



INTEGRATION OF BAYESIAN MODEL AND ADAPTIVE CLUSTERED SAMPLING INTO CONTACT TRACING TO CURB THE SPREAD OF COVID -19 CASES

Olayiwola, O. M.^{1*}, Adekeye, K. S.², Apantaku F. S.,³ Ajayi A. O.⁴, Wale-Orojo O. A.⁵, Ogunsola I. A.⁶ and Hamed B.⁷

^{*1,3,4,5,6}Department of Statistics, College of Physical Sciences, Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.

²Department of Statistics, Redemer University, Ede, Osun State, Nigeria.

⁷Department of Mathematics and Statistics, School of Applied Science, Lagos State University, Ikorodu Lagos, Nigeria.

*Corresponding Author's Email: olayiwolaom@funaab.edu.ng

ABSTRACT

Covid-19 is a communicable virus that causes serious illness (Severe acute respiratory syndrome (SARS)) and middle east respiratory syndrome (MARS)). Its outbreak started in Wuhan, China on December 8, 2019. Fever, cough, tiredness are its signs and symptoms and appear between two to fourteen days after exposure. The severity of COVID-19 can include complications; pneumonia, heart problems, acute kidney injuries. Covid-19 careers should be identified in order to curb the spread of the virus within a population. In this regards, contact tracing is the current technique in use to identify and track the Covid-19 carriers. The aim is to curb the spread of the virus within the population. In order to achieve this goal effectively, appropriate technique is required in the identification of Covid-19 carriers and Modeling. It is known that Covid-19 carriers are hidden, clustered and very difficult to identify in the population. At this point, the Adaptive Cluster Sampling, which is a specialized sampling for identification of hidden and clustered event and Bayesian Model, comes to the practice. Therefore, in this study, Adaptive Cluster Sampling which is capable of tracking hidden and clustered events and Bayesian Model are integrated in contact tracing, and the application on how this technique is used is included.

Keywords: Covid-19, virus, Contact Tracing, Adaptive Cluster Sampling, Bayesian Model.

INTRODUCTION

Covid-19 has been termed pandemic due to the sharp increase in the number of Covid-19 carriers outside China and affecting several numbers of countries across the globe. Covid-19 was caused by a new strain of coronavirus (SARS-CoV-2) first discovered in Wuhan, China, having not been seen earlier in humans (Channappanavar (2017), WHO (2020)), on 31st December 2019. The confirmed cases had increased to 17,755,732 and 677, 752 deaths across the world as at 31st July 2020, 10:48 GMT (Oruonye and Ahmed (2020)). It spreads through droplets generated from a carrier to the other. This virus infects people of all ages. It caused many symptoms, such as illness, difficulty in breathing and even deaths (Yin, 2018).

However, older people with underlying medical conditions are at a higher risk.

According to Worldometer (2020) the total confirmed cases of Covid-19 started increasing from April 2020 across the globe (see figures 1 and 2). From Figures 1 and 2, it is clear that the number of confirmed cases follows the exponential pattern, which implies a growth curve over time. Many measures such as lockdown, physical distancing, closure of schools, markets and places of worship were adopted, still confirm cases are still increasing.

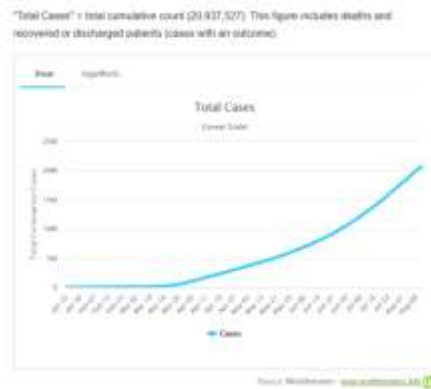


Figure 1: Total confirmed cases in the world

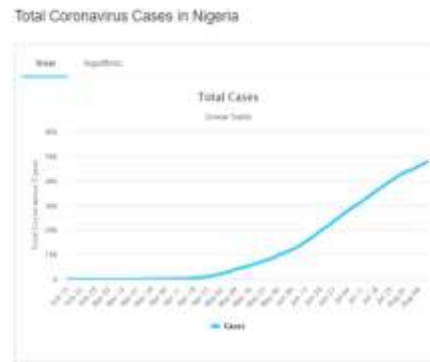


Figure 2: Total confirmed cases in Nigeria.

Contact Tracing Process

Contact tracing is the technique of tracking people who have been exposed to Covid-19 carriers to prevent continuous flow of the virus WHO (2020). It is one of the greatest strategies used in detecting carriers of epidemic diseases such as Covid-19 over the years. Comprehensive techniques, identification and confirmation of cases, isolation of carriers, testing, contact tracing and quarantine of carriers are essential activities to prevent the spread of Corona Virus Diseases. WHO (2020) infers that the method requires systematically identifying people who may have been in contact with Covid-19 carrier and monitoring them for fourteen days from the last point of exposure.

The estimated number of contact persons to be traced, the monetary, technological and physical logistics of getting to the affected communities and contact persons, behavioural and cultural context, socio-political factors, security issues and methods of contact tracing are factors to be considered in composition of workforce requirements for contact tracing. Contact tracing teams include epidemiologist, contact follow up team, Investigation team and Statisticians (WHO, 2015). These teams may require identity cards, transport bus, electronic or paper materials to record information, mobile telephones and credits, masks, hand sanitizers, and gloves.

There are five most prominent contact tracing challenges (Ashley, 2015). Though, they depend on each country's unique economic, social, geopolitical, and cultural context. There is situation of not at home; hard to reach; non response due to sensitive issues; hard core, and many other factors related to non-response in contact tracing of hidden Covid-19 carriers, which require statistical techniques for handling non response and identification of hidden Covid-19 carriers. Hence, this study proposed integration of Adaptive Cluster Sampling into contact tracing to curb the spread of Covid-19 cases and ultimately, this will lead to prevention of Covid-19 outbreaks and putting the cases into final extinction.

Effects of Covid-19 on Economy, Employment, Education, Tourism and Others

The effects of Corona virus diseases (Covid-19) are beyond imagination as it came in at an unexpected period. The whole world is at a great state of pandemic as no one could predict or

stipulate its end time. Great and developed nations with buoyant economy were seriously affected let alone the underdeveloped and developing ones. Many social and economic activities were postponed and seriously crippled. Economic sectors, tourism industries, educational institutions and even religious centres were not left out (CSEA, 2020).

Various researchers had shown that economic meltdown or crisis is often caused by external trade and price shocks, political instability and civil unrest Oruonye and Ahmed (2020). There is no doubt that Covid-19 has also shown that a pandemic could also be a cause for economic crisis. Covid-19 pandemic affected in many ways such as incapability of borrowers' to service loans, decline in oil price, shortage in crucial supplies (e.g. pharmaceutical supplies, spare parts e.t.c.) due to closure of borders, national budget deficit as a result of low oil price and high stock market investors lost (PWC (2020)).

The emergency of the pandemic forced many countries to put some policies in place to curb its spread, such as total lockdown, social distance rules, curfew, and ban on interstate, local and foreign travels (Bishnu, 2020). These policies imposed restrictions in spending and shortage in consumptions. Many industries and companies cut down production and shutdown services rendered while some lay off their staff to reduce personnel costs. These actions resulted into drastic increase in unemployment rate and poverty level.

The children and aged people are the most vulnerable people to Covid-19 (Peterson (2020), WHO (2020)). As a result of this, primary, secondary schools and tertiary institutions were shutdown. This affected the pupils and students intellectually. Teachers and even staff of higher institutions were redundant. Many conferences and seminars were suspended even though some were replaced with online mode. Nevertheless, the pandemic diseases had put a stop to many academic activities in all levels of education in the country.

Adaptive Cluster Sampling

Adaptive cluster sampling (ACS) was developed for hidden and clustered populations. (Thompson, 1990) In ACS the effort is excessively allocated to areas where the hidden event (Covid-19 carriers) in sample unit i meets a pre-specified condition. This allows the contact tracing team to concentrate where the Covid-

19 carrier is present or sufficiently abundant. The final sample size in ACS is random but controlled by the initial selection method and sample size, the form to adapt, and the locality model of adaptively sampled units. Brown and Manly (1998). The ACS is more precise estimates of population parameters when populations are geographically hidden and clustered.

Covid-19 carriers are hidden and clustered, based on this prior knowledge about Covid-19, Bayesian model is appropriate to model it with ACS as the sampling procedure (Goncalves et al (2016)). Thompson introduced the ACS and aggregated the strategy of maximizing survey effort where hidden event is present, developed a design-based analysis of the resulting sampling scheme and showed that adaptive cluster sampling is the better than the convention grid cell sampling (Thompson 1990, 2017). Rare tree species in Nepal was sampled (Acharya et al., 2000) and found Adaptive Cluster Sampling (ACS) to be efficient.

MATERIALS AND METHODS

Model for Estimating Number of Covid-19 Carriers

For Covid-19 contact tracing team to have an idea of the estimated number of Covid-19 carriers in the population, this study fitted a Bayesian model with ACS embedded.

Let a be the expected number of Covid-19 carriers in a country, b be the conditional expected number of country with Covid-19 carriers, λ be the number of confirmed Covid-19 cases in each country, and the $Y = Y_1, Y_2, \dots, Y_p$ the number of country in each Covid-19 carrier continent, where a continent is a collection of one or more countries with Covid -19 carriers.

Bayesian model was fitted using R package (Rapley and Welsh, 2008) for the Number of country in continent, Q by specifying the joint distribution of X, P, Y and N. The ACS mechanism was embedded and lead to adaptive samples = $\{i_1, \dots, i_m\}$ of m out of $M - X + P$ countries. Where M = number of countries, a country is nonempty if it contains at least one Covid-19 carrier and empty otherwise. Let $X \leq M$ be the number of countries with Covid-19 carriers in Q. Let $P \leq X$ be the number of Covid-19 continents (collection of neighbouring countries in Q and $Y = Y_1, Y_2, \dots, Y_p$ denote the number of countries with Covid-19 carriers within each continents so that $X = \sum_{i=1}^p Y_i$. As there are $M - X$ countries without Covid-19. Therefore, there are $M - X + P$ continents. Joint distribution of X, P, Y and N for Q was specified and then we model the structure of continents

with/without Covid-19 carriers conditionally on the continent structure and the number of Covid-19 carriers (N) continents of countries with Covid-19 Olayiwola et al. (2020). The unknown parameters a, b and λ in the model were estimated using a Markov Chain Monte-Carlo (MCMC) (Gelman and Rubin, 1996) algorithm with Winbugs software. Trace plot and correlogram were used for the MCMC Diagnostics to examine goodness of fit of the model. The fitted model was used to obtain a predictive distribution for predicting N. The predictor of N (estimated total number of Covid-19 carriers) is given as:

$$\hat{N} = 1_{p_0}^T N_0 + 1_{p_1}^T N_1 \dots \dots \dots (1)$$

Where T is , N₀ and N₁ is

Proposed Contact Tracing with integration of Bayesian Model and Adaptive Clustered Sampling

WHO in 2015 had suggested involvement of Statisticians in contact tracing team. The Centre for Disease Control (CDC) team members (medical personnel) are solely responsible for contact tracing, which made their effort inefficient to capture all carriers. This study proposed integration of Bayesian Model and Adaptive Cluster Sampling techniques in contact tracing to estimate and capture hidden Covid-19 carriers. The proposed methodology suggested Statisticians to be part of the CDC team to make its effort effective to curb the spread of Covid-19 cases.

Implementation of the Proposed Contact Tracing with integration of Bayesian Model and Adaptive Clustered Sampling

The model in equation (1) will guide statisticians to know the estimated number of Covid-19 carriers within a population together with appropriate sample to be selected strategically and adaptively such that all carriers would be captured. Immediately a carrier or carriers is/are identified, contact tracing will commence by the team of medical personnel with statisticians. All selected samples by contact tracing process and Adaptive Cluster Sampling would be subjected to isolation for 14 days or test depending on availability of the test equipment. These samples would be observed for 14 days and any one with the symptom of Covid-19 would be tested. If positive, admit and treat and if negative discharge. After 14 days, others without the symptoms would be discharged. The proposed procedure is presented in a flow chart below

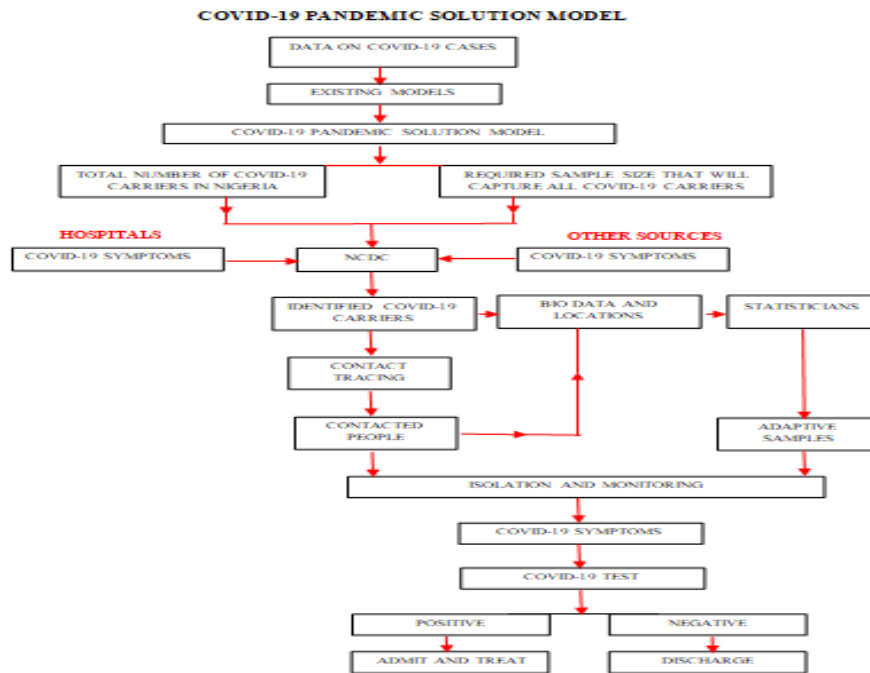


Figure 3: Flow chart for Contact Tracing with integration of Bayesian Model and Adaptive Clustered Sampling

Computation and Interpretation of Reproduction Rates (R) of Covid-19

Involvement of Statisticians in contact tracing will help to get virus. R is a way of rating disease’s ability to spread. It is the number of people one infected person will pass the virus on to, on average. If $R > 1$, it implies the number of cases increases exponentially

If $R < 1$, it implies the number of cases will eventually decline, and if $R = 1$, the spread neither increases nor decreases.

Recommended computation procedure of the Reproduction Rate, R, is as follow

Let:

- n = number of Covid-19 carriers
- m = number of locations of carriers
- T = Contact Tracing process
- A = Adaptive Cluster Sampling
- g = number of priority groups
- n_{ijT} = number of confirmed cases from carrier i contacts at location j captured by contact tracing process

$C_{E_{ij}A}$ = number of confirmed cases from priority group i at location j captured by Adaptive Cluster Sampling,

Then,

$$R = \frac{\sum_{i=1}^n \sum_{j=1}^m C_{ijT} + \sum_{i=1}^n \sum_{j=1}^m C_{E_{ij}A}}{\sum_{j=1}^m C_j} \tag{2}$$

Figure 4, represents proposed contact tracing technique which

appropriate data for computation of R, which will guide policy makers about the spread rate of the comprises Statisticians in CDC team. The inner circle is the focus of contact tracing by CDC team. In this region, as the medical personnel is tracing the contact to carriers, statisticians in their midst will adopt their professional skills to handle all challenges of not at home, hard to reach, non- response due to sensitive issues, hard core, lack of knowledge of the estimated number of Covid-19 carriers and lack of appropriate strategy for sample selection.

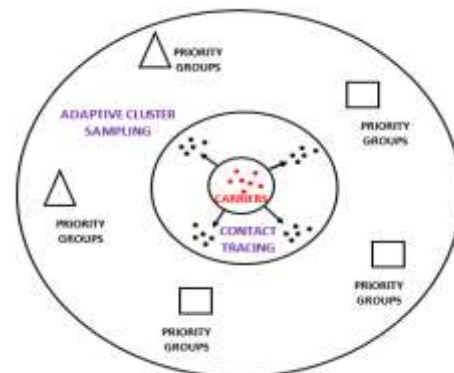


Figure 4: Contact tracing

The outer circle which, comprises of the priority groups, such as big Super Markets, Churches, Mosques, major car park etc,

will be the focus of Statisticians. Statisticians will adopt Adaptive Cluster Sampling to strategically and randomly select adaptive samples from priority groups in the vicinity of identified carriers. Combination of samples from contact tracing and adaptive samples would be subjected to isolation or test depending on availability of test equipment. With this process, required data for computation of R in Eq (2) would be generated and the value of R would be determined and interpreted for decision making to monitor the spread rate of the virus.

RESULTS AND DISCUSSION

Evaluation of the proposed model using Covid-19 cases in Nigeria

The fitted model for predictor of Covid-19 carriers in Equation (1) converged to representation of the target posterior within the 95% highest posterior density (HPD) interval as shown in Figure 4. Trace plot and correlogram in Figure 5 and Figure 6 respectively showed that the chains from the model mixed well and its autocorrelation is quite similar at each lag, which confirmed that the model is efficient for estimating number of Covid-19 carriers in Nigeria.

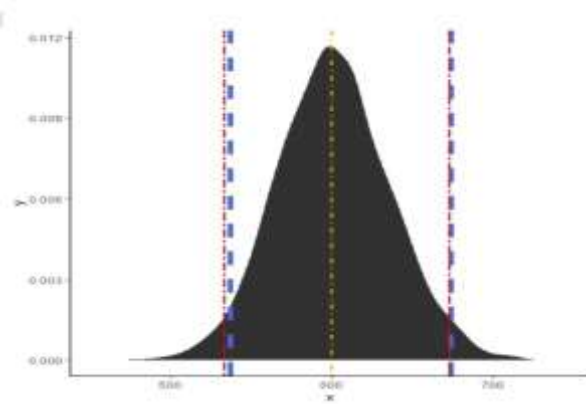


Figure 5: 95% Credible Interval for Posterior Total Estimation for Covid-19 carriers

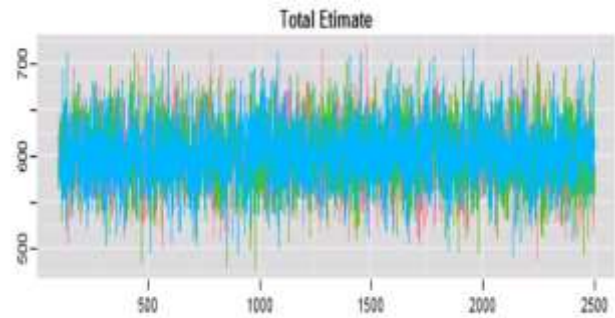


Figure 6: Trace plot

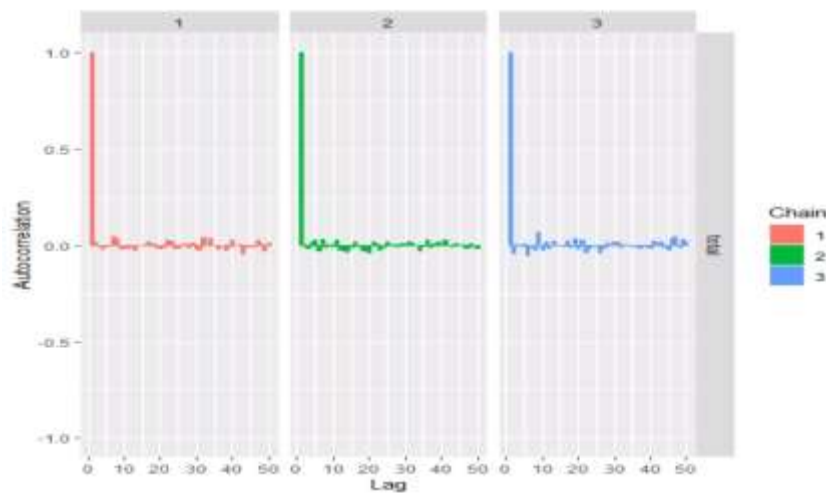


Figure 7: Autocorrelation plot

The predictor in equation (1) gave an estimated number of Covid-19 carriers in Nigeria for July 31, 2020, as shown in Table

1. It was observed in Table 1 and Figure 8, that for each state, the estimated number of carriers is higher than confirmed cases tracked by the NCDC team without involvement of Statisticians

Nairametrics (2020). Thus, the NCDC team (medical personnel) faced which could only be resolved by Statisticians. were unable to capture all the carriers due to challenges they

Table 1: Estimated numbered Covid-19 carriers versus confirmed cases in Nigeria for July 31, 2020.

State	Estimated number of Covid-19 carriers	Confirmed Covid-19 cases
Abuja	126	93
Lagos	109	78
Oyo	71	47
Edo	18	7
Rivers	20	9
Kano	16	6
Delta	21	9
Kaduna	94	54
Ogun	24	9
Plateau	92	64
Ondo	52	32
Enugu	18	6
Bauchi	35	19
Nasarawa	15	5
Adamawa	40	23
Osun	3	1
Cross River	15	0
Zamfara	7	0
Yobe	7	0
Taraba	42	0
Ebonyi	21	0
Kwara	13	0
Katsina	21	0
Borno	7	0
Gombe	54	0
Abia	21	0
Imo	9	0
Benue	21	0
Bayelsa	7	0
Jigawa	7	0
Niger	74	0
Akwainom	27	0
Sokoto	7	0
Anambra	11	0
Ekiti	20	0
Kebi	7	0
Kogi	7	0

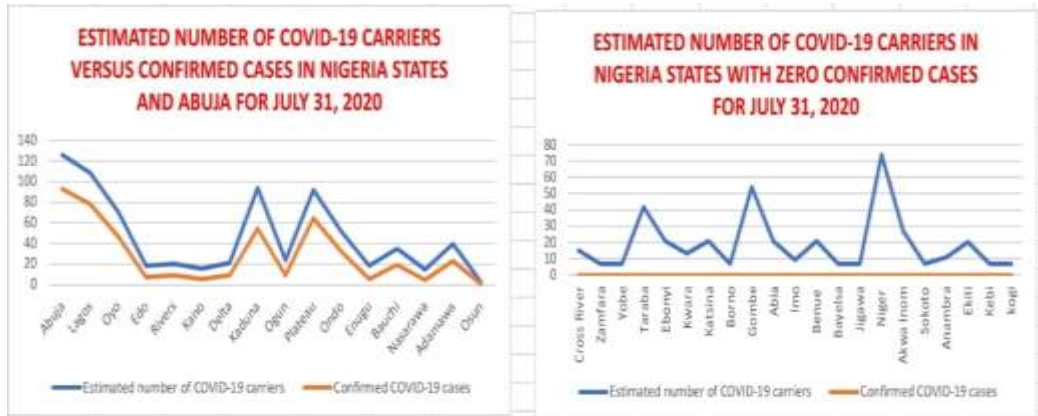


Figure 8: Estimated number of Covid -19 carriers versus confirmed cases

On July 31, 2020, estimated number of Covid-19 carriers in Nigeria was 1159 while confirmed cases was 462 as shown in

Figure 9. This justified that contact tracing process by medical personnel is inefficient to capture carriers. It requires integration of Statistical techniques for carriers to be properly tracked and captured.

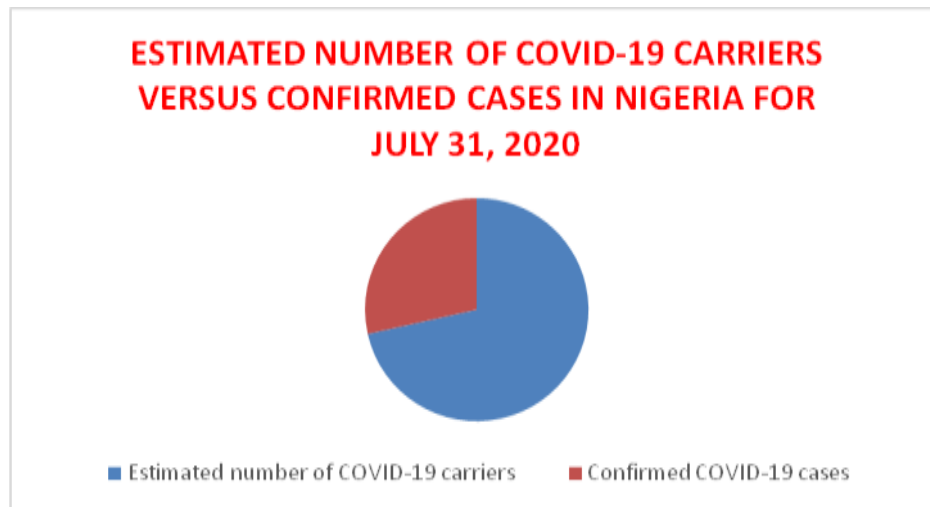


Figure 9: A pie chart showing estimated number of COVID-19 carriers versus confirmed cases in Nigeria

Figure 10 presents the distribution of estimated Covid-19 carriers among Nigeria states for July 31, 2020, Covid-19 carriers are higher in Lagos State and Abuja, while the least is

from Osun State. Figure 9 shows that majority of the states have zero confirmed cases and where there is confirmed cases, it is less than estimated number of carriers, and hence present contact tracing process by medical personnel is inefficient to track Covid-19 carriers in Nigeria.

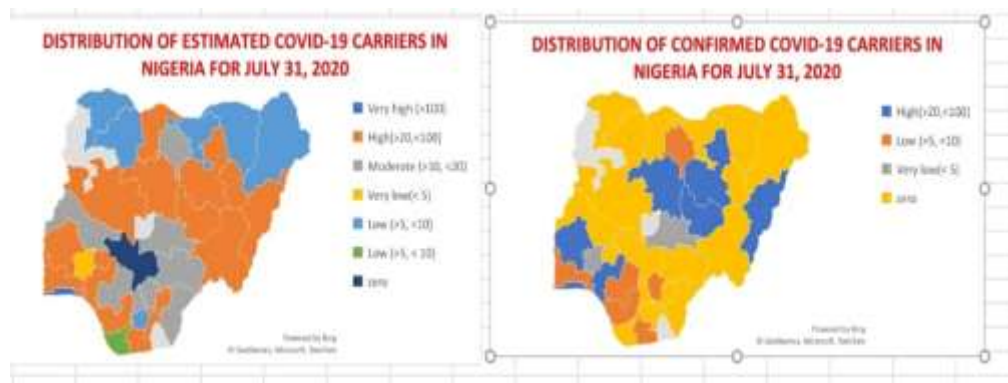


Figure 10: Distribution of Estimated and Confirmed COVID-19 cases carriers in Nigeria for July 31, 2020

CONCLUSION AND SUGGESTIONS

Involvement of Statisticians in contact tracing team was suggested by WHO (2015). The CDC team members (medical personnel) are solely responsible for contact tracing, which made their effort inefficient to capture all carriers. This study has provided the use of Bayesian Model and Adaptive Cluster Sampling techniques in contact tracing to estimate and capture hidden Covid-19 carriers. The proposed methodology suggested Statisticians to be part of the CDC team to make its effort effective to curb the spread of Covid-19 cases.

REFERENCES

- Acharya B, Bhattarai G, Gier A, Stein A. (2000). Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. *Forest Ecology and Management* 137(2000):65-73.
- Ashley L. Greiner, et. al. (2015). Addressing contact tracing challenges – Critical to halting Ebola Virus Diseases transmission. *International Journal of Infectious Diseases*. Vol. 41, Page 53 – 55.
- Bishnu P. K. (2020). Impact of the COVID-19 in Tourism Industry in Nepal and Policy Recommendation. *Journal of Tourism & Adventure*. 3:1, 76-91
- Brown J. A. and Manly B.J.F. (1998). Restricted adaptive cluster sampling. *Environmental and Ecological Statistics* 5:49-63.
- Centre for the Study of Economies of Africa (CSEA). (2020). The implication of Covid-19 on the Nigerian Economy. Retrieved April 13th, 2020, from <http://cseaafrica.org/the-implication-of-covid19-on-the-nigerian-economy>
- Channappanavar R, Perlman S. (2017). Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. *Seminars in immunopathology*. 39(5):529-39.
- Gelman, A. and Rubin, D. (1996) Markov chain Monte Carlo methods in biostatistics. *Statistical Methods in Medical Research*, 5, 339–355
- Goncalves, Kelly C. M., Moura, Fernando A. S. (2016). A Mixture Model for Rare and Clustered Populations Under Adaptive Cluster Sampling. *Bayesian Analysis* 11, Number 2, pp. 519–544.
- Nairametrics (2020). <https://nairametrics.com/2020/08/03/Covid-19-update-in-nigeria/>
- Olayiwola O. M. et al. (2020). Adaptive cluster sampling with model-based approach for estimating total number of hidden COVID-19 carriers in Nigeria. *Statistical Journal of the IAOS*. Vol. 36. S103–S109 S103.
- NCDC (2020), <https://covid19.ncdc.gov.ng>. Accessed June 18th, 2020
- Oruonye E.D. & Ahmed Y.M. (2020). An appraisal of the potential impacts of COVID-19 on Tourism in Nigeria. *Journal of Economics and Technology Research*. Vol. 1 no 1. Pages 32 – 4
- Peterson K. Ozili (2020) COVID-19 pandemic and economic crisis: The Nigerian experiences Structural causes. *Munich Personal RePEc Archive*
- PwC (2020): Nigeria's Webinar COVID-19: Economic Implication and Policy Responses
- R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Rapley, V. and Welsh, A. (2008). Model-Based Inferences from Adaptive Cluster Sampling. *Bayesian Analysis*. 3(4): 717–736. MR2469797. doi: <http://dx.doi.org/10.1214/08-BA327>. 520, 521, 522, 523, 535, 536, 539, 540
- Thompson, S. (1990). "Adaptive cluster sampling." *Journal of the American Statistical Association* 85: 1050–1059.

Thompson S. K (2017): Adaptive and Network Sampling for Inference and Interventions in Changing Populations. *Journal of Survey Statistics and Methodology*. 5 (1): 1-21. 10.1093/jssam/smw035.

World Health Organization (2015). Implementation and management of contact tracing for Ebola Virus disease (September, 2015).

World Health Organization (2020). Contact tracing in the context of Covid-19, interim guidance.

Worldometer (2020).
<https://www.worldometers.info/coronavirus/#countries>.
Accessed on June 23, 2020 2, 2020.

World Health Organization (WHO). Novel Coronavirus – China 2020 [14 January 2020]. Available from: <https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/>.

Yin Y, Wunderink RG. (2018). MERS, SARS and other coronaviruses as causes of pneumonia. *Respirology (Carlton, Vic)*. 23(2):130-7.



©2021 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.