THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

NATIONAL INSTITUTE FOR COMMUNICABLE DISEASES
Division of the National Health Laboratory Service

SUMMARY

Disease surveillance is a core function of the National Institute for Communicable Diseases (NICD), a Division of the National Health Laboratory Service (NHLS). This report includes an analysis of data on laboratory-confirmed COVID-19 deaths to estimate the effective reproductive number (R) of SARS-CoV-2 over time in South Africa nationally and in selected provinces where sufficient data are available. The basic reproductive number (R0) is the average number of secondary infections produced by a typical case of an infection in a population where everyone is susceptible which occurs typically in the first few weeks after introduction of a novel infectious agent into the population. The effective reproduction number (R) is the average number of secondary cases per infectious case in a population composed of both susceptible and non-susceptible hosts (once the infectious agent is circulating). If R>1, the number of new cases per time unit will increase, such as at the start of an epidemic. Where R=1, the number of new cases is stable over time, and where R<1, there will be a decline in the number of new cases per time unit.

This report is based on data collected up to 31 August 2020 (week 36 of 2020). The data were adjusted for the delay from illness onset to death and right censored for 30 days to account for the time lag between onset and reporting of death (R estimated up to 1 August). This analysis updates the report released on 21 August 2020. As in that report, here R is estimated from the data on laboratory-confirmed COVID-19 deaths. It is felt that laboratory-confirmed COVID-19 deaths are likely a relatively stable indicator of COVID-19 epidemic progression, albeit with a substantial time lag, which is a disadvantage of this approach. Note: COVID-19 is the name of the disease and SARS-CoV-2 is the name of the virus.

Highlights

- Nationally, the average R during the period of the stage 5 lockdown was 1.29 (95%CI: 1.06-1.58), rising to near 1.5 by the end of April.
- The daily R dropped steadily during stage 4 lockdown, with an average over this period of 1.27 (95%CI: 1.06 – 1.51).
- During the stage 3 lockdown, the daily R varied slightly, with an average of 1.05 (95%CI: 1.01 – 1.09), between 1 June and 1 August, dropping below 1 during the last weeks of July. This indicates ongoing transmission at a steadily slowing rate over this period, with numbers of new cases per infectious case dropping below 1 towards the end of the period.
- In the Western Cape Province, the average R during the stage 5 lockdown was 1.42 (95%CI: 1.08-1.86), dropping steadily throughout stage 4 and 3 lockdown. During the stage 3 lockdown, the daily R remained close to the threshold value of 1, with an average of 0.98 (95%CI: 0.93 – 1.02) from 1 June to 1 August. This indicates substantial slowing of transmission towards the end of June and through July.
- In other provinces where estimation was possible for stages 5, 4 and 3 (Gauteng, Eastern Cape and KwaZulu-Natal provinces), the R during the stage 5 and 4 lockdown ranged between 1.5 and 1.0. Generally R showed reductions during the stage 3 lockdown indicating slowing of transmission in all three provinces by the end of July.
- This report includes new estimates of R during the stage 3 lockdown for Mpumalanga and Free State provinces, indicating steady declines in R over this period.
- Reasons for the declines in R over the course of the stage 3 lockdown may include good public adherence to physical distancing, mask use and other measures, increasing population-level immunity, or other factors, including residual biases in the data.
- This analysis has important limitations. Changes in the ascertainment rate of COVID-19 deaths, the average risk profile of cases, as well as the delay between symptom onset and reporting of death may change over time, potentially affecting R estimation. The introduction of dexamethasone treatment and use of high flow nasal oxygen since mid-June may lower mortality, making estimation for recent timepoints less reliable. No local data are available from which to calculate the serial interval. Therefore, it is important to interpret these findings together with other sources of data on transmission. Caution should be exercised in interpreting comparisons between different timepoints and provinces as these could be affected by differential testing and reporting practices.

Methods

Daily R estimation

We used data from the first confirmed death in late March until 31 August 2020, based on the national deaths line list maintained by the National Department of Health (NDoH). The NDoH mortality data was linked with the national DATCOV dataset to obtain dates of symptom onset for 26% of fatal cases. The data were adjusted for the delay from illness onset to death and right censored for 30 days to account for reporting delays (last date of estimation: 1 August 2020). Missing dates of symptom onset were imputed using chained equations multiple (1000) imputations [i,ii]. In addition, 148 cases (1.0%) were missing both death date and date of symptom onset and were excluded from the analysis. The analyses for the Northern Cape and North West province were based on fatal cases in the DATCOV dataset because the NDoH data available at the time of compilation was incomplete for these provinces.

The time series is summarized in figures based on the median values for each date. A negative-binomial model was fitted to confirmed COVID-19 cases for which the date of symptom onset was available and used to impute the dates of symptom onset for cases with missing information. The model predictors for imputation were: health sector where death occurred (private, public, outside hospital, or unclassified), age group, month of death, and province. The daily R was estimated using the method of Thompson et al. (EpiEstim v. 2.2-3) [iii,iv] for each imputed dataset (1000 time series generated through the multiple imputation process). For the serial interval we used a gamma distribution with mean of 5.3 (s.d. 2.1) and standard deviation 1.8 (s.d. 0.6) to account for the variability (uncertainty) of the selected serial interval values [v]. We report the medians of the central values and the 2.5th-97.5th percentiles of the estimated daily R values obtained from 1000 imputed datasets [i,ii].

Limitations

The main limitation of this analysis is that the ascertainment rate of COVID-19 deaths may change over time, potentially affecting R estimation. Along with the ascertainment rate, the delay between symptom onset and reporting of death may change over time; cases at the end of the time series will be under-estimated according to the proportion of laboratory-confirmed COVID-19 deaths with delays longer than 30 days between symptom onset and reporting of death. Furthermore, the introduction of dexamethasone treatment in mid-June and use of oxygen administration via high flow nasal cannula may alter mortality outcomes, making estimation for recent timepoints less reliable. In addition, no local data are available from which to calculate the serial interval. The level of variation in the serial interval estimates used here reflects the range of estimates observed in mainland China [vi]. The time series of deaths was based on line lists provided by the National Department of Health (NDoH), linked with the DATCOV hospitalization dataset to ascertain, where possible, the dates of symptom onset for laboratory-confirmed COVID-19 deaths. Lastly, it was not possible to link deaths to the importation status of cases, affecting the reliability of initial R0 estimates, which are therefore not reported here.

Results

The daily number of laboratory-confirmed COVID-19 deaths steadily increased until mid-July 2020, following which daily numbers of new deaths has begun to decrease (Figure 1).

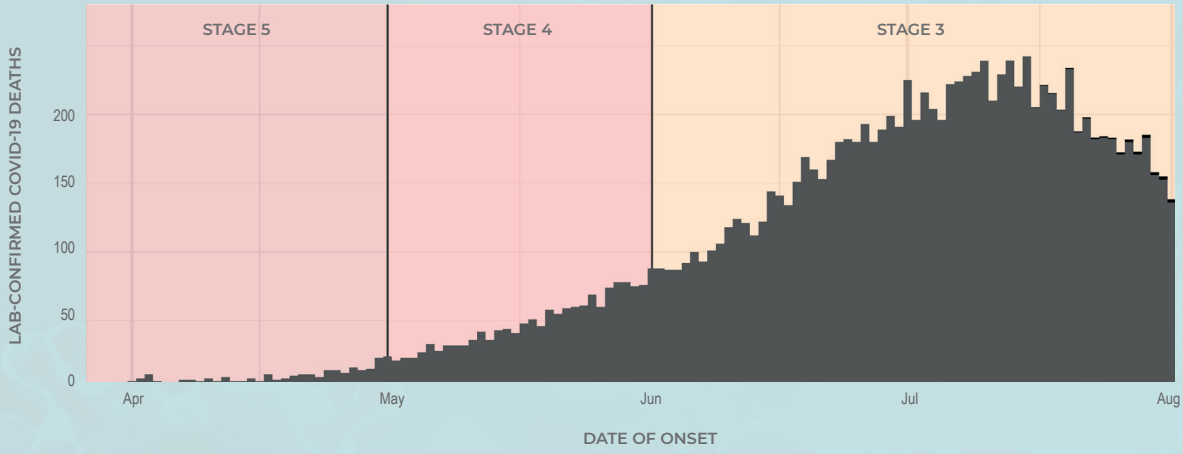

Figure 1. Daily number of COVID-19 deaths by date of symptom onset (missing data imputed; median of imputed time series is shown), South Africa (last date included: 1 August 2020). Black area shows adjustment for expected future deaths by date of onset.

THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

Nationally, the average R during the period of the stage 5 lockdown was 1.29 (95%CI: 1.06-1.58), rising to near 1.5 by the end of April (Figure 2). The daily R then dropped steadily throughout stage 4 lockdown, with an average over this period of 1.27 (95%CI: 1.06 – 1.51). During the stage 3 lockdown, the daily R varied slightly, with an average of 1.05 (95%CI: 1.01 – 1.09), between 1 June and 1 August. As of 1 August, R was estimated to be 0.88 (95% CI: 0.80-0.97).

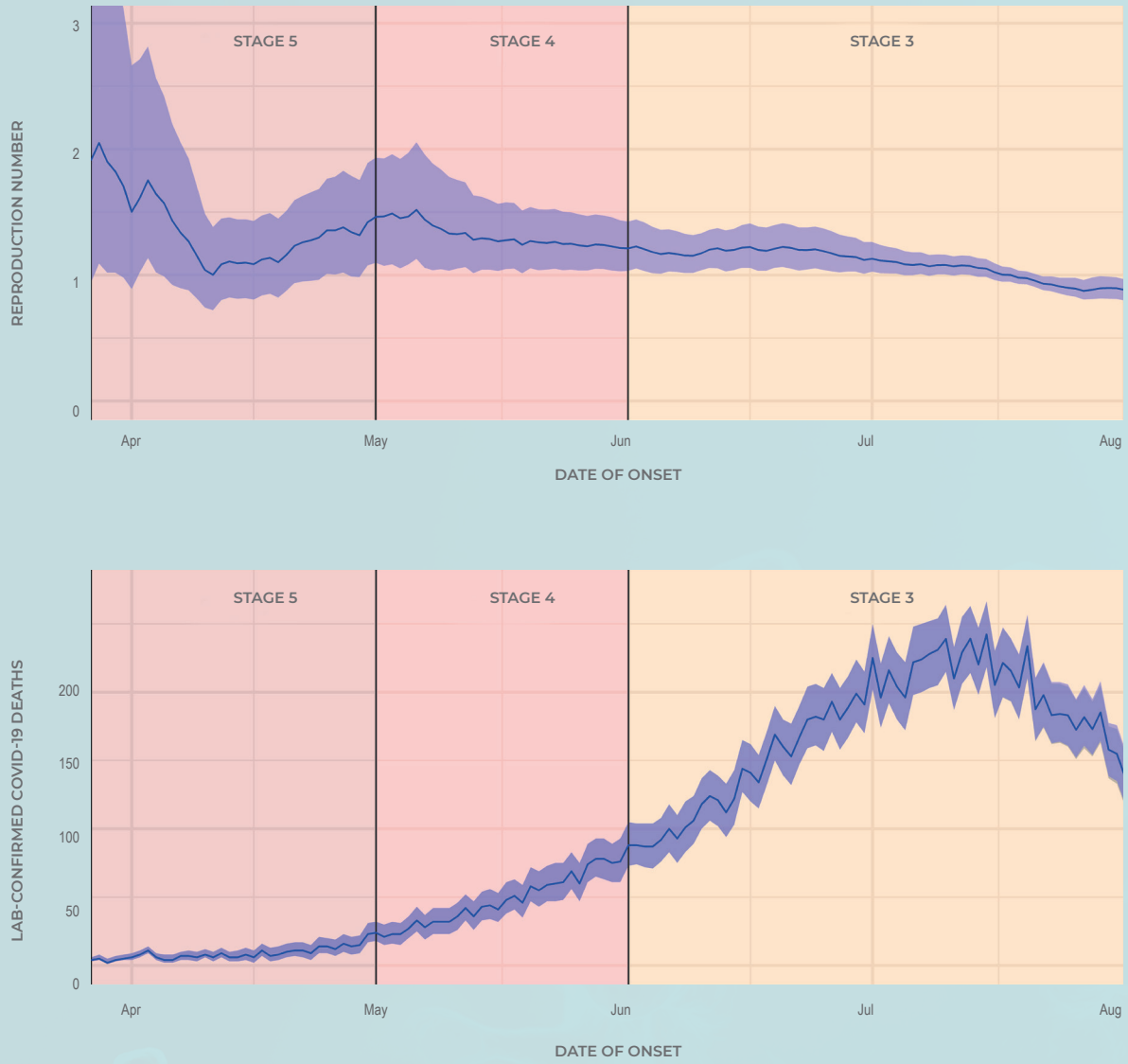

Figure 2. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, South Africa (last date included in the estimation: 1 August 2020). Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data imputed. The median and 95% range for the imputed datasets are shown.

THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

WEEK **38** 2020

In the Western Cape Province, the average R during the stage 5 lockdown was 1.42 (95%CI: 1.08-1.86), with a value near 1.4 at the end of April (Figure 3). The daily R then dropped steadily throughout stage 4 lockdown, with an average over this period of 1.22 (95%CI: 1.04 – 1.45). During the stage 3 lockdown, the daily R remained close to the threshold value of 1, with an average of 0.98 (95%CI: 0.93 – 1.02) from 1 June to 1 August. As of 1 August, R was estimated to be 0.96 (95% CI: 0.83-1.11).

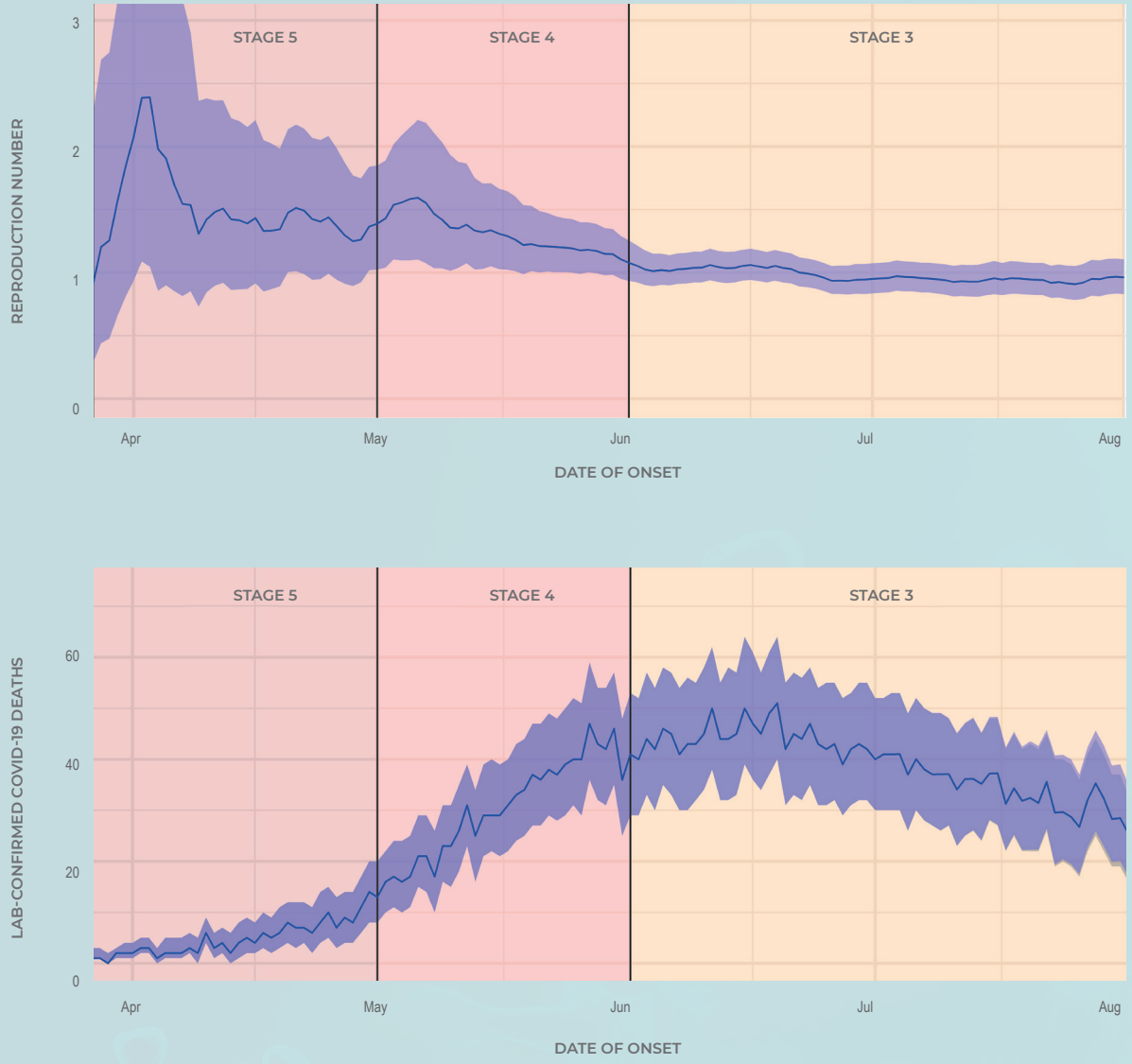

Figure 3. Western Cape (last date included in the estimation: 1 August 2020). Upper panel: Estimated daily R based on deaths, with 95% confidence intervals. Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data imputed. The median and 95% range for the imputed datasets are shown.

THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

In Gauteng, the average R during the stage 5 lockdown was 1.12 (95%CI: 0.71 – 1.67), with a value near 1.75 at the end of April (Figure 4). The daily R remained relatively steady over the period of the stage 4 lockdown, with an average over this period of 1.56 (95%CI: 1.11 – 2.18). During the stage 3 lockdown, the daily R declined, with an average of 1.06 (95%CI: 1.00 – 1.14) from 1 June to 1 August, dropping below 1 toward the end of July. As of 1 August, R was estimated to be 0.89 (95% CI: 0.78-1.01).

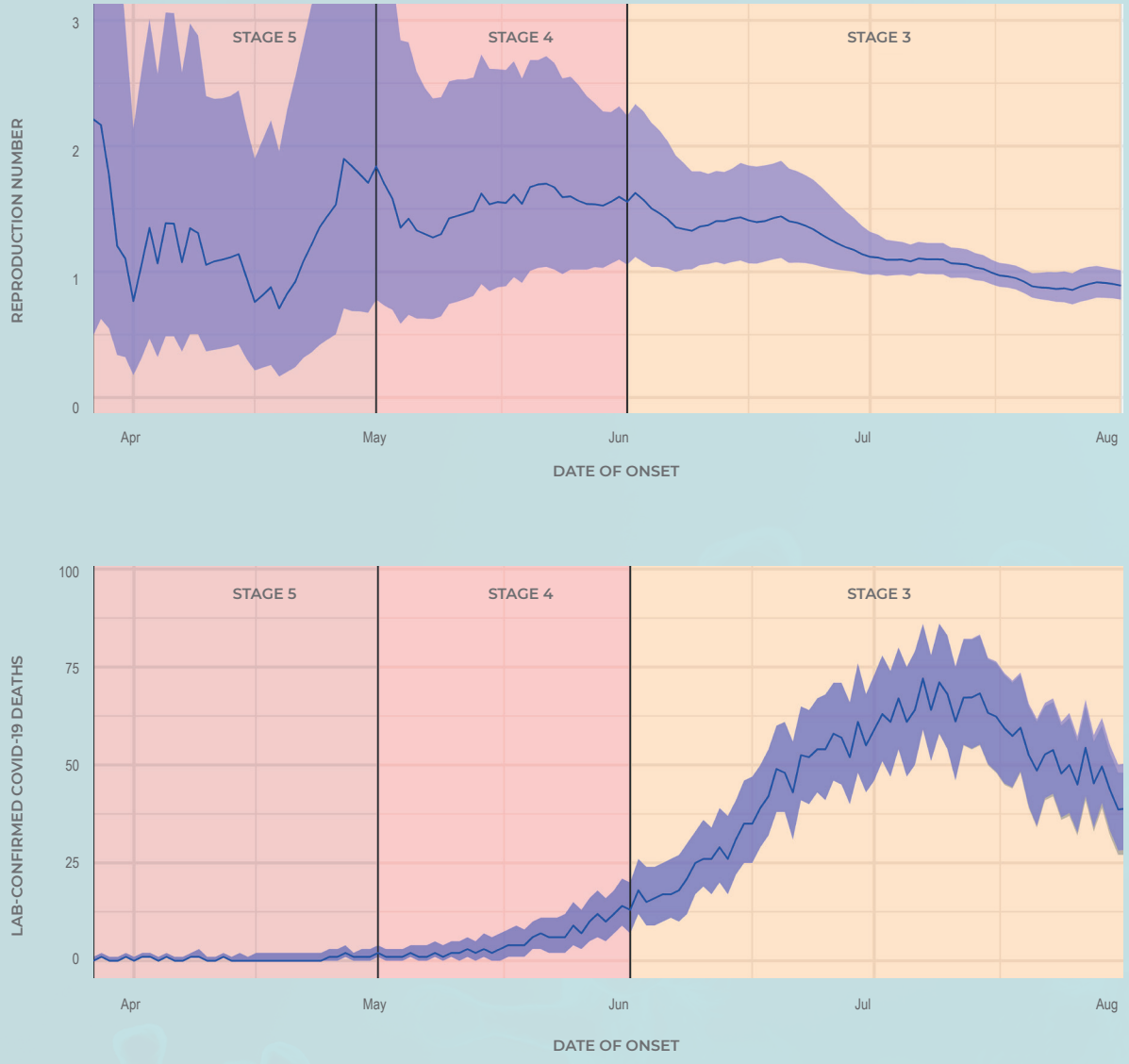

Figure 4. Gauteng (last date included in the estimation: 1 August 2020). Upper panel: Estimated daily R based on deaths, with 95% confidence intervals. Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data imputed. The median and 95% range for the imputed datasets are shown.

THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

In the Eastern Cape, the average R during the stage 5 lockdown was 1.35 (95%CI: 0.98 – 1.88), with a value near 1.5 at the end of April (Figure 5). The daily R fluctuated during stage 4 lockdown, with an average over this period of 1.28 (95%CI: 1.05 – 1.56). During the stage 3 lockdown, the daily R decreased gradually, with an average of 1.01 (95%CI: 0.97 – 1.05) from 1 June to 1 August, dropping below 1 in the second half of July. As of 1 August, R was estimated to be 0.70 (95% CI: 0.54-0.91).

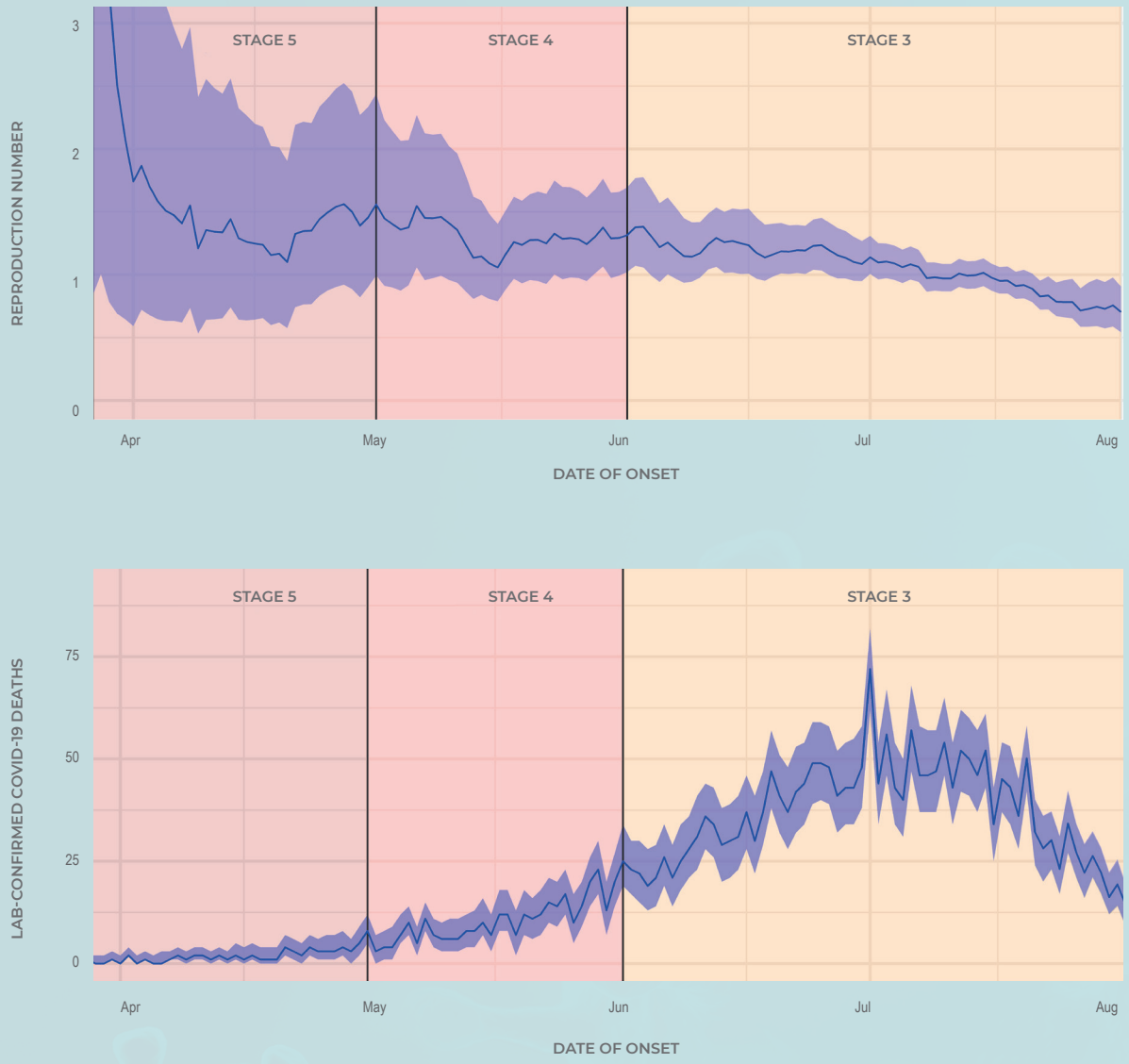

Figure 5. Eastern Cape (last date included in the estimation: 1 August 2020). Upper panel: Estimated daily R based on deaths, with 95% confidence intervals. Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data imputed. The median and 95% range for the imputed datasets are shown.

THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

In KwaZulu-Natal, the daily R fluctuated during the stage 4 lockdown, with an average over this period of 1.20 (95%CI: 0.85 – 1.65). During the stage 3 lockdown, the daily R remained initially well above 1, then gradually decreased until close to 1, with an average of 1.12 (95%CI: 1.02 – 1.25) from 1 June to 1 August. As of 1 August, R was estimated to be 0.93 (95% CI: 0.81-1.05).

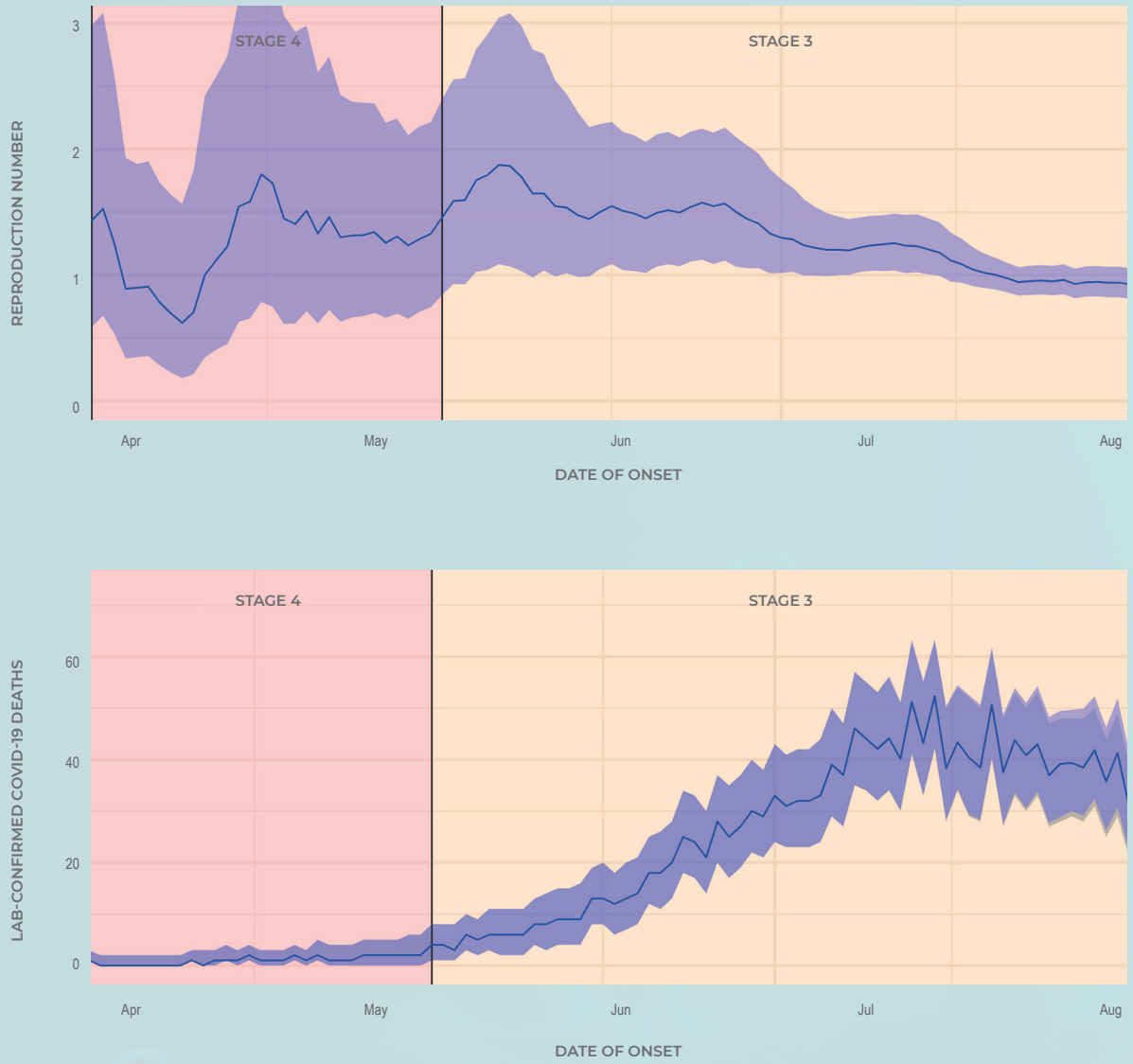

Figure 6. KwaZulu-Natal (last date included in the estimation: 1 August 2020). Upper panel: Estimated daily R based on deaths, with 95% confidence intervals. Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data imputed. The median and 95% range for the imputed datasets are shown. Plots show estimates from 1 May, as daily R estimates are unstable early in the time series due to small numbers of deaths.

In Mpumalanga, the daily R decreased gradually from approximately 2 at the beginning of the stage 3 lockdown, to near 1 at the end of August. The average R during the stage 3 lockdown was 1.15 (95%CI: 0.99 – 1.34) from 1 June to 1 August. As of 1 August, R was estimated to be 0.95 (95% CI: 0.70-1.26).

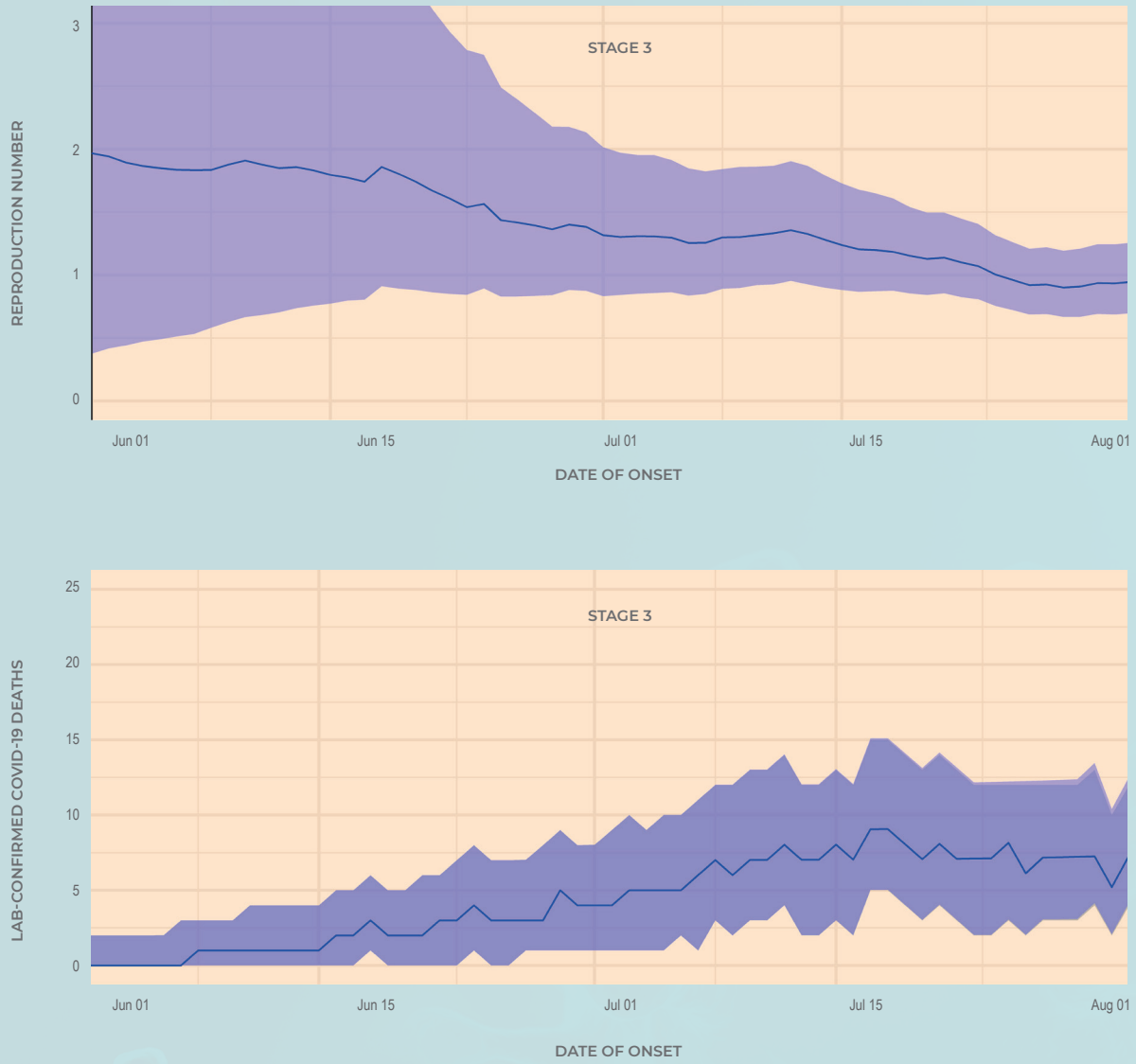

Figure 7. Mpumalanga (last date included in the estimation: 1 August 2020). Upper panel: Estimated daily R based on deaths, with 95% confidence intervals. Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data imputed. The median and 95% range for the imputed datasets are shown. Plots show estimates from 1 June, as daily R estimates are unstable early in the time series due to small numbers of deaths.

THE DAILY COVID-19 EFFECTIVE REPRODUCTIVE NUMBER (R) IN SOUTH AFRICA

WEEK **38** 2020

In Free State, the daily R fluctuated during the stage 3 lockdown, with an average of 1.20 (95%CI: 1.03 – 1.43) from 1 June to 1 August. As of 1 August, R was estimated to be 0.97 (95% CI: 0.80-1.16).

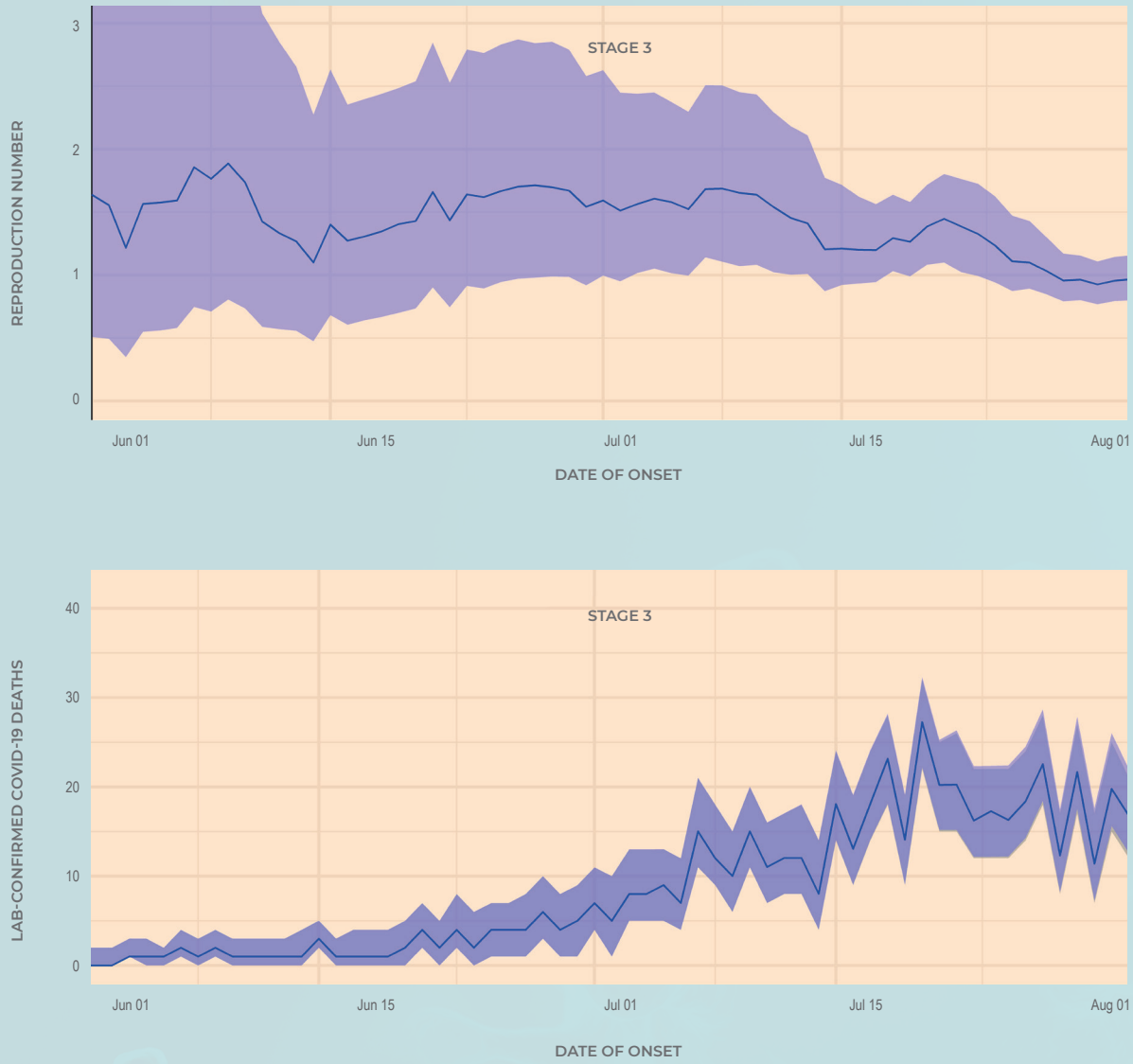

Figure 8. Free State (last date included in the estimation: 1 August 2020). Upper panel: Estimated daily R based on deaths, with 95% confidence intervals. Lower panel: estimated number of laboratory-confirmed COVID-19 deaths by onset date with missing data. The median and 95% range for the imputed datasets are shown. Plots show estimates from 1 June, as daily R estimates are unstable early in the time series due to small numbers of deaths.

This report was jointly prepared by the National Institute for Communicable Disease (NICD) and the DSI-NRF Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA). Inquiries should be referred to Prof Cheryl Cohen (cherylc@nicd.ac.za).

References

i van Buuren, S. (2018) Flexible Imputation of Missing Data, Second Edition. Chapman and Hall/CRC: New York. DOI: 10.1201/9780429492259. https://stefvanbuuren.name/fimd

ii Kleinke, K, & Reinecke, J. (2015) Multiple imputation of overdispersed multilevel count data. In: Uwe Engel (Ed.), Survey Measurements. Techniques, Data Quality and Sources of Error (pp. 209–226). Frankfurt A. M.: Campus/The University of Chicago Press. http://press.uchicago.edu/ucp/books/book/distributed/S/bo22196267.html

iii Thompson, RN, JE Stockwin, RD van Gaalen, JA Polonsky, ZN Kamvar, PA Demarsh, E Dahlqwist, S Li, E Miguel, T Jombart, J Lessler, S Cauchemez, and A Cori. (2019) Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics 29: 100356. DOI: 10.1016/j.epidem.2019.100356

iv Cori, A. (2020) EpiEstim: Estimate time varying reproduction numbers from epidemic curves. R package version 2.2-3. https://github.com/mrc-ide/EpiEstim

v Pitzer, VE, Chitwood, M, Havumaki, J, Menzies, NA, Perniciaro, M, Warren, JL, Weinberger, DM, and T Cohen. (2020) The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. medRxiv https://doi.org/10.1101/2020.04.20.20073338

vi Ali, ST, L Wang, EHY Lau, XK Xu, Z Du, Y Wu, GM Leung, and BJ Cowling. (2020) Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. Science eabc9004. DOI: 10.1126/science.abc9004