

COVID-19 PREDICTION IN SOUTH AFRICA: ESTIMATING THE UNASCERTAINED CASES- THE HIDDEN PART OF THE EPIDEMIOLOGICAL ICEBERG

XUELIN GU

Department of Biostatistics, University of Michigan, Ann Arbor, USA
Email: xuelingu@umich.edu

BHRAMAR MUKHERJEE

Department of Biostatistics, Department of Epidemiology, University of Michigan, Ann Arbor, USA
Email: bhramar@umich.edu

SONALI DAS

Department of Business Management, University of Pretoria, Pretoria, South Africa
Email: sonali.das@up.ac.za

JYOTISHKA DATTA*

Department of Statistics, Virginia Polytechnic Institute and State University, USA
Email: jyotishka@vt.edu

SUMMARY

Understanding the impact of non-pharmaceutical interventions as well as accounting for the unascertained cases remain critical challenges for epidemiological models for COVID-19 spread. In this paper, we propose a new epidemiological model (eSEIRD) that extends the widely used extended Susceptible-Infected-Removed model (eSIR) and SAPHIRE models. We fit these models to the daily ascertained infected and removed cases from March 15, 2020 to December 31, 2020 in South Africa, the ‘worst-hit’ country in the WHO African region. Using the eSEIRD model, the COVID-19 transmission dynamics in South Africa was characterized by the estimated basic reproduction number (R_0) starting at 3.22 (95%CrI: [3.19, 3.23]) then dropping below 2 (95%CrI: [1.36, 1.39]) following a mandatory lockdown implementation and subsequently increasing to 3.27 (95%CrI: [3.27, 3.27]) by the end of 2020. The estimated trajectory of R_0 suggests the effect of early interventions and the subsequent relaxation and emergence of a new coronavirus variant. The estimated ascertainment rate was found to vary from 1.65% to 9.17% across models and time periods. The overall infection fatality ratio (IFR) was estimated as 0.06% (95%CrI: [0.04%, 0.22%]) accounting for unascertained cases and deaths while the reported case fatality ratio was 2.88% (95% CrI: [2.45%, 6.01%]). The models predict that from December 31, 2020, to April 1, 2021, the predicted cumulative number of infected would reach roughly 70% of total population in South Africa. Besides providing insights on the COVID-19 dynamics in South Africa, we develop powerful forecasting tools that enable estimation of ascertainment rates and IFR while quantifying the effect of intervention measures.

Keywords and phrases: COVID-19; South Africa; forecasting; unascertained cases; under-reporting factors; infection fatality ratio

* Corresponding author

© Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.

1 Introduction

The coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was first reported in early December 2019 in Wuhan, China and then quickly spread to majority countries worldwide. At the end of January 2021, over a hundred million people worldwide have been diagnosed with COVID-19 (World Health Organization, 2021b), yet the true number of infections in the population remains underestimated, owing to a combination of selection bias from unascertained cases, and lack of access to tests early on during the pandemic.

South Africa: We focus on the COVID-19 transmission dynamics in South Africa, the ‘epicenter of the outbreak in the African continent’ (World Health Organization, 2020c). The first case was confirmed in South Africa on March 5, 2020. As of February 23, 2021, there are 1,504,588 confirmed cases of COVID-19 (cumulative total) with 49,150 deaths confirmed in Africa (World Health Organization, 2021a). South Africa remains the worst-hit African country with the largest number of confirmed cases and deaths, by the end of 2020, contributing to 54% of the total confirmed cases and 44% of deaths in the WHO African region, while accounting for only 5% of population (World Health Organization, 2021a). A seroprevalence survey on 4,858 blood donors in South Africa estimated the prevalence as of mid-January 2021, by province, as 63% in the Eastern Cape, 52% in the Kwa-Zulu Natal, 46% in the Free State and 32% in the Northern Cape (Sykes et al., 2021) while the number of reported cases is 1.83% of the total population during the same time, implying the possibility of a large degree of under-reporting/undetected cases in South Africa. Thus, understanding the key epidemiological constructs for COVID-19 outbreak is paramount for containing the spread of COVID-19 in South Africa, as well as explaining the disparity between seroprevalence estimates and reported number of cases. Two critical factors emerge from analyzing the majority of available evidence of the public health crisis: (1) the unascertained cases and deaths and (2) the role of non-pharmaceutical interventions.

Unascertained cases and deaths: Based on the clinical characteristics of COVID-19, a majority of patients are symptomatic (roughly 84% according to a recent study He et al., 2020a), most of whom have mild symptoms (World Health Organization, 2020c) and tend to not seek testing and medical care. While private hospitals have reached maximum capacity, public and field hospitals beds have still some margin left with additional challenges due to scarcity of staff (Daniel, 2020). Several recent studies (Hao et al., 2020; Rahmandad et al., 2020; Bhattacharyya et al., 2021) reported that a non-negligible proportion of unascertained cases contributed to the quick spreading of COVID-19. It is suggested that only 1 in 4 mildly ill cases would be detected in South Africa (South African COVID-19 Modelling Consortium, 2020). The relatively lower testing rate in South Africa (Table 1, Figure 1) coupled with a very high test positivity rate especially in July and August 2020 (Ritchie et al., 2020), suggests inadequacy of testing, as well as the possibility of a large unobserved number of unascertained cases (World Health Organization, 2020a). Thus, modeling both ascertained and unascertained cases and deaths can measure infection fatality ratios (IFRs, the proportion of deaths among all infected individuals (World Health Organization, 2020b)) of COVID-19, leading

to a better understanding of the clinical severity of the disease.

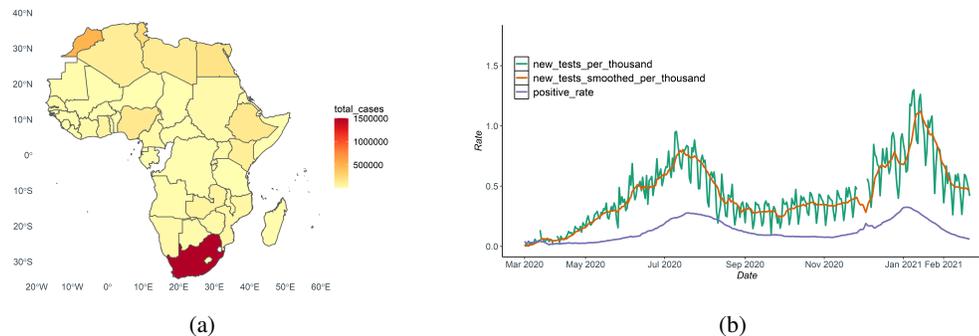


Figure 1: (a) Total cases by country in the African continent; (b) The 7-day average testing positive rate of COVID-19 in South Africa during the study period.

Interventions: With a universal goal to ‘flatten the curve’, a series of non-pharmaceutical interventions were implemented by the government in South Africa, that have been gradually lifted since early May 2020 (South Africa Department of Health, 2021). On March 27, 2020, South Africa adopted a three-week nationwide hard-lockdown (level 5) along with closure of its international borders, which was extended to April 30, 2020. Thereafter, to balance the positive health effects of strict interventions against their economic costs (Arndt et al., 2020), South Africa began a gradual and phased recovery of economic activities with the lockdown restriction eased to level 4 (South Africa Department of Health, 2021), allowing inter-provincial travel only for essential services. From June 1, 2020, national restrictions were lowered to level 3 allowing for inter-provincial travel and school opening, and eased to level 2 and level 1 from August 18 and September 21, 2020 (Table 1). Face-mask wearing remained mandatory in public places at all times, with limitations on gatherings, and sale of alcohol and cigarettes were restricted under harder lockdown (Garba et al., 2020). Although these interventions implemented at an early stage had a higher potential for pandemic containment, previous studies (Garba et al., 2020; Mukandavire et al., 2020; Zhao et al., 2020) reported a consistently large value for the estimated basic reproduction number (R_0) ranging from 2.2 to 3.2 in South Africa by models trained with data in relatively early time windows. Using data observed under various intervention scenarios over a longer period of time, we carry out a thorough investigation to assess the current COVID-19 spread and the effect of these interventions, which will provide valuable insights into the transition dynamics of COVID-19 and intervention deployment in South Africa, and beyond.

Epidemiological models: Since the early days of the pandemic, researchers have responded to the unprecedented public health crisis by providing forecasts and alternative scenarios to inform decision-making, both locally and globally. This has resulted in hundreds of mathematical models of varying complexity. The Susceptible-Infectious-Removed or SIR model (Kermack et al., 1927)

Table 1: Timeline of COVID-19 preventions and interventions in South Africa.

Date (2020)	Confirmed	Death	Testing rate	Interventions and update
5-Mar	1	0	-	-
10-Mar	7	0	-	Screening at ports of entry has intensified and escalated.
15-Mar	51	0	-	Self-quarantine for COVID-19 is recommended. Visas to visitors from high-risk countries (Italy, Iran, South Korea, Spain, Germany, US, UK) are cancelled and previously granted visas are hereby revoked. Gatherings of more than 100 are prohibited. Mass celebrations are canceled.
16-Mar	62	0	-	Of the 53 land ports, 35 are shut down.
18-Mar	116	0	-	A travel ban on foreign nationals from high-risk countries such as Germany, US, UK and China.
27-Mar	1170	1	-	A national lockdown is implemented. Alert level 5 is in effect from midnight 26 March to 30 April.
1-May	5951	116	0.004	A less strict lockdown is in place. Alert level 4 is in effect from 1 to 31 May. Borders will remain closed to international travel, no travel is allowed between provinces, except for the transport of goods and exceptional circumstances.
1-Jun	34,357	705	0.013	From 1 June 2020 alert level 3 is in effect. Restrictions on many activities, including at workplaces and socially, to address a high risk of transmission.
18-Aug	592,106	12,264	0.059	Alert level 2 is in effect.
21-Sep	661,898	15,992	0.07	Alert level 1 is in effect.

is arguably the most commonly used epidemiological models for modeling the trajectory of an infectious disease, and has been used to model the pandemic trajectory in South Africa (Ding et al., 2021). However, the ordinary SIR model does not take time-varying non-pharmaceutical interventions and quarantine protocols into consideration. To address this drawback, a recent extension called extended-SIR or eSIR (Song et al., 2020), was developed with a Bayesian hierarchical Beta-Dirichlet state-space framework, which was successfully applied to model COVID-19 dynamics in India (Ray et al., 2020). One major advantage of this Bayesian hierarchical structure is that uncertainty associated with all parameters and functions of parameters can be calculated from posterior draws without relying on large-sample approximations (Ray et al., 2020). Extending the simple compartment structure in eSIR model, the SAPHIRE model (Wang et al., 2020a) delineated the full transmission COVID-19 dynamics in Wuhan, China with additional compartments, by introducing

unobserved categories (Hao et al., 2020).

In this article, we extended the eSIR approach to the eSEIRD model to combine the advantages of the two existing models: using a Bayesian hierarchical structure to introduce additional unobserved compartments and characterize uncertainty in critical epidemiological parameters including basic reproduction number, ascertainment rate and IFR, with input data as observed counts for cases, recoveries and deaths. Furthermore, we applied these three models and compared the results of the eSEIRD model with two of the existing alternatives, namely the eSIR and the SAPHIRE model, with the following primary objectives: (i) characterizing the COVID-19 dynamics from March 15 to December 31, 2020 – under different time-varying intervention scenarios; (ii) evaluating the effectiveness of the main non-pharmaceutical interventions such as lockdown, and mandatory wearing of face-mask in public places; (iii) capturing the uncertainty in estimating the ascertainment rate and IFR; and (iv) forecasting the future of COVID-19 spread in South Africa.

The organization of this paper is as follows: we describe the two competing epidemiological models and our proposed extension in Section 2. The study design and parameter settings for modeling COVID-19 transmission in South Africa are described in Section 3, and the results and their possible implications are described in Section 4. We conclude with a discussion of nuances and limitation of the methods and sources of data used here and suggest future directions in Section 5.

2 Statistical Methodology

We propose an extension of the eSIR model, called eSEIRD, and compare it against two existing epidemiological methods, the eSIR and SAPHIRE model. In this section, we describe the dynamic systems and the hierarchical models underlying these three epidemiological models. The schematic diagrams for the three compartmental models are shown in Fig. 2 (eSIR and SAPHIRE) and Fig. 3 (eSEIRD).

2.1 eSIR model

The eSIR model assumes the true underlying probabilities of the three compartments susceptible (S), infectious (I) and removed (R) follow a latent Markov transition process and require observed daily proportions of cumulative infected and removed cases as input (Osthus et al., 2017; Song et al., 2020). The observed proportions of infected and removed cases on day t are denoted by Y_t^I/N and Y_t^R/N (the infected and removed counts Y_t^I and Y_t^R divided by total population size N) respectively. Further, we denote the true underlying probabilities of the three compartments on day t by θ_t^S , θ_t^I and θ_t^R , respectively, and assume that for any t , $\theta_t^S + \theta_t^I + \theta_t^R = 1$, and $\pi(t) \in [0, 1]$ is the transmission rate modifier that reflects the stringency of non-pharmaceutical intervention at time t . The following set of differential equations describe the dynamic system for the usual SIR model on the true proportions.

$$\frac{d\theta_t^S}{dt} = -\beta\pi(t)\theta_t^S\theta_t^I, \quad \frac{d\theta_t^I}{dt} = \beta\pi(t)\theta_t^S\theta_t^I - \nu\theta_t^I, \quad \frac{d\theta_t^R}{dt} = \nu\theta_t^I,$$

where $\beta > 0$ denotes the disease transmission rate, and $\nu > 0$ denotes the removal rate (see Fig. 2(a) for a schematic representation). The basic reproduction number $R_0 = \beta/\nu$ indicates the expected number of cases generated by one infected case in the absence of any intervention and assuming that the whole population is susceptible.

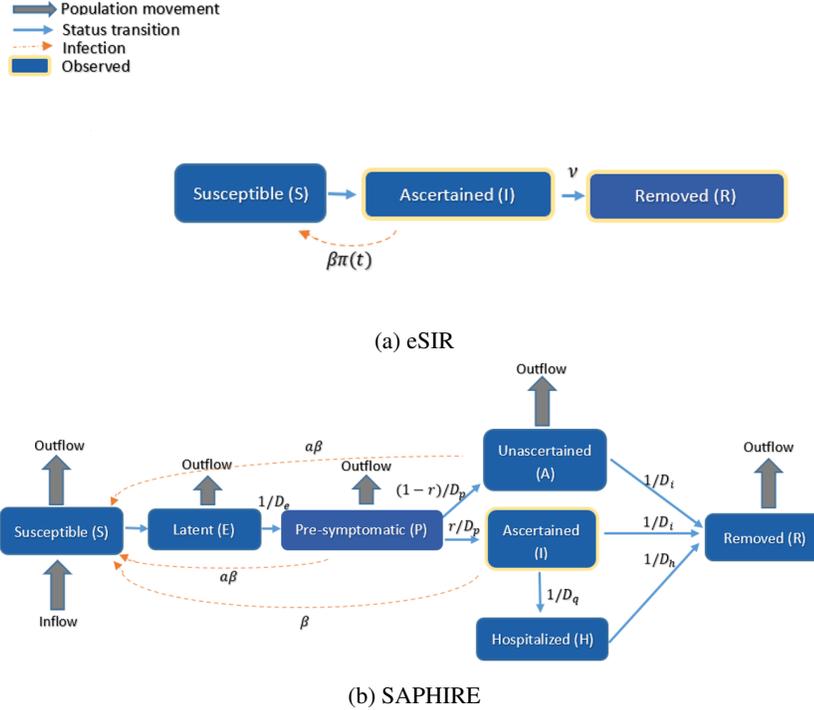


Figure 2: Schematic diagram of the two models (a) eSIR; (b) SAPHIRE.

Hierarchical model The eSIR model works by assuming that two observed time series of daily proportions of infected and removed cases are emitted from two Beta-Dirichlet state-space models, independent conditionally on the underlying process governed by the Markov SIR process:

$$(Y_t^I/N) \mid \theta_t, \tau_1 \sim \text{Beta} \left(\lambda^I \theta_t^I, \lambda^I (1 - \theta_t^I) \right),$$

$$(Y_t^R/N) \mid \theta_t, \tau_1 \sim \text{Beta} \left(\lambda^R \theta_t^R, \lambda^R (1 - \theta_t^R) \right),$$

and the Markov process associated with the latent proportions is built as:

$$\theta_t \mid \theta_{t-1}, \tau_1 \sim \text{Dirichlet} \left(\omega f_1(\theta_{t-1}, \beta, \nu) \right),$$

where θ_t denotes the vector the true underlying probabilities of the compartments on day t whose mean is modeled as an unknown function of the probability vector from the previous time point,

along with the transition parameters: $\tau_1 = (\theta_0^T, \beta, \nu, \lambda, \omega)$. The scale parameters λ^I, λ^R and ω control the variability of the observed and latent processes Y_t^I, Y_t^R and θ_t , respectively.

The function $f_1(\cdot)$ is then solved as the mean transition probability determined by the SIR dynamic system, using a fourth order Runge-Kutta (RK4) approximation (see supplementary Section S.3 for the solution):

$$f_1(\theta_{t-1}, \beta, v) = \begin{pmatrix} \theta_{t-1}^S + 1/6 [k_{t-1}^{S1} + 2k_{t-1}^{S2} + 2k_{t-1}^{S3} + k_{t-1}^{S4}] \\ \theta_{t-1}^I + 1/6 [k_{t-1}^{I1} + 2k_{t-1}^{I2} + 2k_{t-1}^{I3} + k_{t-1}^{I4}] \\ \theta_{t-1}^R + 1/6 [k_{t-1}^{R1} + 2k_{t-1}^{R2} + 2k_{t-1}^{R3} + k_{t-1}^{R4}] \end{pmatrix} := \begin{pmatrix} \alpha_{1(t-1)} \\ \alpha_{2(t-1)} \\ \alpha_{3(t-1)} \end{pmatrix}.$$

Computational details for the eSIR model such as posterior sampling strategy using MCMC algorithm is complemented by the R package publicly available at <https://github.com/lilywang1988/eSIR>.

2.2 SAPHIRE model

The SAPHIRE model (Hao et al., 2020) is an extension of the basic SIR model, with additional compartments to allow for unobserved categories such as, unascertained and hospitalized population. Specifically, in a SAPHIRE model the population is compartmentalized into susceptible (S), exposed (E), presymptomatic infectious (P), ascertained infectious (I), unascertained infectious (A), isolation in hospital (H) and removed (R). Denoting the true underlying accounts of the S, E, P, A, I, H and R compartments on day t by $S_t, E_t, P_t, A_t, I_t, H_t$ and R_t , respectively, the dynamics of these compartments across time t were described by the following set of differential equations (see Fig. 2(b) for a schematic representation):

$$\begin{aligned} \frac{dS_t}{dt} &= n - \frac{\beta S_t(I_t + \alpha(A_t + P_t))}{N} - \frac{nS_t}{N}, \\ \frac{dE_t}{dt} &= \frac{\beta S_t(I_t + \alpha(A_t + P_t))}{N} - \frac{E_t}{D_e} - \frac{nE_t}{N}, \\ \frac{dP_t}{dt} &= \frac{E_t}{D_e} - \frac{P_t}{D_p} - \frac{nP_t}{N}, \\ \frac{dA_t}{dt} &= \frac{(1-r)P_t}{D_p} - \frac{A_t}{D_i} - \frac{nA_t}{N}, \\ \frac{dI_t}{dt} &= \frac{rP_t}{D_p} - \left(\frac{1}{D_i} + \frac{1}{D_q}\right) I_t, \\ \frac{dH_t}{dt} &= \frac{I_t}{D_q} - \frac{H_t}{D_h}, \\ \frac{dR_t}{dt} &= \frac{A_t + I_t}{D_i} + \frac{H_t}{D_h} - \frac{nR_t}{N}. \end{aligned}$$

To fit the SAPHIRE model, the observed number of ascertained cases in which individuals experienced symptom onset on day t , Y_t^I , were assumed to follow a Poisson distribution:

$$Y_t^I \sim \text{Poisson}(\lambda_t^I = rP_{t-1}/D_p),$$

where P_{t-1} denotes the underlying number of pre-symptomatic individuals and r denotes the ascertained rate. For an observation window spanning $t = 1$ to $t = T$, the sampling pseudo-likelihood function for the underlying prevalence parameters is given by:

$$L(\boldsymbol{\beta}, \mathbf{r}) = \prod_{t=1}^T \frac{e^{-\lambda_t^I} \lambda_t^I Y_t^I}{Y_t^I!} \quad (2.1)$$

after plugging $\boldsymbol{\tau}_2 = (\mathbf{X}_0^T, \alpha, D_e, D_p, D_q, D_i, D_h)$ and $\mathbf{X}_t = (S_t, E_t, P_t, A_t, I_t, H_t, R_t)$. Here, \mathbf{X}_t denotes the vector of the underlying population counts of the compartments at time t . With the initial values for the compartments \mathbf{X}_0 set at pre-fixed values, the pseudo-likelihood function in (2.1) can be approximated as a function of the parameters of interest, i.e. $\boldsymbol{\beta}$ and r , by the following steps:

Step 0 The transition parameters $\boldsymbol{\tau}_2 = (\mathbf{X}_0^T, \alpha, D_e, D_i, D_h, D_q)$ and the initial values for the compartments $\mathbf{X}_0 = (S_0, E_0, P_0, A_0, I_0, H_0, R_0)$ are fixed;

Step 1 Use the differential equations to generate the change of each compartment at time $t = 1$, i.e. $d\mathbf{X}_t/dt = (dS_t/dt, dE_t/dt, dP_t/dt, dA_t/dt, dI_t/dt, dH_t/dt, dR_t/dt)$;

Step 2 Compute the state values of each compartment at time $t = 1$: $\mathbf{X}_t = \mathbf{X}_{(t-1)} + \frac{d\mathbf{X}_{t-1}}{d(t-1)}$, and the expected new ascertained cases λ_t^I on day t ;

Step 3 Repeat the step 1-3 for $t = 2, 3, 4, \dots, T$.

For the SAPHIRE model, the MCMC algorithm is implemented with the delayed rejection adaptive Metropolis algorithm implemented in the R package `BayesianTools` (version 0.1.7) to collect posterior samples of the underlying parameters r and $\boldsymbol{\beta}$, then calculate the derived quantities the effective reproduction number R_e based on the posterior draws. We refer the reader to (Hao et al., 2020) for more details.

2.3 eSEIRD model

Similar to the hierarchical structure used in eSIR model, this eSEIRD model (see Fig. 3) works by assuming that the true underlying probabilities of the 7 compartments follow a latent Markov transition process which fits not only the count of daily infected, but also the recovered and death counts.

The dynamics of these 7 compartments across time t were described by the following set of

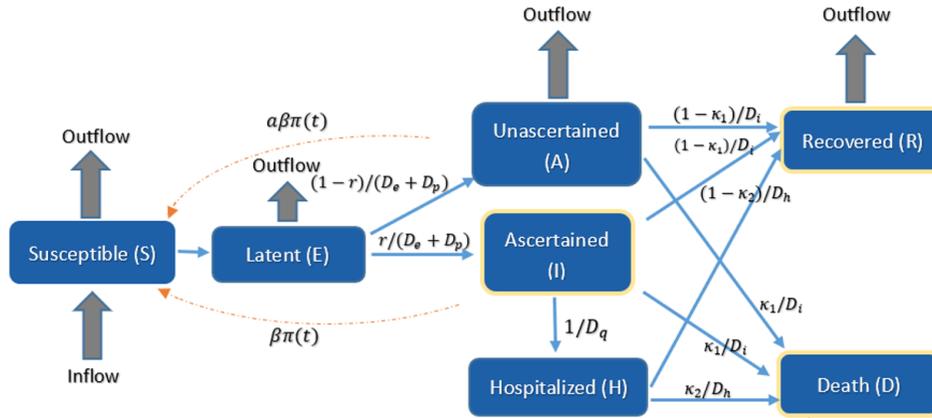


Figure 3: Schematic diagram of the proposed eSEIRD model

ordinary differential equations:

$$\begin{aligned} \frac{dS_t}{dt} &= n - \frac{\beta\pi(t)S_t(I_t + aA_t)}{N} - \frac{nS_t}{N}, \\ \frac{dE_t}{dt} &= \frac{\beta\pi(t)S_t(I_t + aA_t)}{N} - \frac{E_t}{D_e + D_p} - \frac{nE_t}{N}, \\ \frac{dA_t}{dt} &= \frac{(1-r)E_t}{D_e + D_p} - \left(\frac{(1-\kappa_1)}{D_i} + \frac{\kappa_1}{D_i} \right) A_t - \frac{nA_t}{N}, \\ \frac{dI_t}{dt} &= \frac{rE_t}{D_e + D_p} - \left(\left(\frac{(1-\kappa_1)}{D_i} + \frac{\kappa_1}{D_i} \right) + \frac{1}{D_q} \right) I_t, \\ \frac{dH_t}{dt} &= \frac{1}{D_q} I_t - \left(\frac{(1-\kappa_2)}{D_h} + \frac{\kappa_2}{D_h} \right) H_t, \\ \frac{dR_t}{dt} &= \left(\frac{(1-\kappa_1)}{D_i} I_t + \frac{(1-\kappa_1)}{D_i} A_t \right) + \frac{(1-\kappa_2)}{D_h} H_t - \frac{nR_t}{N}, \\ \frac{dD_t}{dt} &= \left(\frac{\kappa_1}{D_i} A_t + \frac{\kappa_1}{D_i} I_t \right) + \frac{\kappa_2}{D_h} H_t. \end{aligned}$$

Hierarchical model: We assumed three observed time series of daily counts of infected, recovered and death cases are emitted from Poisson state-space models, independent conditionally on the underlying process, and the Markov process associated with the latent proportions is constructed

as:

$$\begin{aligned}
 Y_t^I \mid \mathbf{X}_t, \boldsymbol{\tau}_3 &\sim \text{Poisson} \left(\lambda_t^I = \frac{rE_{t-1}}{D_e + D_p} \right), \\
 Y_t^R \mid \mathbf{X}_t, \boldsymbol{\tau}_3 &\sim \text{Poisson} \left(\lambda_t^R = \frac{(1 - \kappa_1)}{D_i} I_{t-1} + \frac{(1 - \kappa_2)}{D_h} H_{t-1} \right), \\
 Y_t^D \mid \mathbf{X}_t, \boldsymbol{\tau}_3 &\sim \text{Poisson} \left(\lambda_t^D = \frac{\kappa_1}{D_i} I_{t-1} + \frac{\kappa_2}{D_h} H_{t-1} \right), \\
 \mathbf{X}_t \mid \mathbf{X}_{t-1}, \boldsymbol{\tau}_3 &\sim \text{Poisson} \left(f_2(\mathbf{X}_{t-1}, \beta, r, \kappa_1, \kappa_2) \right),
 \end{aligned}$$

where $\mathbf{X}_t = (S_t, E_t, A_t, I_t, H_t, R_t, D_t)$ denotes the vector of the underlying population counts of the 7 compartments: $\boldsymbol{\tau}_3 = (\mathbf{X}_0^T, \alpha, \beta, r, D_e, D_p, D_q, D_i, D_h, \kappa_1, \kappa_2)$ denotes the whole set of parameters where \mathbf{X}_0^T denotes the prior for the initial states and κ_1 and κ_2 denote the IFR for non-hospitalized and hospitalized cases, respectively. The function $f_2(\cdot)$ is also solved using RK4 approximation as before (see the solution in Supplementary Section S.4).

$$f_2(\mathbf{X}_{t-1}, \beta, r, \kappa_1, \kappa_2) = \begin{pmatrix} X_{t-1}^S + 1/6 [k_{t-1}^{S1} + 2k_{t-1}^{S2} + 2k_{t-1}^{S3} + k_{t-1}^{S4}] \\ X_{t-1}^E + 1/6 [k_{t-1}^{E1} + 2k_{t-1}^{E2} + 2k_{t-1}^{E3} + k_{t-1}^{E4}] \\ X_{t-1}^A + 1/6 [k_{t-1}^{A1} + 2k_{t-1}^{A2} + 2k_{t-1}^{A3} + k_{t-1}^{A4}] \\ X_{t-1}^I + 1/6 [k_{t-1}^{I1} + 2k_{t-1}^{I2} + 2k_{t-1}^{I3} + k_{t-1}^{I4}] \\ X_{t-1}^H + 1/6 [k_{t-1}^{H1} + 2k_{t-1}^{H2} + 2k_{t-1}^{H3} + k_{t-1}^{H4}] \\ X_{t-1}^R + 1/6 [k_{t-1}^{R1} + 2k_{t-1}^{R2} + 2k_{t-1}^{R3} + k_{t-1}^{R4}] \\ X_{t-1}^D + 1/6 [k_{t-1}^{D1} + 2k_{t-1}^{D2} + 2k_{t-1}^{D3} + k_{t-1}^{D4}] \end{pmatrix} := \begin{pmatrix} \alpha_{1(t-1)} \\ \alpha_{2(t-1)} \\ \alpha_{3(t-1)} \\ \alpha_{4(t-1)} \\ \alpha_{5(t-1)} \\ \alpha_{6(t-1)} \\ \alpha_{7(t-1)} \end{pmatrix}$$

We implemented the MCMC algorithm to sample from the posterior distribution of the underlying parameters r and β , and calculate the derived quantities:

$$R_0 = \beta \left[\alpha(1 - r)(D_i) + \frac{r}{(1/D_i + 1/D_q)} \right].$$

We obtain the posterior mean estimates and credible intervals for the unknown parameters in the model. Because of the hierarchical structure in the state-space model considered in this model, the posterior sampling can be done in a straightforward fashion like eSIR using the R package `rjags`.

3 Modeling COVID-19 Transmission Dynamics in South Africa

3.1 Study design and data source

COVID-19 daily time series data for South Africa were extracted from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (Dong et al., 2020) from the onset of the first 50 confirmed case (March 15, 2020) to April 1, 2021. We

fitted the models using data up to December 31, 2020 and predicted the state of COVID-19 infection in South Africa in a time window, from January 1 to April 1, 2021. To compare the model prediction performance of different models, we used the symmetric mean absolute percentage error (SMAPE), given by:

$$SMAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{(|A_t| + |F_t|)/2},$$

where A_t is the observed value and F_t is the forecast value in the same time period. This design enabled us to select an optimal modeling strategy for South Africa data and check the robustness of prediction performance across different models.

3.1.1 Prior specification

We describe the prior choices and where appropriate, initial values for the model hyper-parameters in this section and a complete summary and list of notations and assumptions are available in Supplementary Table S.1.1. To begin with, we assumed a constant population size ($N = 57,779,622$) for all models and fixed a few transition parameters below in the SAPHIRE and eSEIRD model. First, we set an equal number of daily inbound and outbound travelers (n), in which $n = 4 \times 10^{-4} N$ from March 15 to 25, 2020 estimated by the number of international travelers to South Africa in 2018 (The World Bank, 2020), otherwise $n = 0$ when border closed, i.e. after March 26. We fixed the transmissibility ratio between unascertained and ascertained cases at $\alpha = 0.55$ assuming lower transmissibility for unascertained cases (Li et al., 2020b), an incubation period of 5.2 days, and a pre-symptomatic infectious period of $D_p = 2.3$ days (Li et al., 2020b; He et al., 2020b), implying a latent period of $D_e = 2.9$ days. The mean of total infectious period was $D_i + D_p = 5.2$ days (Li et al., 2020b), assuming constant infectiousness across the pre-symptomatic and symptomatic phases of ascertained cases (Li et al., 2020a), thus, the mean symptomatic infectious period was $D_i = 2.9$ days. We set the period of ascertained cases from reporting to hospitalization $D_q = 7$ days, the same as the median interval from symptom onset to admission reported (Ferretti et al., 2020; Garg et al., 2020). The period from being admitted in hospital to discharge or death was assumed as $D_h = 8.6$ days (Wang et al., 2020b). We fit the SAPHIRE and eSEIRD model in six time periods of 2020: March 15-March 26, March 27- April 30, May 1- May 31, June 1- August 17, August 18-September 20, and September 21-December 31, separated by the change-points of the lockdown strictness level, and denote the ascertained rate and transmission rate in the time periods as $r_1, r_2, r_3, r_4, r_5, r_6, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ and β_6 . In addition, we denote the IFR for non-hospitalized cases $\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{14}, \kappa_{15}, \kappa_{16}$ and for hospitalized cases $\kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{24}, \kappa_{25}, \kappa_{26}$ in eSEIRD model.

Choice of Initial states: For the eSIR model, the prior mean for the initial infected/removed proportion was set at the observed infected/removed proportion on March 15, 2020, and that for the susceptible proportion was the total number of the population minus the infected and removed proportions (Song et al., 2020). For the SAPHIRE model, other than setting prior parameters for initial states, we set the number of initial latent cases $E(0)$ was the sum of those ascertained and unascertained cases with onset during March 15-17, 2020 as $D_e = 2.9$ days (Hao et al., 2020) and the

number of initial pre-symptomatic cases $P(0)$ was that from March 18-19, 2020 as $D_p = 2.3$ days (Hao et al., 2020). The number of ascertained symptomatic cases $I(0)$ was assumed as the number of observed infected cases on March 15, 2020 excluding $H(0)$, $R(0)$ and $D(0)$ (the initial numbers for hospitalized, recovered, and deaths). The initial ascertainment rate (r_0) was assumed as 0.10 as reported in literature (Bhattacharyya et al., 2021; Lau et al., 2020), implying $A(0) = 0.90/0.10I(0)$, and a sensitivity analysis with $r_0 = 0.25$ was conducted to address weak information for r_0 obtained in South Africa and variation of r_0 in different scenarios. $H(0)$ was assumed as 50% of the observed ascertained cases on March 9, 2020 (by assuming the period from reported to hospitalized was 7 days (Ferretti et al., 2020; Garg et al., 2020) at the early stage of the pandemic). In addition, we denoted $R(0)$ as the sum of observed recovered and death cases on March 15. The number of initial susceptible cases $S(0)$ was calculated as the total population (N) minus $E(0)$, $P(0)$, $I(0)$, $A(0)$ and $R(0)$.

In the eSEIRD model, we set the prior mean of initial ascertained, unascertained and hospitalized cases as $I(0)$, $A(0)$ and $H(0)$ discussed above. However, since the latent compartment incorporates the pre-symptomatic cases, the mean of the initial latent cases was set as the sum of those ascertained and unascertained cases with onset during March 15-19, 2020 as $D_e + D_p = 5.2$ days (Hao et al., 2020). The prior mean of initial recoveries and deaths were fixed as the number of observed recovered and death cases on March 15, 2020, respectively. Therefore, the prior mean of initial susceptible compartment was set as the total population excluding the mean of other compartments.

Prior distributions: In the eSIR model, the log-normal priors were used for the removed rate ν and the basic reproduction number R_0 , in particular $\nu \sim \text{LogN}(2.955, 0.910)$, with $E(\nu) = 0.082$ and $SD(\nu) = 0.1$ (Song et al., 2020), and $R_0 (= \beta/\nu) \sim \text{LogN}(0.582, 0.223)$ with $E(R_0) = 3.2$ and $SD(R_0) = 1$ (Ray et al., 2020). Flat Gamma priors were used for the scale parameters of the Beta-Dirichlet distributions as follows (Ray et al., 2020):

$$\omega \sim \text{Gamma}(2, 0.0001), \lambda^I \sim \text{Gamma}(2, 0.0001), \text{ and } \lambda^R \sim \text{Gamma}(2, 0.0001).$$

In the eSEIRD model, for the six time periods, all the transmission rates $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ and β_6 were given a $\mathcal{U}(0, 2)$ prior, ascertained rates r_1, r_2, r_3, r_4, r_5 and r_6 were given Beta(10, 90) prior (Lau et al., 2020; Rahmandad et al., 2020), the IFR for non-hospitalized cases $\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{14}, \kappa_{15}, \kappa_{16} \sim \text{Beta}(0.03, 2.93)$ and for hospitalized cases $\kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{24}, \kappa_{25}, \kappa_{26} \sim \text{Beta}(0.44, 1.76)$ with mean equal to 0.1% and 20% , respectively (U.S. CDC, 2020; Jassat et al., 2020).

In addition, to account for the effect of time-varying contact rate during the prediction period, we set a time-varying contact rate modifier $\pi(t)$ in the eSIR and eSEIRD model where t from January 1 to April 1, 2021: $\pi(t)$ was set as 0.75 since the lockdown was tuned to level 3 after December 28, 2020. Note that the modifier $\pi(t)$ is a conjectural quantity and hence must be guided by empirical studies (Ray et al., 2020). Using MCMC sampling method for the eSIR and eSEIRD model, we set the adaptation number to be 10^4 , thinned by 10 draws to reduce auto-correlation, and set a burn-in period of 5×10^4 draws under 10^5 iterations for 4 parallel chains.

We fit the SAPHIRE model in six time periods as in the eSEIRD model. We used $r_1 \sim$

$Beta(10, 90)$ and reparameterized $r_2, r_3, r_4, r_5,$ and r_6 by

$$\text{logit}(r_i) = \text{logit}(r_{i-1}) + \delta_{i-1}, i = 2, \dots, 6,$$

where $\text{logit}(r) = \log(r/(1-r))$, and $\delta_i \sim \mathcal{N}(0, 1)$, for $i = 1, \dots, 5$. We use non-informative prior distributions for transmission rates $\beta_i \sim \mathcal{U}(0, 2)$ for $i = 1, \dots, 6$ to reflect lack of information about these hyperparameters (Hao et al., 2020). Therefore, β and r were assumed to follow different distributions for these six time periods. Finally, the effective reproduction number can be derived to be:

$$R_e = \beta \left[\alpha D_p + \alpha(1-r)D_i + \frac{r}{(1/D_i + 1/D_q)} \right].$$

We set a burn-in period of 10^5 iterations and continued to run 10^5 iterations with a sampling step size of 10 iterations for the MCMC algorithm of the SAPHIRE model.

A comparison between assumptions of the three models in the Supplementary Section S.2. All analyses were conducted in R (version 4.0.0), and source codes are available at https://github.com/umich-cphds/south_africa_modeling. Posterior mean and corresponding 95% credible interval (95% CrI) were reported for the parameters of interests.

4 Results

Here we present the detailed results for South Africa, subdivided into estimation of key epidemiological parameters, short-term and long-term forecasts, and finally model evaluation in terms prediction and quality of fit.

4.1 Reproduction number and intervention evaluation

The estimated posterior mean of R_0 was 1.18 (95%CrI: [1.09, 1.28]) in the eSIR model throughout the training period while in the eSEIRD model, the value of R_0 started at 3.22 (95%CrI: [3.19, 3.23]) then dropped though still significantly above 1 after the lockdown implementation and increased to 3.27 (95%CrI: [3.27, 3.27]) during the last three months of 2020 (Table 2). It suggests that the effective contact rate decreased by more than 50% over the lockdown time period and attained the lowest point in its trajectory during August to September, 2020 though the lockdown was eased to a relatively less strict level. On the other hand, the effective reproduction number (R_e) in different lockdown periods estimated by the SAPHIRE model demonstrates that a similar trend but the magnitude of the estimated R_e decreased dramatically when r_0 increases from 0.10 to 0.25 (Table 2, Fig. 4), possibly suggesting lack of robustness with respect to the choice of initial r_0 .

4.2 Short-term and long-term forecasts

We first forecast the total cumulative number of infections, including unascertained cases, in the SAPHIRE model up to February 28, 2021 depending on the time-period considered for estimating the trend, and then extend it to April 1, 2021. On February 28, 2021, the estimated cumulative

number of infections are: (a) 43.3 million if the trend of the least strict lockdown (level 1) was assumed, (b) 36.9-37.8 million if the trend of the lockdown level 2-4 was assumed, and (c) 29.3 million if the trend of strict lockdown was assumed ($r_0 = 0.10$). However, the short-term forecasts in SAPHIRE model exhibits lack of robustness under different r_0 settings, for example, when $r_0 = 0.25$ the estimated cumulative number of infections was 10.6 million if the trend of the least strict lockdown (level 1) was assumed and 0.9 million if the trend of the lockdown level 2 was assumed. In the eSEIRD model, the predicted total cumulative number of cases reach 41.2 million under $r_0 = 0.10$, and 41.6 million under $r_0 = 0.25$, along with estimated total deaths counts as 35 or 37 thousand when $r_0 = 0.10$ or 0.25, respectively, by February 28, 2021. Furthermore, we used the eSEIRD model to forecast the epidemic trajectory for a relatively longer time period, where we found that by April 1, the cumulative number of total infected (including the unascertained cases) and deaths would reach roughly 41.3 million (which is around 70% of the total population in South Africa) and 35 thousand, respectively. For comparison, the observed number of cases and deaths on April 1 were 1,549,451 and 52,897, respectively. It is also worth noting that the SIR model used in Ding et al. (2021) predicted the pandemic to cease transmission by April 28, 2021 with an estimated fatality count of 17,072.

4.3 Fitting and prediction performance

All the three candidate models were applied to the COVID-19 data in South Africa with high accuracy as the estimated daily new cases were close to the observed numbers from March to October, 2020 (Fig. 4 (a)-(c)). However, the eSEIRD model showed a poorer fit during the second pandemic wave in South Africa from November to December 2020, compared to the other two models. The eSIR model performed best in terms of fitting the cumulative ascertained cases with the smallest SMAPE (2.43% when $r_0=0.10$) while the SAPHIRE model had the second smallest training SMAPE (Table 4). In terms of predictive accuracy, the SAPHIRE model performed best with the smallest SMAPE (4.41% for 15 days, 5.92% for 31 days and 6.94% for 91 days when $r_0=0.10$) while the eSIR model had the second smallest SMAPE (6.90% for 15 days, 10.78% for 31 days and 8.40% for 91 days when $r_0=0.10$) (Table 4). We note that for a few selected important time points, the predicted number of cases in the eSIR and SAPHIRE model are closer to the observed (Table 3). The predictive performances for the three competing models substantiate their credibility in terms of capturing the transmission dynamics for the time-period considered in this study.

4.4 Unascertained cases and deaths

As demonstrated by SAPHIRE modeling results in Figure 4 (d), the large number of unascertained and pre-symptomatic cases contributed to the rapid spread of disease. The estimated ascertained rates were very low, starting at 8.99% (95% CrI: [8.20%, 9.80%]), decreasing to below 2 during level 5 to level 3 lockdown and then increasing to 15.48% (95% CrI: [15.24%, 15.73%]) during the second pandemic wave in South Africa, respectively (Table 2, Fig. 4 (f)). Similarly, in the eSEIRD model, the estimated ascertained rates were also at a very low level (1.65% to 9.17%) and had a similar trend as in SAPHIRE model (see Table 2). As mentioned before, the estimated ascertained rates

were robust with respect to choices for r_0 in eSEIRD model, but changes drastically in SAPHIRE model with r_0 changed to 0.25: 95.58% (95%CrI: [76.20%, 99.50%]) before lockdown, between 30.18% to 35.93% during level 5 and 4 lockdown, 95.53% (95%CrI: [81.09%, 99.58%]) in level 3 lockdown, and then decreasing from 33.23% to 18.80% in level 2 to level 1 lockdown.

By the eSEIRD model, the overall IFR was estimated as 0.06% (95%CrI: [0.04%, 0.22%]) throughout the whole time period taking the reported and unreported cases and deaths into account while the observed overall case fatality ratio was estimated as 2.88% (95% CrI: [2.45%, 6.01%]) (Fig. 5). Furthermore, the eSEIRD model provided Bayesian estimates for IFR and deaths among hospitalized and non-hospitalized cases. The estimated IFR for hospitalized cases was 15.28% (95%CrI: [0.01%, 69.10%]) before lockdown and increased to 65.86% (95%CrI: [51.00%, 82.91%]) in the first time period of lockdown. After that, the IFR for hospitalized cases decreased from 22.9% (95%CrI: [20.75%, 25.18%]) to 7.46% (95%CrI: [7.22%, 7.71%]) during May to September. By the end of 2020, it again increased to 19.25% (95%CrI: [18.82%, 19.69%]). The IFR of hospitalized cases was much larger than that of non-hospitalized cases (less than 0.01%), and these estimates were robust to the choice of r_0 .

5 Discussion

In this paper, we propose a new infectious disease forecasting model that incorporates the unascertained cases, population movement over different time periods, and the effect of intervention strategies in a unified way and use it to investigate the spread of COVID-19 in South Africa, the hardest hit country on the African continent. The methodological tools developed here can be used to estimate the IFR as well as estimate actual COVID-19 deaths from the reported death counts.

The lockdown intervention and mandatory face-mask wearing in public places employed in South Africa seemed to contain the spread of COVID-19 effectively as the R_e decreased dramatically initially but increased later following the relaxation of lockdown stringency afterwards. However, the R_e was consistently above 1 throughout the whole period analyzed, which implies the interventions failed to dampen the transmission fully, further substantiated by the basic reproduction number estimates in the eSEIRD model as well. This agrees with the public health experts advice of carefully implemented intervention policies while taking account their potential economic costs (Stiegler and Bouchard, 2020).

We also found that the estimated ascertainment rate is very low in South Africa compared to that reported for many other countries (Hao et al., 2020; Bhattacharyya et al., 2021; Rahmandad et al., 2020), also implied by the low testing rate and high testing positive rate in South Africa. As of February 21, the number of total tests conducted is 8.9 million, suggesting that about 15.4% population were tested (Ritchie et al., 2020). Furthermore, the estimated ascertainment rate is consistent with that in other multiple global epicenters under severe pandemic of COVID-19, such as France, the United States, Italy and Spain in March (Lau et al., 2020). The large number of unascertained cases is likely to contribute significantly to the continuing spread of COVID-19 Li et al. (2020b); Cereda et al. (2020); Ioannidis (2021). Our findings suggest that there are around 70% of the total population in South Africa infected by December 31, 2020, which is roughly consistent with the

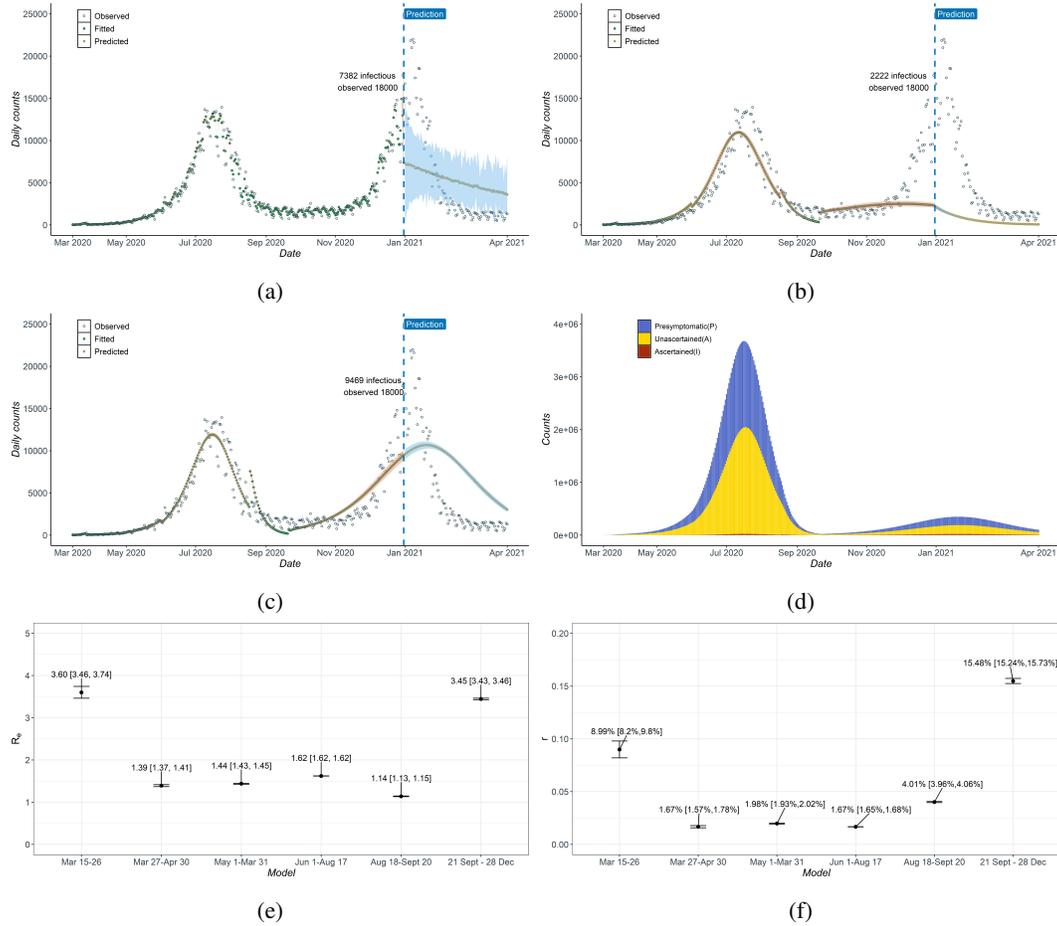


Figure 4: (a)-(c) Daily new number of ascertained infections cases estimated by the models compared with observed data: (a) eSIR, (b) SAPHIRE, and (c) eSEIRD; (d) Current pre-symptomatic/unascertained/ascertained infectious in the SAPHIRE model; (e)-(f) Estimated effective reproduction number (R_e) and ascertained rate (r) in the SAPHIRE model in four time periods. (Assume initial ascertained rate (r_0) equal to 0.10.)

seroprevalence survey conducted in South Africa that the estimated prevalence is around 63% in the Eastern Cape, one of the pandemic centers in South Africa (Sykes et al., 2021). Despite the potential high prevalence of COVID, the second wave of pandemic appeared in South Africa and other pandemic centers like Brazil, which may due to the waning immunity against infection with the time to the first wave increasing, and the coronavirus lineages might have higher inherent transmissibility than the pre-existing lineages and be able to evade the immunity generated in response to previous infection (Sabino et al., 2021). To prevent potential resurgence in the future, addition to the strict

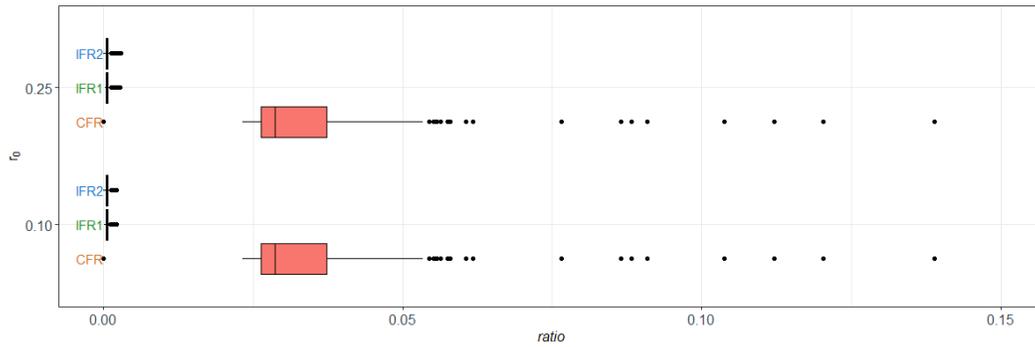


Figure 5: Case fatality ratio (CFR) and estimated infection fatality ratio (IFR) in the eSEIRD model. $CFR = (\text{Number of reported deaths}) / (\text{Number of reported deaths and recovered})$; $IFR1 = (\text{Number of reported deaths}) / (\text{Number of reported and unreported cases})$; and $IFR2 = (\text{Number of reported and unreported deaths}) / (\text{Number of reported and unreported cases})$ [20].

interventions, more surveillance testing and effective testing strategies under conditions of limited test availability, such as contact tracing of the contacts and confirmed cases, will be helpful to curtail the pandemic in South Africa (Garba et al., 2020).

Although highly transmissible and poorly ascertained, the COVID-19 IFR is estimated as 0.06% taking account of unreported cases and deaths in South Africa, comparable to the estimates in other locations with similar low mortality rate based on serological data (Ioannidis, 2021).

The low IFR may be due to the South African population being relatively young which lessens the fatal impact on general population to some extent (Statista, 2021). Our estimates of the IFR of hospitalized cases are much higher than that for non-hospitalized cases, suggesting that the most severe cases may have been admitted to hospitals despite the relatively lack of the testing arrangements. The very low estimated IFR for the non-hospitalized cases also imply that the degree of under-reporting for death by the model is very low (0.24% by April 1, 2021), and likely to be affected by the same factors.

Comparison of the models: The eSIR and the SAPHIRE model have been successfully applied to the data in India and Wuhan, China, separately (Song et al., 2020; Hao et al., 2020). Although SAPHIRE model exhibits superior prediction performance on COVID-19 cases, the estimates of underlying parameters and unascertained cases showed lack of robustness to the change of initial ascertainment rate r_0 . On the other hand, the eSIR model has the best estimation capability in terms of the ascertained cases but a relatively poor predictive capacity for capturing the change in the trend of the epidemic in time for neglecting some important clinical characteristics. The eSEIRD model also has a good fitting performance but a relatively poor prediction capacity. Table 2 also suggests that the estimates in eSEIRD model are robust estimated compared to the SAPHIRE model, probably an artifact of the Bayesian hierarchical model used.

Table 2: The posterior mean and credible intervals of the basic/effective reproduction number (R_0 or R_e) and ascertained rate (r) obtained from different models and settings.

Model	r_0	R_0 or R_e		$r(\%)$		
		Mean	95% CrI	Mean	95% CrI	
eSIR	0.1	1.18	[1.09,1.28]	-	-	
	0.25	1.18	[1.09,1.28]	-	-	
SAPHIRE	Mar 15-26	3.6	[3.46,3.74]	8.99	[8.20,9.80]	
		Mar 27-Apr 30	1.39	[1.37,1.41]	1.67	[1.57,1.78]
	0.1	May 1-31	1.44	[1.43,1.45]	1.98	[1.93,2.02]
		Jun 1-Aug 17	1.62	[1.62,1.62]	1.67	[1.65,1.68]
	Aug 18- Sept 20	1.14	[1.13,1.15]	4.01	[3.96,4.06]	
	Sept 21-Dec 31	3.45	[3.43,3.46]	15.48	[15.24,15.73]	
	0.25	Mar 15-26	1.44	[1.28,1.72]	95.58	[76.20,99.50]
		Mar 27-Apr 30	1.4	[1.35,1.47]	30.18	[23.92,33.17]
	May 1-31	1.39	[1.37,1.41]	35.93	[28.99,38.41]	
	Jun 1-Aug 17	1.07	[1.07,1.07]	94.53	[81.09,99.58]	
	Aug 18- Sept 20	0.8	[0.79,0.80]	33.23	[28.76,34.99]	
	Sept 21-Dec 31	1.25	[1.24,1.25]	18.8	[16.47,20.01]	
	eSEIRD	Mar 15-26	3.22	[3.19,3.23]	4.7	[4.27,5.17]
			Mar 27-Apr 30	1.48	[1.46,1.51]	1.85
0.1		May 1-31	1.52	[1.51,1.54]	2.02	[1.99,2.06]
		Jun 1-Aug 17	1.68	[1.67,1.68]	1.65	[1.64,1.66]
Aug 18- Sept 20		1.38	[1.36,1.39]	2.23	[2.21,2.25]	
Sept 21-Dec 31		3.27	[3.27,3.27]	9.17	[9.01,9.34]	
0.25		Mar 15-26	3.25	[3.23,3.26]	7.2	[6.50,7.95]
		Mar 27-Apr 30	1.57	[1.54,1.59]	2.17	[2.05,2.28]
May 1-31		1.54	[1.53,1.56]	2.07	[2.03,2.11]	
Jun 1-Aug 17		1.68	[1.67,1.68]	1.65	[1.64,1.65]	
Aug 18- Sept 20	1.38	[1.36,1.39]	2.23	[2.20,2.25]		
Sept 21-Dec 31	3.27	[3.27,3.27]	9.14	[8.95,9.31]		

Recent Developments: Anticipating increased inter-provincial holiday travel during December 2020, and Easter 2021, the South Africa government imposed ad-hoc lockdown measures in potential hotspot areas around the country, which did contribute positively to the COVID-19 prevention measures as is evident from the trajectory on Fig. 4. Several reports (e.g. National COVID-19 Modelling Consortium SACEMA, 2021; South African Medical Research Council, 2021) predict

Table 3: Comparison of the models regarding the cumulative ascertained infected and death with the observed (in thousands). Bold-faced entries indicate column winners regarding the closeness to the observed.

Model	r_0	Infected				Death			
		Estimation		Prediction		Estimation		Prediction	
		Dec 31 2020	Jan 31 2021	Feb 28 2021	Apr 1 2021	Dec 31 2020	Jan 31 2021	Feb 28 2021	Apr 1 2021
eSIR	0.1	1052	1256	1403	1549	-	-	-	-
	0.25	1052	1256	1402	1535	-	-	-	-
SAPHIRE	0.1	1058	1379	1624	1776	-	-	-	-
	0.25	1055	1576	2320	3243	-	-	-	-
eSEIRD	0.1	878	917	928	932	31	34	35	35
	0.25	924	966	977	980	33	36	37	37
Observed	-	1052	1454	1513	1549	28	44	50	54

Table 4: Symmetric mean absolute percentage error (SMAPE) of short-term forecasting in 2021. Bold-faced entries indicate column winners regarding prediction performance.

Model	r_0	Cumulative ascertained cases				Cumulative ascertained deaths			
		Training	Testing		Training	Testing			
			Jan 1 15	-Jan 31		Jan 1 15	-Jan 31		
eSIR	0.1	2.43%	6.90%	10.78%	-	-	-	-	
	0.25	2.41%	6.90%	10.78%	-	-	-	-	
SAPHIRE	0.1	3.17%	4.41%	5.92%	-	-	-	-	
	0.25	11.13%	2.40%	2.74%	-	-	-	-	
eSEIRD	0.1	13.17%	28.37%	35.50%	66.80%	5.07%	12.89%		
	0.25	6.03%	23.31%	30.49%	60.30%	5.97%	9.88%		

the third wave of COVID-19 to hit South Africa in the austral-winter of 2021.

Around October-November 2020, the B.1.351 or 501.V2 variant, first identified in South Africa, was considered to be highly transmissible (Tang et al., 2021; Tegally et al., 2021). At the time of writing, international travel is still restricted, and hence also contributing to the containment of spread outside South Africa borders.

The AstraZeneca vaccine rollout was halted by the South African Health Products Regulatory Authority (SAHPRA) as it was examined to be ineffective for the variant prevalent in South Africa,

which delayed the vaccination plan. Thereafter, the Sisonke ('together' in Zulu) trial was underway (South African Medical Research Council, 2021), which is a Phase 3 ENSEMBLE trial managed by the SAMRC aimed to '... make a vaccine available while the licensing process takes place'. At the time of writing, the J&J single-dose vaccine is being administered to healthcare workers. It is estimated that the total cost of immunizing two-thirds of South Africa's population to achieve herd-immunity could be around ZAR22 billion, excluding logistical cost of deployment. Another important logistical constraint that remains in South Africa, albeit for all of sub-Saharan Africa, is the requirement of low storage temperatures for mRNA vaccines (Fontanet et al., 2021) thus restricting the options from the variety of vaccines available currently.

Strengths and Limitations: The key methodological innovation for the proposed method is revealed by a quick comparison between the schematic diagrams for eSEIRD model (Fig. 3) and SAPHIRE model (Fig. 2(b)). Broadly speaking, eSEIRD incorporates $\pi(t)$, the transmission rate modifier as well as splits the 'removed' compartment into 'recovered' and 'deaths' while accounting for separate rates for ascertained, unascertained and hospitalized cases.

Despite the superior performance and robustness exhibited by the models examined here, there are some important limitations. First, the model assumptions were elicited from previous reports from other countries because of the dearth of such information for South Africa, especially for the fixed values for hyper-parameters. Though the estimation of parameters and prediction of infections seem to be robust to these assumptions to some extent, the inference and prediction would be much more convincing when based on accurate information specific to South Africa.

Second, the ascertained rate was assumed to follow the same distribution in a long time period in the eSEIRD model although in reality it might be time-varying depending on the accumulating knowledge and deployment of clinical resources for COVID-19, given the spatial variation within South Africa regarding the population density and movement, as well as regarding location of COVID-19 hotspots and hospital resources. Further, the population density is highly heterogeneous in different regions in South Africa with higher concentration near high-density economic-hub cities, such as Johannesburg, Cape Town and Durban. COVID-19 cases are also diversely spread. For instance, Gauteng Province is spatially very small, and is a highly dense province, with roughly 30% of total cases in the nation, while about 49% of confirmed cases cluster in KwaZulu-Natal, Eastern Cape and Western Cape Provinces. In addition, the seroprevalence study also suggested that the prevalence may vary from city to city: 63% in the Eastern Cape, 52% in the KwaZulu-Natal and 32% in the Northern Cape (Sykes et al., 2021). Without considering these heterogeneities and potential confounding factors in individual regions, the conclusion on the national data might be biased. The burden of HIV and tuberculosis comorbidity, particularly among the less privileged socio-economic population, also adds to the complexity of analyzing the COVID-19 data from South Africa (Boulle et al., 2020).

Third, in this paper we implicitly assumed that the recovered cases would not be infected again, but this assumption is still inconclusive based on extant research for COVID-19 (Gousseff et al., 2020). It might lead to a resurgence if this assumption is not valid and the interventions are totally lifted. Thus, it might be necessary to conduct more national serological surveys on COVID-

19 among the general population in South Africa to confirm the national, as well as provincial, seroprevalence. Such large-scale studies will also provide more powerful evidence to examine the evolving benefits of non-pharmaceutical intervention decisions and provide guidance to manage provincial level disparity.

Finally, from the early stage of this pandemic to now, there has been an explosive development in COVID-19 forecasting models, but systematic comparison between the available models in terms of out-of-sample prediction and inference has been rare (see e.g. Friedman et al., 2020), as are carefully executed simulation studies where the ‘ground truth’ is known. Lack of simulation studies comparing the candidate methods is also a limiting feature of this paper, and we hope to pursue this in a future endeavor.

Acknowledgements

This work was supported by grants from the National Science Foundation [grant numbers DMS-1712933 (to B.M.) and DMS-2015460 (to J.D.)] and from National Institute of Health – 1 R01 HG008773-01 (to B.M.). The first and second author (X.G. and B.M.) would also like to thank the Center for Precision Health Data Sciences at the University of Michigan School of Public Health, The University of Michigan Rogel Cancer Center and the Michigan Institute of Data Science for internal funding that supported this research.

References

- Arndt, C., Davies, R., Gabriel, S., Harris, L., Makrelov, K., Robinson, S., Levy, S., Simbanegavi, W., van Seventer, D., and Anderson, L. (2020), “Covid-19 lockdowns, income distribution, and food security: An analysis for South Africa,” *Global Food Security*, 26, 100410.
- Bhattacharyya, R., Kundu, R., Bhaduri, R., Ray, D., Beesley, L., Salvatore, M., and Mukherjee, B. (2021), “Incorporating false negative tests in epidemiological models for SARS-CoV-2 transmission and reconciling with seroprevalence estimates,” *Scientific Reports*, 11(1), 9748.
- Boulle, A., Davies, M.-A., Hussey, H., Ismail, M., Morden, E., Vundle, Z., Zweigenthal, V., Mahomed, H., Paleker, M., Pienaar, D., et al. (2020), “Risk factors for COVID-19 death in a population cohort study from the Western Cape Province, South Africa,” *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*.
- Cereda, D., Tirani, M., Rovida, F., Demicheli, V., Ajelli, M., Poletti, P., Trentini, F., Guzzetta, G., Marziano, V., Barone, A., et al. (2020), “The early phase of the COVID-19 outbreak in Lombardy, Italy,” *arXiv preprint arXiv:2003.09320*.
- Daniel, L. (2020), “New booze, curfew rules for Covid hotspots: Govt to take decision today,” <https://bit.ly/34fxDFK>, accessed 2020-12-01.

- Ding, W., Wang, Q.-G., and Zhang, J.-X. (2021), “Analysis and prediction of COVID-19 epidemic in South Africa,” *ISA Transactions*, S0019–0578(21)00060–4.
- Dong, E., Du, H., and Gardner, L. (2020), “An interactive web-based dashboard to track COVID-19 in real time,” *The Lancet infectious diseases*, 20, 533–534.
- Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., and Fraser, C. (2020), “Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing,” *Science*, 368, eabb6936.
- Fontanet, A., Autran, B., Lina, B., Kieny, M. P., Karim, S. S. A., and Sridhar, D. (2021), “SARS-CoV-2 variants and ending the COVID-19 pandemic,” *The Lancet*, 397, 952–954.
- Friedman, J., Liu, P., Gakidou, E., and Team, I. C.-. M. C. (2020), “Predictive performance of international COVID-19 mortality forecasting models,” *medRxiv*, 2020.07.13.20151233.
- Garba, S. M., Lubuma, J. M.-S., and Tsanou, B. (2020), “Modeling the transmission dynamics of the COVID-19 Pandemic in South Africa,” *Mathematical Biosciences*, 328, 108441.
- Garg, S., Kim, L., Whitaker, M., O’Halloran, A., Cummings, C., Holstein, R., Prill, M., Chai, S. J., Kirley, P. D., Alden, N. B., et al. (2020), “Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019—COVID-NET, 14 States, March 1–30, 2020,” *Morbidity and mortality weekly report*, 69, 458.
- Gousseff, M., Penot, P., Gallay, L., Batisse, D., Benech, N., Bouiller, K., Collarino, R., Conrad, A., Slama, D., Joseph, C., et al. (2020), “Clinical recurrences of COVID-19 symptoms after recovery: viral relapse, reinfection or inflammatory rebound?” *Journal of Infection*, 81, 816–846.
- Hao, X., Cheng, S., Wu, D., Wu, T., Lin, X., and Wang, C. (2020), “Reconstruction of the full transmission dynamics of COVID-19 in Wuhan,” *Nature*, 584, 420–424.
- He, J., Guo, Y., Mao, R., and Zhang, J. (2020a), “Proportion of asymptomatic coronavirus disease 2019: A systematic review and meta-analysis,” *J Med Virol*, 93, 820–830.
- He, X., Lau, E. H. Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X., Mo, X., Chen, Y., Liao, B., Chen, W., Hu, F., Zhang, Q., Zhong, M., Wu, Y., Zhao, L., Zhang, F., Cowling, B. J., Li, F., and Leung, G. M. (2020b), “Temporal dynamics in viral shedding and transmissibility of COVID-19,” *Nature Medicine*, 26, 672–675.
- Ioannidis, J. P. (2021), “Infection fatality rate of COVID-19 inferred from seroprevalence data,” *Bulletin of the World Health Organization*, 99, 19.
- Jassat, W., Cohen, C., Masha, M., Goldstein, S., Kufa, T., Savulescu, D., Walaza, S., Bam, J.-L., Davies, M.-A., Prozesky, H. W., Naude, J., Mnguni, A. T., Lawrence, C. A., Mathema, H. T., and Blumberg, L. (2020), “COVID-19 in-hospital mortality in South Africa: the intersection of communicable and non-communicable chronic diseases in a high HIV prevalence setting,” *medRxiv*, 2020.12.21.20248409.

- Kermack, W. O., McKendrick, A. G., and Walker, G. T. (1927), “A contribution to the mathematical theory of epidemics,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115, 700–721.
- Lau, H., Khosrawipour, T., Kocbach, P., Ichii, H., Bania, J., and Khosrawipour, V. (2020), “Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters,” *Pulmonology*, iSBN: 2531-0437 Publisher: Elsevier.
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S., Lau, E. H., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., Tu, W., Chen, C., Jin, L., Yang, R., Wang, Q., Zhou, S., Wang, R., Liu, H., Luo, Y., Liu, Y., Shao, G., Li, H., Tao, Z., Yang, Y., Deng, Z., Liu, B., Ma, Z., Zhang, Y., Shi, G., Lam, T. T., Wu, J. T., Gao, G. F., Cowling, B. J., Yang, B., Leung, G. M., and Feng, Z. (2020a), “Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia,” *N Engl J Med*, 382, 1199–1207.
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020b), “Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2),” *Science*, 368, 489–493.
- Mukandavire, Z., Nyabadza, F., Malunguza, N. J., Cuadros, D. F., Shiri, T., and Musuka, G. (2020), “Quantifying early COVID-19 outbreak transmission in South Africa and exploring vaccine efficacy scenarios,” *PLOS ONE*, 15, e0236003.
- National COVID-19 Modelling Consortium SACEMA (2021), “COVID-19 modelling update: Considerations for a potential third wave,” http://www.sacema.org/node/National_COVID-19_Modelling_Consortium.
- Osthus, D., Hickmann, K. S., Caragea, P. C., Higdon, D., and Valle, S. Y. D. (2017), “Forecasting seasonal influenza with a state-space SIR model,” *Ann. Appl. Stat.*, 11, 202–224.
- Rahmandad, H., Lim, T. Y., and Sterman, J. (2020), “Estimating COVID-19 under-reporting across 86 nations: implications for projections and control,” *medRxiv*.
- Ray, D., Salvatore, M., Bhattacharyya, R., Wang, L., Du, J., Mohammed, S., Purkayastha, S., Halder, A., Rix, A., Barker, D., Kleinsasser, M., Zhou, Y., Bose, D., Song, P., Banerjee, M., Baladandayuthapani, V., Ghosh, P., and Mukherjee, B. (2020), “Predictions, Role of Interventions and Effects of a Historic National Lockdown in India’s Response to the the COVID-19 Pandemic: Data Science Call to Arms,” *Harvard Data Science Review*.
- Ritchie, H., Ortiz-Ospina, E., Beltekian, D., Mathieu, E., Hasell, J., Macdonald, B., Giattino, C., Appel, C., Rodés-Guirao, L., and Roser, M. (2020), “Coronavirus Pandemic (COVID-19),” <https://ourworldindata.org/coronavirus>.
- Sabino, E. C., Buss, L. F., Carvalho, M. P., Prete, C. A., Crispim, M. A., Fraiji, N. A., Pereira, R. H., Parag, K. V., da Silva Peixoto, P., Kraemer, M. U., et al. (2021), “Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence,” *The Lancet*, 397, 452–455.

- Song, P. X., Wang, L., Zhou, Y., He, J., Zhu, B., Wang, F., Tang, L., and Eisenberg, M. (2020), “An epidemiological forecast model and software assessing interventions on COVID-19 epidemic in China,” *MedRxiv*.
- South Africa Department of Health (2021), “COVID-19 Online Resource and News Portal,” <https://sacoronavirus.co.za>, accessed: February 14, 2021.
- South African COVID-19 Modelling Consortium (2020), “Estimating cases for COVID-19 in South Africa Long-term national projections,” <https://bit.ly/2T7qQLM>, accessed: February 21, 2021.
- South African Medical Research Council (2021), “Sisonke - Protecting Healthcare Workers,” [http://sisonkestudy.samrc.ac.za/#\[object%20object\]](http://sisonkestudy.samrc.ac.za/#[object%20object]).
- Statista (2021), “South Africa: Average age of the population from 1950 to 2050,” <https://www.statista.com/statistics/578976/average-age-of-the-population-in-south-africa/>.
- Stiegler, N. and Bouchard, J.-P. (2020), “South Africa: Challenges and successes of the COVID-19 lockdown.” *Ann Med Psychol (Paris)*, 178, 695–698.
- Sykes, W., Mhlanga, L., Swanevelder, R., Glatt, T. N., Grebe, E., Coleman, C., Pieterse, N., Cable, R., Welte, A., van den Berg, K., et al. (2021), “Prevalence of anti-SARS-CoV-2 antibodies among blood donors in Northern Cape, KwaZulu-Natal, Eastern Cape, and Free State provinces of South Africa in January 2021.” ””.
- Tang, J. W., Toovey, O. T., Harvey, K. N., and Hui, D. D. (2021), “Introduction of the South African SARS-CoV-2 variant 501Y. V2 into the UK,” *The Journal of Infection*, 82, e8–e10.
- Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay, S., San, E. J., Msomi, N., et al. (2021), “Detection of a SARS-CoV-2 variant of concern in South Africa,” *Nature*, 592, 438–443.
- The World Bank (2020), “International tourism, number of arrivals - South Africa,” <https://data.worldbank.org/indicator/ST.INT.ARVL?locations=ZA>.
- U.S. CDC (2020), “Provisional Death Counts for Coronavirus Disease 2019 (COVID-19),” <https://www.cdc.gov/nchs/nvss/vsrr/covidweekly/index.htm>.
- Wang, C., Liu, L., Hao, X., Guo, H., Wang, Q., Huang, J., He, N., Yu, H., Lin, X., and Pan, A. (2020a), “Evolving epidemiology and impact of non-pharmaceutical interventions on the outbreak of Coronavirus disease 2019 in Wuhan, China,” *MedRxiv*.
- Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., and Xiong, Y. (2020b), “Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China,” *JAMA*, 323, 1061–1069.

- World Health Organization (2020a), “COVID-19 virtual press conference on 30 March 2020,” https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-30mar2020.pdf?sfvrsn=6b68bc4a_2, accessed February 21, 2021.
- (2020b), “Estimating mortality from COVID-19: scientific brief, 4 August 2020,” <https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci-Brief-Mortality-2020.1>.
- (2020c), “Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19),” <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>, accessed May 5, 2021.
- (2021a), “South Africa: WHO Coronavirus disease (COVID-19) Dashboard,” <https://covid19.who.int/region/afro/country/za>, accessed February 21, 2021.
- (2021b), “WHO Coronavirus disease (COVID-19) Dashboard,” <https://covid19.who.int>, accessed: February 21, 2021.
- Zhao, Z., Li, X., Liu, F., Zhu, G., Ma, C., and Wang, L. (2020), “Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya,” *Science of The Total Environment*, 729, 138959.

Received: March 1, 2021

Accepted: April 28, 2021