Estimating the under-ascertainment number of COVID-19 cases in Kano, Nigeria in the fourth week of April 2020: a modelling analysis of the early outbreak

Salihu S Musa

Hong Kong Polytechnic University

Shi Zhao

Chinese University of Hong Kong

Nafiu Hussaini

Bayero University

Zian Zuang

Hong Kong Polytechnic University

Yushan Wu

Chinese University of Hong Kong

Maggie H Wang

Chinese University of Hong Kong

Daihai He (✉ daihai.he@polyu.edu.hk)

Hong Kong Polytechnic University <https://orcid.org/0000-0003-3253-654X>

Research article

Keywords: COVID-19, under-ascertainment, reproduction number, Nigeria, modelling.

Posted Date: October 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-34391/v2>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The coronavirus disease 2019 (known as COVID-19) pandemic caused by Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2) appeared in Wuhan, China has rapidly spread to over 200 countries and territories. In Nigeria, the Kano State Ministry of Health has confirmed its first case of COVID-19 on April 11, 2020, and since then there might have been issues of under-ascertainment that occurred roughly from 22 to 27 April 2020. As of 4 October 2020, there were 1738 reported COVID-19 cases in Kano with 54 associated deaths. In this work, we estimate the number of under-ascertainment cases and the basic reproduction number, B , of COVID-19 in Kano, Nigeria. We also predict the number of COVID-19 cases in the short term.

Methods: We employ the exponential growth and modelled the outbreak curve of COVID-19 cases, in Kano, Nigeria from 11 to 30 April 2020. We estimated the number of under-ascertainment cases using the maximum likelihood estimation. We adopted the SI estimated for Hong Kong as approximations of the unknown SI for COVID-19 in Kano to estimate the a . We use ARIMA model to provide a short term (15 days) prediction of the COVID-19 cases in Kano, Nigeria.

Results: We revealed that the initial growth phase mimic an exponential growth pattern. We found that the under-ascertainment was likely to have resulted in 213 (95% CI: 106–346) unreported cases from 22 to 27 April 2020. The reporting rate after 27 April 2020 increase up to 10-fold compared to the scenario from 22 to 27 April 2020 on average. We estimated the c of COVID-19 in Kano as 2.74 (95% CI: 2.53–2.96). We forecasted that the total number of COVID-19 cases in Kano to be 1067 (95% CI: 883, 2137) by June 6, 2020.

Conclusion: The under-ascertainment likely exists during the fourth week of April, 2020 and should be regarded in the future analysis/investigation.

1. Introduction

Coronaviruses are a group of related Ribonucleic acid (RNA) viruses that belongs to the *Coronaviridae* family (and the order *Nidovirales*) and widely disseminated in human beings (Huang et al., 2020). Most of the coronavirus infections in human have mild symptoms. The outbreaks of the two other coronaviruses (known as beta-coronaviruses), which include severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), have caused more than 10,000 cumulative cases in the past two decades, with the death rates of about 10% for SARS-CoV and 37% for MERS-CoV (Huang et al., 2020). The novel coronavirus disease 2019 (named COVID-19) started in Wuhan, Hubei province of China and spread worldwide, have been declared a Public Health Emergency of International Concern by the WHO on January 30, 2020 (WHO, 2020a), and later named the 1st pandemic caused by coronaviruses on March 11, 2020 (WHO, 2020b). As of June 3, 2020, COVID-19 affected over 6.5 million people with more than 386000 fatalities worldwide (WHO, 2020c).

Nigeria has reported its first (imported) case of COVID-19 on February 27, 2020 following the case definition by the NCDC (NCDC, 2020; Ohia, 2020). After which community transmission takes place due to inadequate contacts tracing from index case and lack of early closure of borders to prevent further spread. The fact that the control of COVID-19 pandemic rely heavily on a country's health care system. Nigeria is currently witnessing a rapid increase of COVID-19 cases probably due to the poor health care system, making it more vulnerable to the virus especially with population of over 200 million people (the highest in Africa). In northern region, Kano state has confirmed its first case of COVID-19 on April 11, 2020 (NCDC, 2020), and since then there might have been issues of under-ascertainment that exists roughly from 22 to 27 April 2020. This is likely due to the lack of sufficient health care facilities (such as test kits, gowns, and facemasks), limited diagnostic testing of suspected patients, and some other unknown reasons. Being the commercial center and the most populace state in Nigeria (with more than 10 million people), Kano is likely the most vulnerable for COVID-19 in northern Nigeria (Gilbert, 2020). As of June 3, 2020, there have been 970 cases of COVID-19 infections confirmed in Kano, Nigeria including 45 deaths (NCDC, 2020). Further the fact that few COVID-19 diagnostic testing was done in Nigeria. In particular, As of May 26, 2020 only about 69,801 people were tested throughout the country (NCDC, 2020). This together with the fact that autopsies and testing of deceased individuals are not generally carried out in the country (in some cases, for traditional or religious reasons), clearly suggests a gross under-ascertainment of the true scenario of the pandemic in the country.

Recently, there were numerous researches focusing on mathematical and statistical modelling to study the dynamics transmission of COVID-19 pandemic since its emergence in Wuhan, China (Li, 2020; Zhao et al., 2020a; Zhao et al., 2020b; Ngonghala et al., 2020; Eikenberry et al., 2020; Tang et al., 2020; Wu et al., 2020; Musa et al., 2020; Gilbert et al., 2020; Lin et al., 2020). Some of these studies focused on estimation of basic reproduction number by using the serial intervals and intrinsic growth rate (Zhao et al., 2020a; Zhao et al., 2020b) or using ordinary differential equations and Markov Chain Monte Carlo methods (Ngonghala et al., 2020; Eikenberry et al., 2020; Tang et al., 2020; Lin, 2020). However, few studies have been done to understand the transmission of COVID-19 pandemic in Africa (Gilbert et al., 2020; Musa et al., 2020).

In this study, we aim to investigate the epidemiological patterns, and estimate the number of under-ascertainment cases and R_0 of the COVID-19 outbreak in Kano, Nigeria. We hope our results in this study will be useful to inform the world community of the under-ascertainment issues and the value of R_0 in order to help to curtail the spread of the virus. In addition, our study will make a short-term forecast of COVID-19 cases in order to predict possible scenario and informed decision makers in the country about the importance of sustaining stringent measures as recommended by the WHO and other health related organizations as the virus is yet to have effective treatment or vaccination. All measures are currently directed primarily to non-pharmaceutical interventions (NPIs), like social (physical) distancing, community lockdown, quarantine of suspected cases, contact tracing, isolation of confirmed cases and the use of facemasks in general public.

2. Data And Methods

According to the NCDC report, the total cumulative number of COVID-19 cases in Kano stand at 73 between 22 to 24 April 2020, and stand at 77 between 25 to 27 April 2020 (after addition of four new cases), i.e., no new case was reported between 22 to 24 April 2020, as well as between 25 to 27 April 2020, which appears weird considering the rapid increase of the outbreak curve since the index case on 11 April 2020 in Kano (NCDC, 2020). We presume that the COVID-19 cases in Kano were under-ascertained probably from 22 to 27 April 2020. In this paper, we estimate the number of under-ascertainment cases and of COVID-19 in Kano, Nigeria from 22 to 27 April 2020 based on the available data during the early phase of the epidemic.

We used the time series data for cumulative confirmation compiled by the NCDC (2020) from 11 April to 30 April 2020. All cases data were confirmed from the laboratory according to the definition of COVID-19 cases by the NCDC which is available at <https://covid19.ncdc.gov.ng/report/>. The data chosen for this study was from 11 to 30 April 2020 instead of including up to the present date, this is due to the fact that the diagnostic testing have improved significantly since towards the end of April 2020, and also sufficient personal protective equipment's (PPEs) were provided to the frontline health workers.

We suspected that there was a number of under-ascertainment of COVID-19 cases, denoted by, likely from 22 to 27 April 2020. The cumulative confirmation of the total number of cases, represented by C_i of the i -th day since 11 April 2020 is the summation of the cumulative cases reported/ascertained, represented by C_i^r , and cumulative number of under-ascertainment cases, represented by C_i^u . Thus, following the previous studies (Trotter et al., 2005; Zhao et al., 2020b), the relation/formula for computing the expected number of under-ascertainment of cases is given as $C_i^u = C_i - C_i^r$, where C_i^r is observed from the data, and C_i^u is 0 for i before 22 April and for i after 27 April 2020. We employed the approach used in previous works (Zhao et al., 2020a; Zhao et al., 2020b; De Silva, 2009) and we modelled the outbreak curve. The C_i^r series is used as an exponential growing Poisson process. The data from 22 to 27 April 2020 seems constant probably due to the poor testing facilities and some other unknown reasons, thus, these data were ignored in exponential growth fitting. The C_i^r and the intrinsic growth rate (represented by γ) of the exponential growth were to be estimated using the log-likelihood estimation (ℓ), from the Poisson distributed likelihood framework on number of cases. We estimated the 95% confidence interval (95% CI) of γ based on the profile likelihood estimation technique with cutoff threshold computed by a Chi-square quantile, given by $\chi^2_{pr = 0.95, df = 1}$ (Fan & Huang, 2005). We obtained the R_0 based on the estimation of γ following similar approach as in (Zhao et al., 2020a; Zhao et al., 2020b; Musa et al., 2020). Therefore, using similar approach (Zhao et al., 2020b), we have the basic reproduction number, R_0 , given by $R_0 = \frac{1}{1 - \gamma \tau}$ with 100% susceptibility presumed at the early stage for COVID-2019 outbreak, where τ denotes the serial interval (SI) of COVID-19 follows a probability density function $h()$. We note that this formula has been derived theoretically as well as adopted in previous studies (Zhao et al., 2020a; Zhao et al., 2020d; Wallinga & Lipsitch, 2007).

Since the transmission chain of COVID-19 in Africa still remains fully uncovered, we adopted the SI information of COVID-19 from previous works, see for instance (Du et al., 2020; Nishura et al., 2020). We note that were modelled as a lognormal distribution with mean of 5.0 days and standard deviation (SD) of 1.9 days (Du et al., 2020; Nishura et al., 2020). It is important to note that slightly changing the SI information may not affect our main results and conclusion. In this work, we also aimed to evaluate the trends of the daily number of cases, in this case, represented by C_i for the i -th day, and given as $C_i = C_{i-1} + \dots$. A simulations algorithm based on previous study (Zhao et al., 2020b) was formulated for the iterative Poisson distribution given by λ , here the function λ represent the expectation. For details of the simulations framework see (Zhao et al., 2020b).

Furthermore, we employed Autoregressive Integrated Moving Average (ARIMA) model (a time series model), which has been used in previous study to make short-term prediction (Maleki, 2020). The model consists of three key parameters: p (autoregressive order), d (the degree of difference) and q (moving average order). We assigned these parameters different values to form candidate models. To fit models with given values of p , d , and q , conditional-sum-of-squares method is used to find starting values of other parameters, then maximum likelihood method is applied passing the starting parameter values to calculate the final estimates (Shephard, 1997). As Akaike's Information Criterion (AIC) (Akaike, 1998) is widely used in selecting the best model among alternatives (Wang et al., 2018; Ömer, 2010; Mondal et al., 2014), we calculated AIC value for models and chose the model with the lowest AIC. The 95% confidence intervals are obtained based on the assumption that residuals of the model are normally distributed. When making predictions, ARIMA model assumes the residuals normally distributed, which could make some predictions go below zero. Here, according to the actual situation, we only kept the non-negative predicted value.

The **R** statistical software (version 3.5.1) was used for the simulations in this study.

3. Results

We estimated the total number of COVID-19 under-ascertainment cases as 213 (95% CI: 106–346) likely from 22 to 27 April 2020, see Figure 1a. Clearly, this result was notably greater than zero. We also estimated the R_0 as 2.74 (95% CI: 2.53–2.96), see Figure 1b, which is largely consistent with previous findings (Zhao et al., 2020a; Zhao et al., 2020b; Musa et al., 2020). In Figure 2a, with the value of R_0 as 2.74 and C_0 as 213, the exponential growing framework fitted the cumulative number of COVID-19 cases (C_i) eloquently well, considering the McFadden's pseudo- R -squared value of 0.99. Figure 2b showed the fitting results using the exponential growth of the daily number of COVID-19 cases in Kano, Nigeria.

The estimation of R_0 rely hugely on the SI of COVID-19. In this study, we adopted the SI information of the COVID-19 from previous works (Du et al., 2020; Nishura et al., 2020) as approximations to that of Kano, considering the fact that the estimation of SI requires sufficient time. This is because the computation of SI needs the knowledge of disease transmission chain which requires adequate number of patient (09), and thus this is difficult to be done in a short period.

However, using the SI of Hong Kong as approximation could provide a reasonable insight into the transmission potential and features of COVID-19 in Kano at the early phase of the outbreak. We found that changing the mean and SD of SI very slightly would not change our main results. The R_0 of COVID-19 in Kano was estimated at 2.74 (95% CI: 2.53–2.96), and this is consistent with previously computed R_0 (Zhao et al., 2020a; Zhao et al., 2020b).

In Figure 3, we depicted 15-days forecast of cumulative covid-19 cases (red line) and observed data (blue dots). The 95% confidence interval is marked in red. We found the values of indicators that evaluated the performance of the model used given by: Mean Error (ME) as -4.66; Root Mean Squared Error (RMSE) as 7.18; Mean Absolute Error (MAE) as 6.07; and Mean Absolute Scaled Error (MASE) as 0.33. Furthermore, we presented the auto-correlation function (ACF) and partial autocorrelation function (PACF) plots of the reported COVID-19 cases in Kano, see Appendix Figure 1. We presented the Q-Q plot of residual that show the residuals were normally distributed and Ljung-Box test provided the p value as 0.54, see Appendix Figure 2.

4. Discussion

The total number of COVID-19 under-ascertainment cases estimated as 213 (95% CI: 106–346), see Figure 1a. This result insinuated the existence of under-ascertainment cases likely from 22 to 27 April 2020. After considering the effect of under-ascertainment, our estimation of the R_0 as 2.74 (95% CI: 2.53–2.96), see Figure 1b, is also consistent with previous findings (Zhao et al., 2020a; Zhao et al., 2020b; Li et al., 2020; Ma, 2020; Lin et al., 2020; Musa et al., 2020; You et al., 2020).

We provided the simulated daily number of COVID-19 cases () as shown In Figure 2b. We found that, the parameter is approximately equaled the observed daily number of cases after 27 April 2020, but was larger than the observations of cases from 22 to 27 April 2020. Our finding highlight that the under-ascertainment probably exists in the fourth week of April 2020. Thus, the reporting rate was estimated after 27 April 2020, which was found to have increased by up to 10-fold (95% CI: 5-16) compared to the scenario before 27 April 2020 on average. One possible reasons was due to the poor testing facilities for diagnosing daily new cases, and lack of adequate PPEs for the frontline health workers. The newly reported daily cases started increasing very fast after 27 April 2020, see Figure 2b.

Although under-ascertainment is difficult to be ameliorate for some diseases (Sethi et al., 1999), e.g., infectious intestinal disease (IID), this is unlikely for COVID-19 (Sethi et al., 1999). In general, solving the problem of under-ascertainment, cases and controls need to be appropriately included in cases-control studies. In addition, knowing the exact number of ascertainment cases or maintaining a low risk under-ascertainment cases is very crucial in controlling disease pandemic (<https://catalogofbias.org/biases/ascertainment-bias/>). Considering the general asymptomatic (mild) nature of the COVID-19 infection, it is possible that different reporting controls have different criteria for finding the kind of strategies to be used in order to avoid or reduce under-ascertainment issues. We assert future analysis /investigation. Furthermore, estimation of key

epidemiological parameters in populations during disease epidemics by using routine data requires knowledge of when, where and to what extent these data represent the true scenario of disease, and in some instances it is necessary to make some adjustment in order to avoid underestimation. Multiplication factors can also be used to adjust notification and surveillance data to provide more realistic estimates of incidence (Gibbons et al., 2014).

We conducted the first-order difference and Augmented Dickey-Fuller Test to the reported COVID-19 cases in Kano, Nigeria. Processed data was proved to be the stationary time series (p value 0.017). According to the lowest AIC, the best model among all models was ARIMA (2,1,0) (Table 1), which is consistent with the ACF and PACF result (Appendix Figure 1) of the Processed data. The Q-Q plot of residual (Appendix Figure 2) revealed that the residuals were normally distributed and Ljung-Box test provided the p -value as 0.54, suggesting that the residuals are likely to be white noises. Furthermore, we forecasted that the total number of COVID-19 cases will be 1067 (95% CI: 883, 2137) by June 6, 2020. Figure 3 indicates that the fitted models seems very appropriate. We also found the values of indicators that computed the performance of the model used, which are given by: Mean Error (ME) as -4.66; Root Mean Squared Error (RMSE) as 7.18; Mean Absolute Error (MAE) as 6.07; and Mean Absolute Scaled Error (MASE) as 0.33, showing the accuracy measures for the forecast model, this further highlight the good performance of the estimated models of the total confirmed cases.

Table 1. Calculated AIC value of candidate ARIMA models.

Candidate models	AIC value
ARIMA(0,1,0)	388.86
ARIMA(0,1,1)	376.87
ARIMA(0,1,2)	377.94
ARIMA(0,1,3)	377.72
ARIMA(0,1,4)	378.42
ARIMA(0,1,5)	380.42
ARIMA(1,1,0)	385.70
ARIMA(1,1,1)	378.38
ARIMA(1,1,2)	377.62
ARIMA(1,1,3)	379.20
ARIMA(1,1,4)	380.42
ARIMA(2,1,0)	376.50
ARIMA(2,1,1)	378.43
ARIMA(2,1,2)	378.84
ARIMA(2,1,3)	380.49
ARIMA(3,1,0)	378.46
ARIMA(3,1,1)	378.57
ARIMA(3,1,2)	379.98
ARIMA(4,1,0)	379.80
ARIMA(4,1,1)	379.35
ARIMA(5,1,0)	380.11

5. Conclusions

The pandemic of COVID-19 is still causing serious damage to global public health and socioeconomic

fast-spreading virus, there is need to assess and maintain effective non-pharmaceutical intervention measures to curtail/control further spread of the virus. In this study, we estimated under-ascertainment cases of the COVID-19 in Kano, Nigeria, which probably existed and resulted in 213 (95% CI: 106–346) unreported cases, likely from 22 to 27 April 2020. The reporting rate after 27 April 2020 was likely to have increased up to 10-fold (95% CI: 5–16) in comparison with the scenario from 22 to 27 April 2020 on average, and it should be regarded in future analysis/investigation. We estimated the R_0 of COVID-19 in Kano as 2.74 (95% CI: 2.53–2.96), which shows the potential of COVID-19 to spread and persist in a community. Using the ARIMA model, we forecasted that the total number of COVID-19 cases will be 1067 (95% CI: 883, 2137) by 6 June 2020, which is crucial for preparedness against future epidemic.

Abbreviations

R_0 : Basic reproduction number; SI: Serial interval; CI: Confidence interval; n : number of under-ascertainment of COVID-19 cases; C_t : Cumulative number of cases; Σ : Summation of the cumulative reported cases; U : cumulative unreported cases; γ : Intrinsic growth rate; ℓ : log-likelihood estimation; $(\bullet)$: moment generating function of the probability distribution for the serial interval; $h()$: Probability distribution for the serial interval; $h(k)$: Lognormal distribution; 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.

Declarations

Ethics approval and consent to participate

No ethics approval or consent to participate was required.

Consent to publish

All authors read and gives consent for publication of this manuscript.

Availability of data and materials

All data are publicly available via <https://covid19.ncdc.gov.ng/report/>.

Competing interests

All authors declare that they have no competing interests.

Funding

Not applicable.

Authors' contributions

Loading [MathJax]/jax/output/CommonHTML/jax.js

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

Acknowledgements

None.

Authors' information

^aDepartment of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China. ^bDepartment of Mathematics, Kano University of Science and Technology, Wudil, Nigeria. ^cJC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China. ^dShenzhen Research Institute of Chinese University of Hong Kong, Shenzhen, China. ^eDepartment of Mathematical Sciences, Bayero University, Kano, Nigeria. Corresponding author: daihai.he@polyu.edu.hk (DH).

References

1. Huang C, Wan Y, Li X, Ren L, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet Resp Med*. 2020;8(5):475-481. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
2. World Health Organization. WHO director-general's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV) – 30 January 2020a. [https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-\(2019-ncov\)](https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-(2019-ncov)). Accessed June 2020.
3. World Health Organization. WHO director-general's opening remarks at the media briefing on COVID-19 – 11 March 2020b. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020>. Accessed June 2020.
4. World Health Organization. Coronavirus disease (COVID-19) Situation Reports. 2020c. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>. Assessed October 2020.
5. Nigeria Center for Disease Control. Coronavirus disease (COVID-19) situation reports. 2020. <https://covid19.ncdc.gov.ng/report/>. Assessed October, 2020.
6. Center for Disease Control and Prevention. Coronavirus disease (COVID-19) 2020 interim case definition, approved April 5, 2020. 2020. <https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/>. Assessed October, 2020.
7. Ohia C, Bakarey AS, Ahmad T. COVID-19 and Nigeria: putting the realities in context. *Int J Infect Dis*. 2020;95: 279–281. <https://doi.org/10.1016/j.ijid.2020.04.062>.

8. Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *Lancet*. 2020;395(10227):871-77. [https://doi.org/10.1016/S0140-6736\(20\)30411-6](https://doi.org/10.1016/S0140-6736(20)30411-6).
9. Li Q, Guan X, Wu P, Wang X, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 2020;382:1199-1207. <https://doi.org/10.1056/NEJMoa2001316>.
10. Zhao S, Lin Q, Ran J, Musa SS, 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 Dis*. 2020a;92:214-217. <https://doi.org/10.1016/j.ijid.2020.01.050>.
11. Zhao S, Musa SS, Lin Q, Ran J, 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*. 2020b;9:388. <https://doi.org/10.3390/jcm9020388>.
12. Zhao S, Gao D, Wu Y, Zuang Z, et al. Estimating the serial interval of novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. *Frontiers in Phys*. 2020c;8:207. <https://doi.org/10.3389/fphy.2020.00347>.
13. Zhao S, Musa SS, Fu H, He D, 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 Vect*. 2019d, 12, 344. <https://doi.org/10.1186/s13071-019-3602-9>.
14. Ngonghala CN, Iboi E, Eikenberry S, Scotch M, et al. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *Math Biosci*. 2020;325:108364. <https://doi.org/10.1016/j.mbs.2020.108364>.
15. Eikenberry SE, Mancuso M, Iboi E, Phan T, 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 Dis Model*. 2020;5:293-308. <https://doi.org/10.1016/j.idm.2020.04.001>.
16. Tang B, Wang X, Li Q, Bragazzi NL, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. *J Clin Med*. 2020;9(2):462. <https://doi.org/10.3390/jcm9020462>.
17. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*. 2020;395(10225):689-697. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9).
18. Musa SS, Zhao S, Wang M, Habib AG, et al. Estimation of exponential growth rate and basic reproduction number of the coronavirus disease 2019 (COVID-19) in Africa. *Infect Dis Povert*. 2020;9(96). <https://doi.org/10.1186/s40249-020-00718-y>.
19. Lin Q, Zhao S, Gao D, Lou Y, et al. A conceptual model for the outbreak of Coronavirus disease 2019 (COVID-19) in Wuhan, China with individual reaction and governmental action. *Int J Infect Dis*. 2020;93:211-216. <https://doi.org/10.1016/j.ijid.2020.02.058>.
20. Fan J, Huang T. Profile likelihood inferences on semiparametric varying-coefficient partially linear

21. Ma J. Estimating epidemic exponential growth rate and basic reproduction number. *Infect Dis Model.* 2020;5:129-141. <https://doi.org/10.1016/j.idm.2019.12.009>.
22. Cowling BJ, Fang VJ, Riley S, Peiris JM, et al. Estimation of the serial interval of influenza. 2009;20(3):344-347. <https://dx.doi.org/10.1097%2FEDE.0b013e31819d1092>.
23. Du Z, Xu X, Wu Y, Wang L, et al. Serial interval of COVID-19 among publicly reported confirmed cases. *Emerg Infect Dis.* 2020;26(6):1341. <https://doi.org/10.3201/eid2606.200357>.
24. Sethi D, Wheeler J, Rodrigues LC, Fox S, Roderick P, Investigation of under-ascertainment in epidemiological studies based in general practice. *Int J Epidemiol.* 1999;28(1):106-12. <https://pubmed.ncbi.nlm.nih.gov/10195673/>.
25. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis.* 2020;93:284–286. <https://doi.org/10.1016/j.ijid.2020.02.060>.
26. Maleki M, Mahmoudi MR, Wraith D, Pho K-H. Time series modelling to forecast the confirmed and recovered cases of COVID-19. *Travel Med Infect Dis.* 2020, 101742. <https://doi.org/10.1016/j.tmaid.2020.101742>.
27. Shephard N. The relationship between the conditional sum of squares and the exact likelihood for autoregressive moving average models. 1997. <http://www.nuffield.ox.ac.uk/Economics/papers/1997/w6/ma.pdf>.
28. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G, editors. *Selected Papers of Hirotugu Akaike*. New York, NY: Springer New York. 1998;199-213.
29. Wang Y-w, Shen Z-z, Jiang Y. Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *PLoS ONE.* 2018;13(9):e0201987. <https://doi.org/10.1371/journal.pone.0201987>.
30. Ömer Faruk D. A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence.* 2010;23(4):586-94. <https://doi.org/10.1016/j.engappai.2009.09.015>.
31. Mondal P, Shit L, Goswami SJIJoCS, Engineering, Applications. Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. 2014;4(2):13. <https://doi.org/10.5121/ijcsea.2014.4202>.
32. Gibbons CL, Mangan MJ, Plass D, Havelaar AH, et al. Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health.* 2014;14(1):147. <https://doi.org/10.1186/1471-2458-14-147>.
33. Harapan H, Itoh N, Yufika A, Winardi W, et al. Coronavirus disease 2019 (COVID-19): A literature review. *J Infect Public Health.* 2020;13: 667–673. <https://doi.org/10.1016/j.jiph.2020.03.019>
34. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med.* 2020;382:727-33. <https://doi.org/10.1056/NEJMoa2001017>.
35. Chen N, Zhou M, Dong X, Qu J, et al. Epidemiological and clinical characteristics of 99 cases of 2019 Loading [MathJax]/jax/output/CommonHTML/jax.js , China: a descriptive study. *Lancet.* 2020;395(10223):507-

513. [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7).

36. Rodriguez-Morales AJ, Cardona-Ospina JA, Gutiérrez-Ocampo E, Villamizar-Peña R, et al. Clinical, laboratory and imaging features of COVID-19: A systematic review and meta-analysis. *Travel Med Infect Dis*. 2020;34:101623. <https://doi.org/10.1016/j.tmaid.2020.101623>.
37. Kuiken T, Fouchier RA, Schutten M, Rimmelzwaan GF, et al. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet*. 2003;362:263–70. [https://doi.org/10.1016/S0140-6736\(03\)13967-0](https://doi.org/10.1016/S0140-6736(03)13967-0).
38. You C, Deng Y, Hu W, Sun J, et al. Estimation of the time-varying reproduction number of COVID-19 outbreak in China. *Int J Hyg Env Health*. 2020;113555. <https://doi.org/10.1016/j.ijheh.2020.113555>.

Figures

Figure 1

The estimates of the number of under-ascertainment cases between 22 and 27 April 2020, and the basic reproduction number (R_0). Panel (a) shows 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 and R_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.

Figure 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) show the exponential growth fitting results of the cumulative number of cases (C_i) and the daily number of cases (σ_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.

Figure 3

15-days advance forecast of cumulative number of COVID-19 cases (red curve and red area) and observed data (blue dots and black curve). The 95% confidence interval is marked in red.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendices.pdf](#)