,8]. The role played by the Huanan Seafood Wholesale Market in the transmission of SARS-CoV-2 during the early outbreaks is yet to be fully uncovered [9]. A genomic study highlights that SARS-CoV-2 was initially introduced into the market from an unknown place (yet to be verified) with human-to-human transmission likely occurred earlier than reported [9]. The disease spreads very rapidly, starting from China (the first epicenter) including clusters of cases from individuals and health workers, and later spreads to all parts of the world [3,9]. Human-to-human transmission usually occurs via respiratory droplets from an infected person when he/she coughs or sneezes. SARS-CoV-2 can stay on the surface of a contaminated object for relatively a long time (or hours) indicating fomites likely being the main source of transmission [9–11].

Nigeria reported its 1st case of COVID-19 on 27th February 2020 [12,13]. The patient is an Italian citizen who came from Milan through the Murtala Muhammad Airport, Lagos. The index case was confirmed by the Nigeria Center for Disease Control (NCDC) following the case definition of SARS-CoV-2 [12]. That is, a COVID-19 patient is an individual with laboratory confirmation of SARS-CoV-2 infection detected using reverse transcription-polymerase chain reaction (RT-PCR) test [12]. SARS-CoV-2 ribonucleic acid (RNA) can also be detected using a molecular amplification detection test [14,15]. During the early epidemic period, community transmission of COVID-19 in Nigeria occurs very fast likely due to the lack of effective contact tracing and delay of border/school closure as part of the proactive measures against the spread of the virus, that helps to prevent further outbreaks. The fact that the management and control of the pandemic depend largely on a country's health care system. By 20th February 2021, there were 150,908 and 1,813 cases and deaths, respectively, in Nigeria [12]. Currently, Nigeria is observing the second wave of the pandemic which is rapidly increasing. Low quality of health care services such as unsafe clinical facilities and practices is one of the cogent reasons for the rapid spread of the virus in Nigeria. This makes the country more vulnerable to COVID-19 infection [13]. Nigeria has the highest population in Africa with over 200 million individuals. Kano state is one of the epicenters in Nigeria during the early epidemic phase and reported its first COVID-19 case on 11th April 2020. Since then, there might have been a problem in reporting the number of cases that are likely under-ascertained from 22 to 27 April 2020.

The under-ascertainment of cases for infectious diseases usually occurs when surveillance systems fail to capture case scenarios at two distinct levels of the surveillance pyramid from a population since many infected persons do not consider a medical investigation. There are several methods to estimate the extent of under-ascertainment for diseases in a population, some of which will be discussed in the current study [16]. The scenario of under-ascertainment or under-reporting for COVID-19 cases during the current pandemic situation is an important issue in terms of individuals and public health as well as socio-economic growth. Thus, it needs urgent attention by researchers considering its impact on disease spread and the role it plays in prevention and control strategy. The possible existence of COVID-19 under-ascertainment in Kano was plausibly due to the provision of inadequate facilities in the health sector (such as facemasks, test kits, and gowns, especially for the front-line health workers), insufficient diagnostic testing centers, and most importantly suspension of the diagnostic testing centers due to the contamination (infection) by the health workers during the early phase of the outbreaks [17]. For this reason, samples have to be taken to the Federal Capital Territory (Abuja, Nigeria) for testing and that causes delay and uncertainty. Kano state is the center of commerce and the most populous in Nigeria, making it one of the most vulnerable for COVID-19 infection [18]. By 20th February 2021, there were 3636 COVID-19 cases reported in Kano including 99 related deaths [12].

Although Nigeria has done remarkably well in terms of diagnostic testing for COVID-19 cases, there is a still need for improvement to test more people that are COVID-19 suspected or exposed, especially with the asymptomatic behavior of the virus which makes the infected person to spreads it even without symptoms. By 20th February 2021, the number of people tested for COVID-19 in Nigeria was 1,441,013, which accounts for 0.075% of the total population [12]. For this reason, and the fact that testing/autopsies of dead people that are rarely or not be taking place, in some situations, for religious or traditional reasons. This highlights the possible existence of under-ascertainment or under-reporting of the COVID-19 pandemic situation in Nigeria. Also, during April 2020, there were claims about the rises in the number of deaths (suspected to be COVID-19 related [19]) in Kano state of which the state officials at its preliminary investigation urged that the situation about the mysterious deaths during the early outbreaks was not related to COVID-19 pandemic, but rather were likely due to complications arising from other diseases, such as diabetes, hypertension, malaria, meningitis, and so on [20,21]. This obliged the federal government to assigned some delegates to investigate the true scenario in the state. Some reports later show that the rise in the number of COVID-19 deaths during the early outbreaks in Kano was due to shortages of medical services for other diseases as a result of the pandemic crisis [22].

Currently, numbers of epidemiological modelling studies have been done and focus on the use of mathematical/statistical model to investigate the COVID-19 transmission dynamics since its appearance in Wuhan, China [18,23–30]. Several works focused on the estimation of R_0 by adopting the SI and intrinsic growth rate estimates [30,31]; or employing dynamic models and the Markov chain Monte Carlo technique [23,25,27,28]. However, several studies have currently been reported to get insights/understand the dynamics behavior of COVID-19 transmission in Africa [18,26]. In the current work, we aim to study and analyse some vital biological or epidemiological parameters and trends of COVID-19 epidemics during the early phase. We will also estimate the COVID-19 under-ascertainment scenarios in Kano as well as estimate the R_0. We suggest that our results should be vital to inform the world community (especially the public health sector) about the under-ascertainment situation, and also to reveal insights on the spread and control of COVID-19 transmission. Further, we aim to make a short-term prediction for COVID-19 cases to forecast likely future situations to inform public health and decision-makers of the significance of the impacts of NPIs measures compliance. Since the COVID-19 vaccine (and treatment) is still under development, most of the control measures are directed primarily on the use of NPIs, such as social (physical) distancing, community lockdown, quarantine, contact tracing, isolation, and facemasks.

2. Material and methods

Based on the report by the NCDC, the total COVID-19 cases in Kano remains at 73 from 22 to 24 April 2020 (i.e., during the early outbreak), and remains at 77 from 25 to 27 April 2020 (after an increase of four cases). We observed that zero cases were reported from 22 to 24 April 2020, as well as from 25 to 27 April 2020, which seems uncanny looking at the fast-spreading nature of COVID-19 during the early phase, which has been increasing since the confirmation of the index case in Kano which occurred on 11 April 2020 [12]. Thus, we hypothesized that the cases of COVID-19 in the state were under-ascertained (or under-estimated) likely from 22 to 27 April 2020. Therefore, we estimate η and R_0 of COVID-19 in Kano based on the reported data situation from 11 to 30 April 2020.

The time series case data for the COVID-19 confirmed cases obtained via the NCDC situation report was used (from 11 April 2020 until 9 January 2021). The cases data were based on the laboratory confirmation based on the COVID-19 case definition by the NCDC [12] (11). The data selected for the estimation of under-ascertainment, in this study, was from 11 to 30 April 2020, and that for the prediction was from 25 December to 9 June 2020. Note that we did not include the data up to the present date for the estimation of η, because the testing centers were ameliorated remarkably towards May 2020. Also, the provision of personal protective equipment (PPEs) largely improved in the health sector.

Here, we presumed that the COVID-19 cases in Kano were under-ascertained (η) and most likely exist from 22 to 27 April 2020. COVID-19 cumulative total reported cases (C_i) with i representing the day (since 11 April 2020) consist of the summation of cumulative cases ascertained/reported (α_i) and cumulative cases under-ascertained/under-reported (τ_i). Hence, following previous works [31], we used the following relation for estimating the total cases under-ascertained, and is given by $C_i = \alpha_i + \tau_i$, where α_i is observed from the cases data, τ_i is 0 for i before 22 April and η for summation from 22 to 27 April 2020. By adopting previous methods, see for instance [5,30–32], we model the epidemic curve of the disease's pattern in Kano state, Nigeria. The C_i series was used as an exponential growing Poisson process. We observed that the cases from 22 to 27 April 2020 happened to be constant likely due to the aforesaid reasons. Therefore, these data were not included in the exponential growth fitting scheme. The η and the intrinsic growth rate (denoted by γ) of the exponential growth were to be estimated by using the log-likelihood estimation (ℓ) from the Poisson distributed likelihood technique on the daily number of cases. The 95% confidence interval (CI) for η was estimated according to the technique of the profile likelihood estimation with the cutoff threshold calculated using a Chi-square quantile [33,34]. The R_0 was estimated according to the estimation of γ using a similar technique as in [26,30,31]. And, thus, is given by $R_0 = \frac{1}{\int_0^\infty e^{-\gamma\phi} h(\phi) \mathrm{d}\phi}$, which assumed 100% susceptibility during the early outbreak [31], where ϕ represents the COVID-19 SI and follows a probability density function, given by $h(\phi)$. We observed that the above relation was obtained theoretically using previous works [30,34–36].

Subsequently, since the COVID-19 transmission trends in Africa is yet to be fully uncovered (for instance low cases and deaths despite low compliance for NPIs measure), we used the COVID-19SI information from previous estimates, see for example [37,38]. It is worth noting that $h(\phi)$ were modelled as a lognormal distribution which has a mean of 5.0 days and SD of 1.9 days [37,38]. It is also imperative to know that varying the SI information to a small degree would not impact our main results. Further, we observed that the daily cases, denoted by β_i for the i-th day, so that the cumulative confirmations is given by $C_i = C_{i-1} + \beta_i$. A simulation algorithm was formulated for the iterative Poisson distribution given by $E[\beta_i] = C_i \cdot (e^{\gamma} - 1)$ [31]. Where the function $E[\cdot]$ denotes the expectation. For more details on the simulation schemes, see [31].

Moreover, we used the ARIMA model (autoregressive integrated moving average model) to make a short-term forecast based on the observed data [39–41]. The ARIMA model is a well-known time series model framework that was introduced in the 1970s by Box and Jenkins [42], and has been adopted in many studies [40,41]. It is used to investigate the non-stationary part of the data that has one or two differencing possibilities. The model comprises three important parameters (quantities). They are autoregressive order (denoted by p), degree of difference (denoted by d), and the moving average order (denoted by q) [39,41]. Firstly, we applied $log(x+1)$ transformation to the data. Then we differenced the data and set d as the orders of difference when the data reaches stationary. We employed the Augmented Dickey-Fuller test to examine the stationary status of the processed series (i.e., first-order differential of the daily cases of COVID-19 in Kano).

We adopted the autocorrelation function (ACF) and the partial autocorrelation function (PACF) to determine the possible value for parameters p and q. These parameters were assigned with different values to obtain the candidate models. The conditional-sum-of-squares approach was applied to fit these models with given values of p, d, and q, and also to determine the initial values of other key parameters. Thus, the maximum likelihood framework was employed by passing the values of the initial parameters to determine the final estimates [43] (36). Then we calculated the performance indicators (ME, RMSE, MAE, MASE) for candidate models. Akaike's information criterion (AIC) (see [44]) was used in previous studies to choose the best model among alternatives [45–47], and to determine the AIC value for the models. Finally, the best model's combination and the order of the models were obtained based on all indicators (see **Appendix** Fig. 1). The 95% CI were obtained according to the assumption that the model's residuals are normally distributed. To obtain the predicted values, the ARIMA model presumed that the residuals are normally distributed, and could forecast some values that are less than zero. Thus, based on the actual scenarios, we only kept the positive values forecasted and ignore the negative values.

The **R** statistical software (version 3.5.1 with packages '*deSolve*', '*sfsmisc*', '*tseries*', '*forecast*', '*astsa*', '*ggplot2*', and '*DMwR*') was employed to obtain and present the simulation results in this work.

3. Results

Using Fig. 1**a**, we obtained the estimated value of η of 213 (95% CI: 106–346) from 22 to 27 April 2020. In Fig. 1**b**, we estimated the R_0 as 2.74 (95% CI: 2.53–2.96). This result is hugely in line with previous estimates [26,30,31]. In Fig. 2**a**, using the estimated values of R_0 as 2.74 and η as 213, the exponential growing schemes fitted the cumulative number of cases C_i significantly well, and with McFadden's pseudo-R-squared value of 0.99 which indicates the quintessential of the fitting result. Fig. 2**b**, shows the fitting results, using the exponential growth, of the daily confirmation in Kano, Nigeria.

Considering the fact that the estimates of R_0 depend largely on the estimates of the SI of COVID-19. In the current study, we employed the COVID-19 SI information from previous works [37,38] as approximations estimates of the SI to that of Kano. Since the SI estimates need ample time, as well as the information on infection trends, which needs sufficient patient samples and time for follow-up [48], and thus a bit strenuous to be done within a short interval of time. Notwithstanding, using the estimates of the SI for Hong Kong as an approximation to that of Kano could give sensible insight into the dynamics of COVID-in the state. We observed that slightly varying the mean and SD of SI would not vary our main results significantly. Thus, COVID-19 estimates of R_0 in Kano, Nigeria was likely 2.74 (95% CI: 2.53–2.96), which seems in line with previous estimates of R_0, see, for instance [5,30,31].

As for the ARIMA model, we performed the Augmented Dickey-Fuller Test on the processed data and found that the time series data used was stationary with the first-order difference (and the p value of 0.01). The ACF result of the processed data (Appendix Fig. 1a) shows a single spike at the first lag and the PACF (**Appendix** Fig. 1**b**) shows a tapering pattern. An ARIMA (0,1,1) model was initially determined. Then we also included ARIMA (1,1,1) and ARIMA (0,1,2) as candidate models. According to the performance indicators (Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Scaled Error (MASE)), **Appendix Table 1** shows that the ARIMA (0,1,1) was the best model among all models (the highlighted model is selected according to the lowest AIC), and this result was in line with the results of ACF and PACF in **Appendix** Fig. 1.

Fig. 3 presented a 15-day prediction for the cumulative confirmations of COVID-19 cases (red line) based on the ARIMA model (0,1,1), i.e., from 26 December 2020 to 9 January 2021, as well as the observed data (in blue dots). With 95% CI shown in red. The Q-Q plot of residual was plotted in **Appendix** Fig. 2 which shows that the residuals were normally distributed and the Ljung-Box test provided the p-value as 0.39.

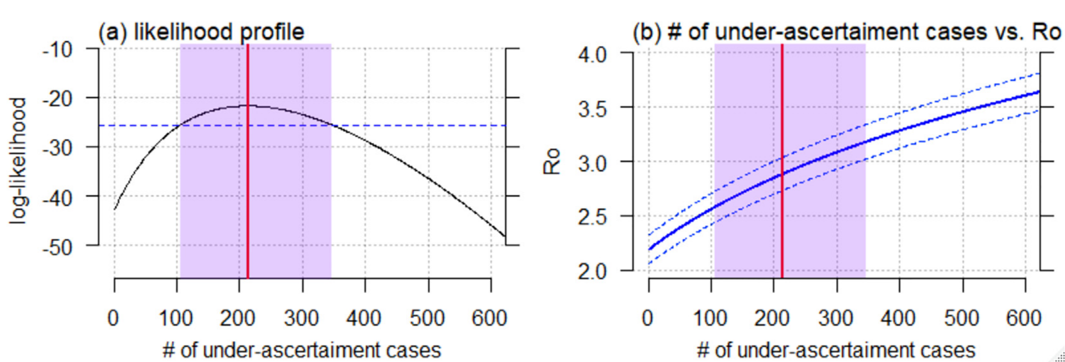

Fig. 1 Estimation of under-ascertainment of COVID-19 cases in Kano, Nigeria from 22 to 27 April 2020, and the basic reproduction number (R_0). Panel (**a**) presents the likelihood profile (ℓ, dark black curve) of the estimated number of unreported cases (η), and the cutoff threshold (horizontal blue dashed line) for the 95% CI. The relationship between η andR_0, where the bold blue curve is the mean estimation, and the dashed blue curves are the 95% CI of estimated R_0. In panels (**a**) and (**b**), the purple shading area on the horizontal axis represents the 95% CI, and the vertical red line represents the maximum likelihood estimate (MLE) of the number of under-ascertainment cases. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

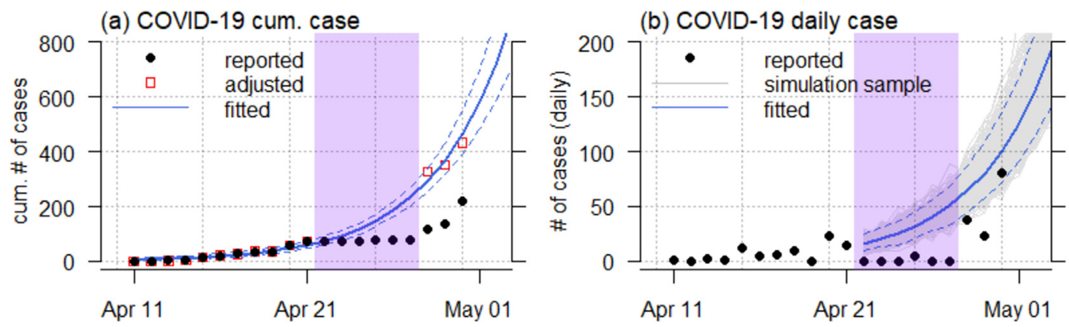

Fig. 2 Time series fitting results of the number of COVID-19 cases in Kano, Nigeria. With the MLE of R_0 at 2.74, panels (a) and (b) presents the exponential growth fitting results of the cumulative number of cases (Ci) and the daily number of cases (β_i), respectively. In panels (**a**) and (**b**), the black bold dots are the reported cases, the blue bold curve represents the median of the fitting results, the dashed blue curves are the 95% CI of the fitting results, and the purple shading area represents the time window from 22 to 27 April 2020. In panel (**a**), the red squares are the cumulative total number of cases estimated. In panel (**b**), the grey curves represent the 1000 simulation samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

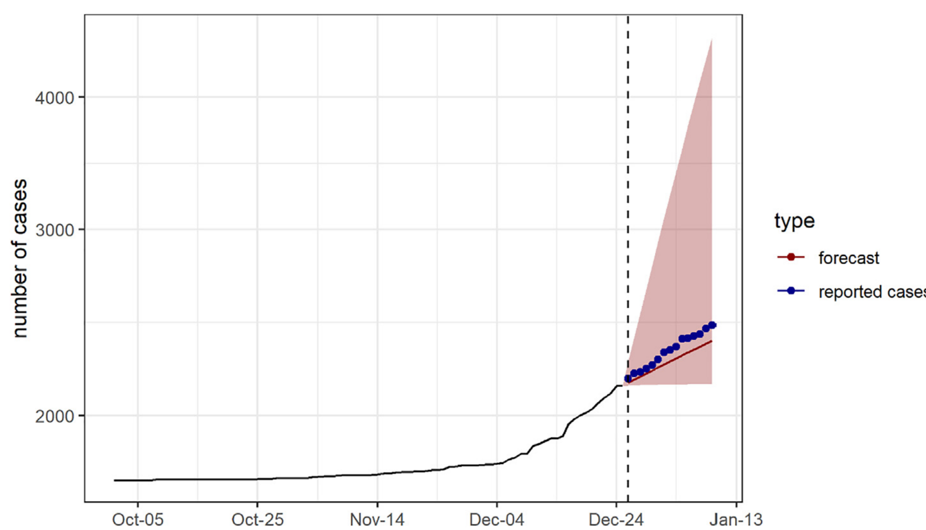

Fig. 3 15-days advance forecast of cumulative number of COVID-19 cases (red curve and pink area) and observed data (blue dots and black curve). The 95% prediction interval is marked in pink. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Furthermore, we found the values of indicators which was used to evaluate the performance of the model and are given by ME as 5.37; RMSE as 12.25; MAE as 9.70; and MASE as 10.67.

4. Discussion

Fig. 1**a** shows the estimated total number of under-ascertainment of COVID-19 cases in Kano and was found to be 213 (95% CI: 106–346). This result implied that the η likely exist and found to be from 22 to 27 April 2020. Fig. 1**b** provides the estimate of the R_0 and found to be 2.74 (95% CI: 2.53–2.96), this result is in line with previous estimates of the basic reproduction number [24–26,30,31,49,50]. Our estimates for the under-ascertainment and R_0 were notably greater than zero, indicating the potential of the virus to spreads across the city.

Fig. 2**b** shows the simulated daily number of cases (β_i) and found that the parameter β_i approximately equaled the observed daily cases after 27 April 2020, and larger than the observations from 22 to 27 April 2020. Our result indicates that under-ascertainment likely exists during the fourth week of April 2020. Hence, the reporting rate was estimated after 27 April 2020 and showed an increase by up to about 10-fold (95% CI: 5–16) compared to the situation before 27 April 2020 on average. This is likely due to inadequate testing centers and insufficient provision of PPEs during the initial phase of the outbreaks. The newly confirmed daily cases begin to increase more rapidly after 27 April 2020, see Fig. 2**b**.

The estimation of under-ascertainment for disease dynamics may be simple for some diseases, like infectious intestinal disease (IID), and seems to be strenuous for some other diseases, such as COVID-19 [51]. Specifically, estimating the under-ascertainment in disease dynamics, cases and controls should be properly/relevantly considered, for instance, in case-control studies. Furthermore, finding the accurate or estimated values of cases ascertained or at least or sustaining a low-risk under-ascertainment is vital for prevention and control of disease's spreads [52]. Looking at the asymptomatic feature of COVID-19 infection, it is plausible that a different reporting controls strategy have different technique for determining the type of control strategies to be employed to prevent/reduce the problems of under-ascertainment in disease's spreads. Thus, we suggest that our results should be regarded in future studies. Besides, estimating key epidemiological quantities or parameters during disease outbreaks using conventional (random) cases data, requires sufficient information on when, where, and to what level the data can represent the actual situation, and in some cases, it is mandatory to consider some adjusted (assumed) values prevent/minimize underestimation or under-reporting scenario. Multiplication factors can be employed to make such adjustments in surveillance (and notification) of case data to impart estimates that are closer to the real situation [16].

The results in **Appendix** Fig. 2 indicate that the residuals probably seem to be white noises, which further highlights that the ARIMA models provide a good predictive model for the COVID-19 prevalence in Kano state. In addition, we predicted that the total cumulative cases in Kano will be around 2353 (95% CI: 2141, 4564) by 9 January 2021. Fig. 3 reveals that the fitted models can provide some epidemiological insights on disease spreads in the state, which will help in the design of effective preventive and control measures. Furthermore, the indicators of the model obtained show the accuracy of the predicted values, which further show the good performance of the estimated models for the case scenarios. The current study has some limitations, which are: the estimates did not provide the daily under-ascertained situation (the exact number of daily cases ascertained), but rather gives the cumulative total number of cases under-ascertained likely from 22 to 27 April 2020. Further, our predictions were done for a short period (i.e., from 26 December 2020 to 9 January 2021) instead of long-term prediction, this is because we are interested more in inferring the unreported situation during the early phase of the outbreaks.

Based on the lessons learned, African countries (in general) have slightly strengthened their preparedness plans against unexpected outbreaks such as the COVID-19 pandemic, which include a timely response for disease spread to prevent community transmission; airport surveillance and temperature screening at ports of entry; improvement of diagnostic centers for COVID-19 testing/mass testing; the strengthened collaboration between African countries and other countries like China, United State of America and UK (for instance deployment of medical personnel from the United Kingdom, donations of testing kits and PPEs by Jack Ma foundation, etc.) [18,53].

Some of the main challenges regarding the fight against the COVID-19 pandemic in Africa include the imposition of the hand-hygiene policy when a majority of Africa's population lacks portable and sufficient water supply, temperature screening at the airports for travelers without considering asymptomatic carriers of the virus, the impossibility of work from home policy and insufficient quantity and quality of PPEs especially for frontline health workers, insufficient data sharing and Illegal crossing of borders at non-official points of entry [53,54].

5. Conclusions

COVID-19 pandemic has caused serious destruction to public health and socio-economic situations worldwide. With vaccines currently being developed for use against SARS-CoV-2, it is paramount important to investigate the effectiveness of NPIs measures to mitigate the spread of the pandemic. This paper investigated the COVID-19 epidemic in Kano, Nigeria during the initial phase of the outbreaks. We estimated the number of under-ascertainment of COVID-19 cases in Kano as 213 (95% CI: 106–346) from 22 to 27 April 2020. Our estimates highlight that the reporting rate after 27 April 2020 likely increased up to about 10-fold (95% CI: 5–16) compared to the situation from 22 to 27 April 2020, and thus should be considered for future epidemiological investigation. The basic reproduction number, R_0, was estimated at 2.74 (95% CI: 2.53–2.96), which indicates the potential of the virus to cause large outbreaks. Our estimates are essential for preparedness planning against future outbreaks.

Declarations

Ethics approval and consent to participate

No ethics consent or approval for participation was needed.

Availability of data and materials

All data used are obtained from public domain available from https://covid19.ncdc.gov.ng/report/.

Funding

DH was supported by an Alibaba (China) Co. Ltd Collaborative Research grant (ZG9Z).

CRediT authorship contribution statement

Salihu S. Musa: Conceptualization, Formal analysis, Writing - original draft. **Shi Zhao:** Conceptualization, Formal analysis, Writing - original draft. **Nafiu Hussaini:** Formal analysis, Writing - review & editing. **Zian Zhuang:** Formal analysis, Writing - original draft. **Yushan Wu:** Formal analysis. **Abdurrahman Abdulhamid:** Formal analysis, revision. **Maggie H. Wang:** Formal analysis, Writing - review & editing. **Daihai He:** Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

MHW is a shareholder of Beth Bioinformatics Co., Ltd. DH received support from an Alibaba (China) Co. Ltd. Collaborative Research grant. Other authors declared that they have no competing interests.

Acknowledgements

None.

Appendix A. .

See Figs. A1 and A2.

See Table A1.

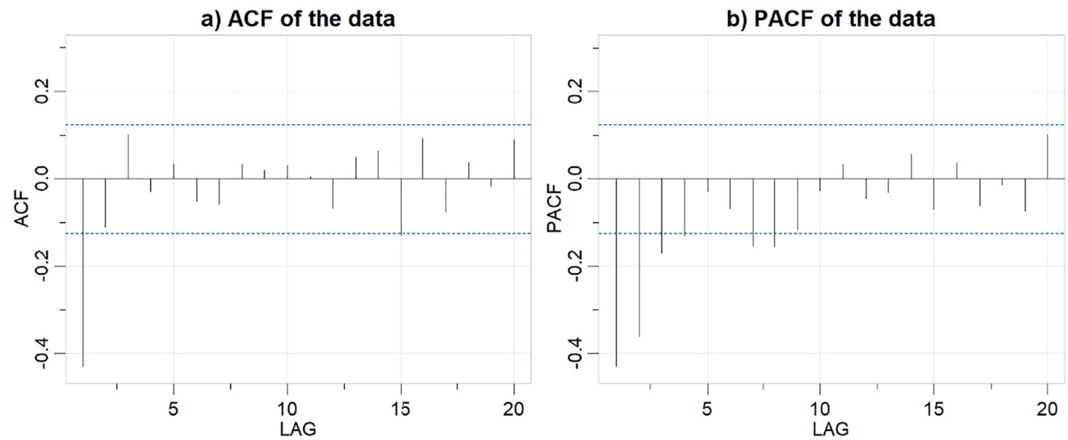

Fig. A1 The ACF and PACF of the COVID-19 situation in Kano, Nigeria.

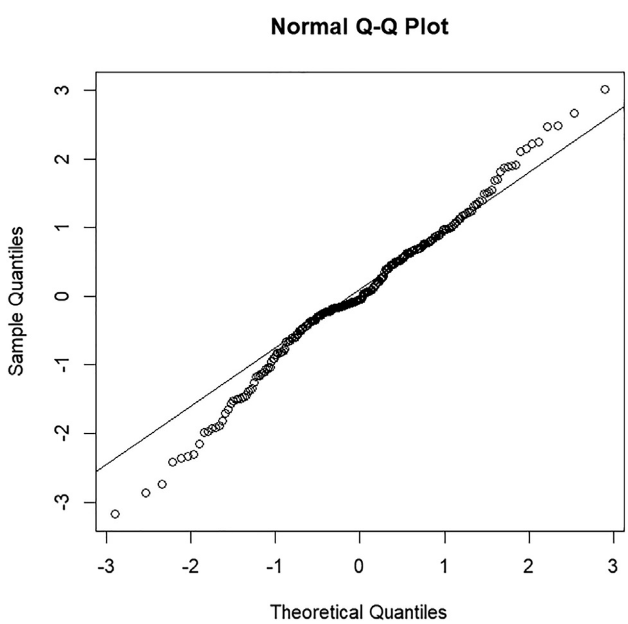

Fig. A2 The Q-Q plot for the residuals of the ARIMA model.

Table A1 Calculated AIC, ME, RMSE, MAE and MASE value of candidate ARIMA models. The highlighted model is selected according to the lowest AIC.

Candidate models	AIC	ME	RMSE	MAE	MASE
ARIMA(0,1,1)	752.57	0.035	1.02	0.78	0.85
ARIMA(1,1,1)	753.13	0.038	1.02	0.78	0.86
ARIMA(0,1,2)	753.04	0.038	1.02	0.78	0.86

References

[1] V.G. da Costa, M.L. Moreli, M.V. Saivish, The emergence of SARS, MERS and novel SARS-2 coronaviruses in the 21st century, Arch. Virol. 165 (7) (2020) 1517–1526.

[2] C. Huang et al, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, The Lancet 395 (10223) (2020) 497–506.

[3] World Health Organization (who). Coronavirus disease (COVID-19) Dashboard. 2021; Available from: https://covid19.who.int/.

[4] H. Harapan et al, Coronavirus disease 2019 (COVID-19): A literature review, J. Infect. Public Health 13 (2020) 667–673.

[5] World Health Organization (WHO). WHO director-general's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV) – 30 January 2020, 2020; Available from: https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-(2019-ncov).

[6] World Health Organization. WHO director-general's opening remarks at the media briefing on COVID-19 – 11 March 2020. 2020; Available from: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.

[7] N. Zhu et al, A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. 382 (2020) 727–733.

[8] N. Chen et al, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, The Lancet 395 (10223) (2020) 507–513.

[9] A.J. Rodriguez-Morales et al, Clinical, laboratory and imaging features of COVID-19: A systematic review and meta-analysis, Travel Med. Infect. Dis. 34 (2020) 101623.

[10] T. Kuiken et al, Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome, The Lancet 362 (9380) (2003) 263–270.

[11] N. Van Doremalen et al, Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1, N. Engl. J. Med. 382 (16) (2020) 1564–1567.

[12] Nigeria Center for Disease Control (NCDC). Coronavirus disease (COVID-19) pandemic, 2021; Available from: https://ncdc.gov.ng/.

[13] S.S.Q.S. Musa, S. Zhao, A. Yusuf, U.T. Mustapha, D. He, Mathematical modeling of COVID-19 epidemic with effect of awareness programs, Infect. Disease Model. 6 (2021) 448–460.

[14] Center for Disease Control and Prevention (CDC). Coronavirus disease (COVID-19) 2020 interim case definition, approved April 5, 2020, 2020; Available from: https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/.

[15] C. Ohia, A.S. Bakarey, T. Ahmad, COVID-19 and Nigeria: putting the realities in context, Int. J. Infect. Diseases 95 (2020) 279–281.

[16] C.L. Gibbons et al, Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods, BMC Public Health 14 (1) (2014) 1–17.
[17] BBC News, What is behind Nigeria's unexplained deaths in Kano. Assessed 10 May 2020, 2020; Available from: https://www.bbc.com/news/world-africa-52454259.
[18] M. Gilbert et al, Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study, The Lancet 395 (10227) (2020) 871–877.
[19] Reuters, Nigeria investigating jump in deaths in Kano city, coronavirus suspected. Assessed 6 June 2020, 2020; Available from: https://www.reuters.com/article/health-coronavirus-nigeria-kano-idUSL5N2CB6SW.
[20] Reuters, Nigeria Kano state's rise in deaths not due to coronavirus -state government, 2020; Available from: https://www.reuters.com/article/health-coronavirus-nigeria-kano/nigeria-kano-states-rise-in-deaths-not-due-to-coronavirus-state-government-idUSL5N2CE0LO?edition-redirect = uk.
[21] O.O. Odukoya, U.A. Omeje, COVID-19 pandemic in Nigeria: A case study of Kano State-Challenges and lessons learned, J. Clin. Sci. 17 (4) (2020) 91.
[22] Yahoo News, Virus cases leap in Nigeria's Kano. 2020; Available from: https://news.yahoo.com/virus-cases-leap-nigerias-kano-140755914.html? guccounter = 1&guce_referrer = aHR0cHM6Ly9lbi53a WtpcGVkaWEub3Jn Lw&guce_referrer_sig = AQAAAJ3H ZANpnoviX4u FMPtsMA bm4gUuP1N– 7yRaL69vTejCGyUtG5Hq JxY7oJ6vwWqnAbRJ7E2tVCGgTgB1e QbblWLk8Aiw1v9tv0 EXVUJR2tdjPG1ZjgCgi0PtnMSCkJkOT Fr5V3uL4DdKlkd8fcb2c 14SNAUA3s45CWRx4k0g4Xn.
[23] S.E. Eikenberry et al, To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic, Infect. Disease Model. 5 (2020) 293–308.
[24] Q. Li et al, Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia, N. Engl. J. Med. 382 (1199–1207) (2020).
[25] Q. Lin et al, A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action, Int. J. Infect. Diseases 93 (2020) 211–216.
[26] S.S. Musa et al, Estimation of exponential growth rate and basic reproduction number of the coronavirus disease 2019 (COVID-19) in Africa, Infect. Diseases Poverty 9 (1) (2020) 1–6.
[27] C.N. Ngonghala et al, Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus, Math. Biosci. 325 (2020) 108364.
[28] B. Tang et al, Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, J. Clin. Med. 9 (2) (2020) 462.
[29] J.T. Wu, K. Leung, G.M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, The Lancet 395 (10225) (2020) 689–697.
[30] S. Zhao et al, Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak, Int. J. Infect. Diseases 92 (2020) 214–217.
[31] S. Zhao et al, Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak, J. Clin. Med. 9 (2) (2020) 388.
[32] U. De Silva et al, A preliminary analysis of the epidemiology of influenzaA(H1N1)vvirusinfectioninThailandfromearlyoutbreak data,June-July2009,Eurosurveillance14(31)(2009)19292.
[33] J. Fan, T. Huang, Profile likelihood inferences on semiparametric varying-coefficient partially linear models, Bernoulli 11 (6) (2005) 1031–1057.
[34] S. Zhao, et al., Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020, Front. Phys., 2020. 8(207).
[35] J. Wallinga, M. Lipsitch, How generation intervals shape the relationship between growth rates and reproductive numbers, Proc. Roy. Soc. B: Biol. Sci. 274 (1609) (2007) 599–604.
[36] S. Zhao et al, Simple framework for real-time forecast in a data-limited situation: the Zika virus (ZIKV) outbreaks in Brazil from 2015 to 2016 as an example, Parasites Vectors 12 (1) (2019) 1–13.
[37] Z. Du et al, Serial interval of COVID-19 among publicly reported confirmed cases, Emerg. Infect. Dis. 26 (6) (2020) 1341.
[38] H. Nishiura, N.M. Linton, A.R. Akhmetzhanov, Serial interval of novel coronavirus (COVID-19) infections, Int. J. Infect. Diseases 93 (2020) 284–286.
[39] X. Li et al, A comparative time series analysis and modeling of aerosols in the contiguous United States and China, Sci. Total Environ. 690 (2019) 799–811.
[40] A.F. Lukman et al, COVID-19 prevalence estimation: Four most affected African countries, Infect. Disease Model. 5 (2020) 827–838.
[41] M. Maleki et al, Time series modelling to forecast the confirmed and recovered cases of COVID-19, Travel Med. Infect. Dis. 37 (2020) 101742.
[42] G.E. Box, et al., Time series analysis: forecasting and control, John Wiley & Sons, Hoboken, NJ, 2008.
[43] N. Shephard, The relationship between the conditional sum of squares and the exact likelihood for autoregressive moving average model, 1997.
[44] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: Selected papers of hirotugu akaike, Springer, 1998, pp. 199–213.
[45] D.Ö. Faruk, A hybrid neural network and ARIMA model for water quality time series prediction, Eng. Appl. Artif. Intell. 23 (4) (2010) 586–594.
[46] P. Mondal, L. Shit, S. Goswami, Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices, Int. J. Comput. Sci., Eng. Appl. 4 (2) (2014) 13.
[47] Y.-W. Wang, Z.-Z. Shen, Y. Jiang, Comparison of ARIMA and GM (1, 1) models for prediction of hepatitis B in China, PLoS ONE 13 (9) (2018) e0201987.
[48] B.J. Cowling et al, Estimation of the serial interval of influenza, Epidemiology (Cambridge, Mass.) 20 (3) (2009) 344.
[49] J. Ma, Estimating epidemic exponential growth rate and basic reproduction number, Infect. Disease Model. 5 (2020) 129–141.
[50] C. You et al, Estimation of the time-varying reproduction number of COVID-19 outbreak in China, Int. J. Hyg. Environ. Health 228 (2020) 113555.
[51] D. Sethi et al, Investigation of under-ascertainment in epidemiological studies based in general practice, Int. J. Epidemiol. 28 (1) (1999) 106–112.
[52] Catalogue of bias, Ascertainment bias, Systematic differences in the identification of individuals included in a study or distortion in the collection of data in a study. Assessed 16 May 2020. 2020; Available from: https://catalogofbias.org/biases/ascertainment-bias/.
[53] World Health Organization (WHO) Regional Office for Africa. COVID-19 Strategic Response Plan in the WHO African Region. Assessed on 13 December 2020. 2020; Available from: https://www.afro.who.int/health-topics/coronavirus-covid-19#sitrep.
[54] R. Nyarko et al, The Impact on Public Health and Economy Using Lockdown as a Tool against COVID-19 Pandemic in Africa: A Perspective, J. Epidemiol. Public Health Rev. 5 (3) (2020).