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

The Daily COVID-19 Effective Reproductive Number (R) in the South Africa

Week 39 of 2022

Summary

Overview of report

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 cases, hospital admissions, and deaths to estimate the effective reproductive number (R) of SARS-CoV-2 over time in South Africa at the national and provincial levels. The basic reproductive number (R_0) 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 30 September 2022 (week 39 of 2022). The data were adjusted for the delays from illness onset to case report, hospital admission, and death and right censored for 2, 7, and 7 days respectively to account for the time lag between each outcome (test result, admission, or death) and the time of reporting (last date of estimation based on laboratory-confirmed cases: 28 September 2022). This analysis updates the report released on 1st August 2022. **In this report, R is estimated from the data on laboratory-confirmed COVID-19 cases as well as laboratory-confirmed COVID-19 hospital admissions and hospital-based deaths. In addition, while earlier reports included only primary infections in cases-based analyses, this report, as with the previous three reports, includes reinfections.** There may be non-overlapping sources of bias for the three data sources, which motivates a comparison of R estimates. Note: COVID-19 is the name of the disease and SARS-CoV-2 is the name of the virus.

Highlights

- R estimates based on cases were above 1 nationally and in Gauteng and Western Cape provinces at the end of the estimation period. The credible intervals for R estimates at the end of the estimation period included 1 for all provinces.
- R estimates based on admissions were close to 1 nationally and in all four of the provinces for which numbers of admissions were sufficient to produce reasonable estimates, at the end of the admissions-based estimation period.
- The numbers of tests being conducted are at the lowest level since the beginning of the pandemic and are decreasing. The low number of tests leads to increased uncertainty in the R estimates. The decreasing trend in testing reduces the apparent rate of epidemic growth and may bias R estimates downward. The results presented in this report must be interpreted with these caveats in mind.

Methods

Daily R estimation

We used data from 1 February 2022 until 30 September 2022, based on the national DATCOV dataset on hospitalized cases, and the linelist of laboratory-confirmed cases, datasets maintained by the National Institute for Communicable Diseases (NICD). This report includes both primary infections and suspected reinfections; positive tests were classified as being associated with a suspected reinfection if the time since the most recent positive test for that patient was at least 90 days [iv]. The laboratory-confirmed case data was linked with the national DATCOV dataset to obtain dates of symptom onset. Following data linkage, symptom onset data were available for 6% of laboratory-confirmed cases, while dates of onset were available for 61% of hospitalized cases, and 27% of fatal cases in the DATCOV dataset. The data were adjusted for the delay from symptom onset to reporting of test result / hospital admission and right censored for 2, 7, and 7 days (for cases, hospital admissions, and deaths respectively) to account for reporting delays (last date of estimation based on laboratory-confirmed cases: 28 September 2022). The provincial level DATCOV data have different end-dates, so the provincial time series were individually adjusted for right-censoring; national-level analyses were based on the pooled provincial-level time series, truncated to the earliest end date of provincial time series. Missing dates of symptom onset were imputed using chained equations multiple imputations (50) [i,ii]. 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. Separate imputations were done for the case and admissions datasets. The hospital-based deaths data set is a subset of admissions, so the same set of imputations were used. The model predictors for the imputation procedure on the cases dataset were: health sector where sample collection occurred

(private or public), age group (in ten-year age bands), month of case report, patient sex, and province. The model predictors for the imputation procedure on the hospital admissions dataset were: health sector where admission occurred (private or public), age group (in ten-year age bands), patient outcome, day of the week of hospital admission, month of hospital admission, and health district in which admission occurred. The daily R was estimated using the method of Thompson *et al*. (EpiEstim v. 2.2-3) [iii,iv] for each imputed dataset. For the serial interval we used a gamma distribution with mean of 6.6 (s.d. 3.3) and standard deviation of 0.5 (s.d. 0.27) to account for the variability (and uncertainty) of the selected serial interval values. Parameters were estimated by fitting a gamma distribution to data from PHIRST-C, a community cohort study of COVID-19 transmission [vii]. We report the medians of the central values and the 2.5^{th}-97.5^{th} percentiles of the estimated daily R values obtained from the imputed datasets [i,ii]. Small numbers of hospital-based deaths in Mpumalanga and Northern Cape led to large fluctuations and wide credible intervals in deaths-based R estimates for these provinces; as a result, this report does not include deaths-based R estimates for Mpumalanga and Northern Cape provinces.

Previous versions of this report have included descriptions of trends in daily R values during lockdown levels 5 through 1, and the adjusted level 3 lockdown, as well as the 2021 lockdown levels 1 through 4. The current report focuses on more recent trends in daily R values, starting on 1 March 2022 (for more details regarding the COVID-19 lockdowns in South Africa, please refer to the South African government website [vi]).

Results

Nationally, R estimates increased slightly through March 2022, increasing to just below 2 in April 2022, then decreased gradually until the second week of May 2022, when R estimates dropped below 1. R began increased at the beginning of July, then remained steady through late August before increasing to slightly above 1, where it remained through the end of the estimation period.

Trends at the province level were generally similar, with R increasing in late March 2022, peaking in late April 2022 or early May 2022, and then declining with timing varying somewhat by province. R estimates began gradually increasing in all provinces in approximately July, and again in August or September, with R above 1 in Gauteng and Western Cape at the end of the estimation period, though the credible interval for R estimates based on cases included 1 for all provinces at the end of the estimation period. R estimates for several provinces showed considerable levels of uncertainty to due to relatively low numbers of cases and hospital admissions (Table 1).

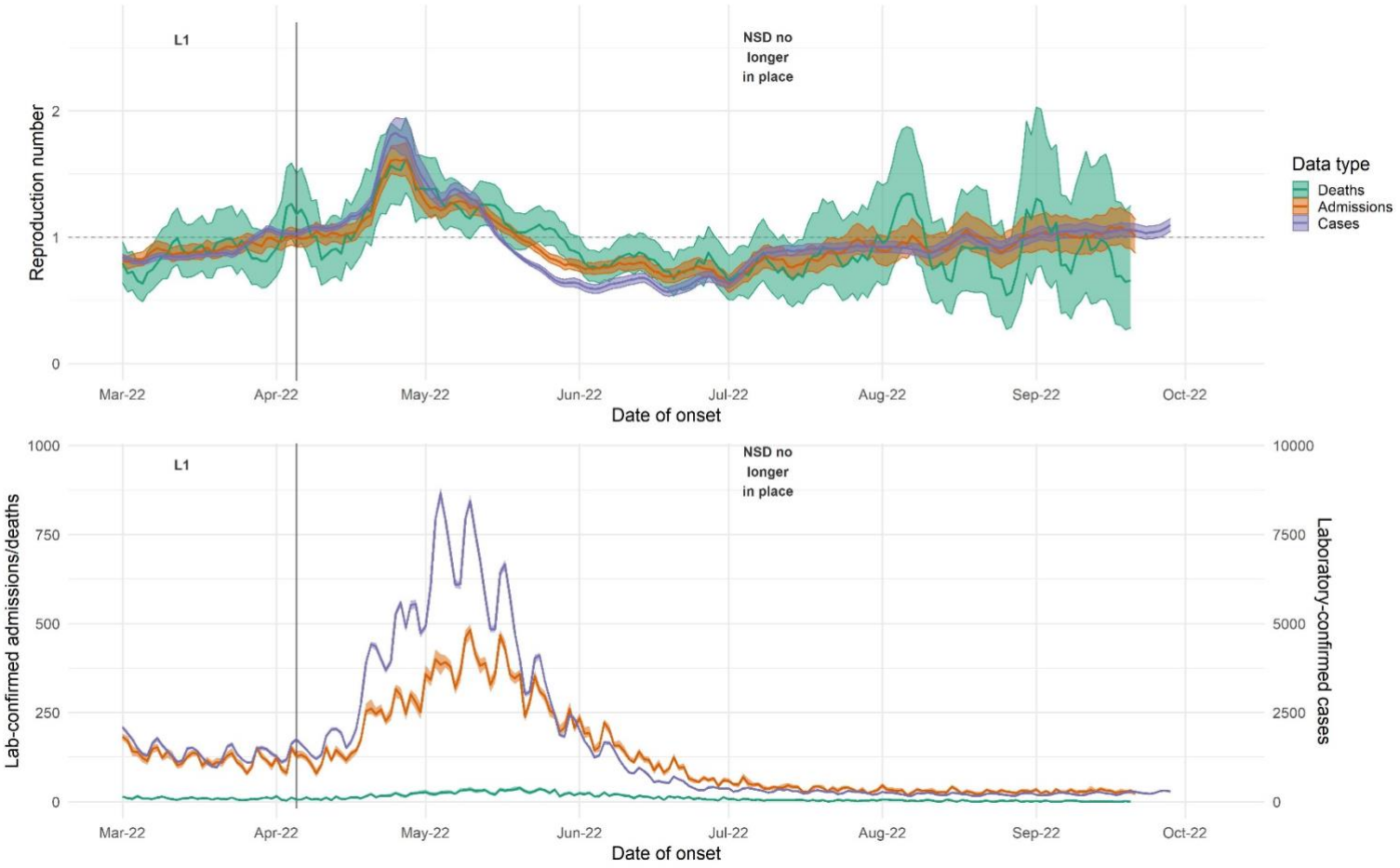

Figure 1. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, South Africa (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

Table 1: Daily R estimates by province for R based on laboratory-confirmed cases and laboratory-confirmed hospital admissions from healthcare and laboratory service providers. Each cell contains median values with 95% credible intervals.

	Cases	Cases	Admissions
	28 September 2022	21 September 2022	21 September 2022
National	1.10 (1.05,1.15)	1.05 (1.00,1.10)	1.01 (0.87,1.14)
Eastern Cape	1.05 (0.83,1.28)	0.93 (0.75,1.17)	1.27 (0.76,1.96)
Free State	0.99 (0.76,1.25)	1.00 (0.78,1.25)	N/A
Gauteng	1.14 (1.07,1.22)	1.09 (1.01,1.17)	1.05 (0.84,1.28)
KwaZulu-Natal	0.95 (0.83,1.08)	0.95 (0.82,1.07)	0.87 (0.62,1.20)
Limpopo	1.05 (0.77,1.38)	0.92 (0.68,1.23)	N/A
Mpumalanga	0.79 (0.59,1.01)	1.03 (0.82,1.29)	N/A
Northern Cape	0.85 (0.33,1.71)	0.86 (0.37,1.70)	N/A
North West	0.88 (0.70,1.10)	0.90 (0.71,1.12)	N/A
Western Cape	1.16 (1.04,1.29)	1.21 (1.07,1.36)	1.02 (0.71,1.37)

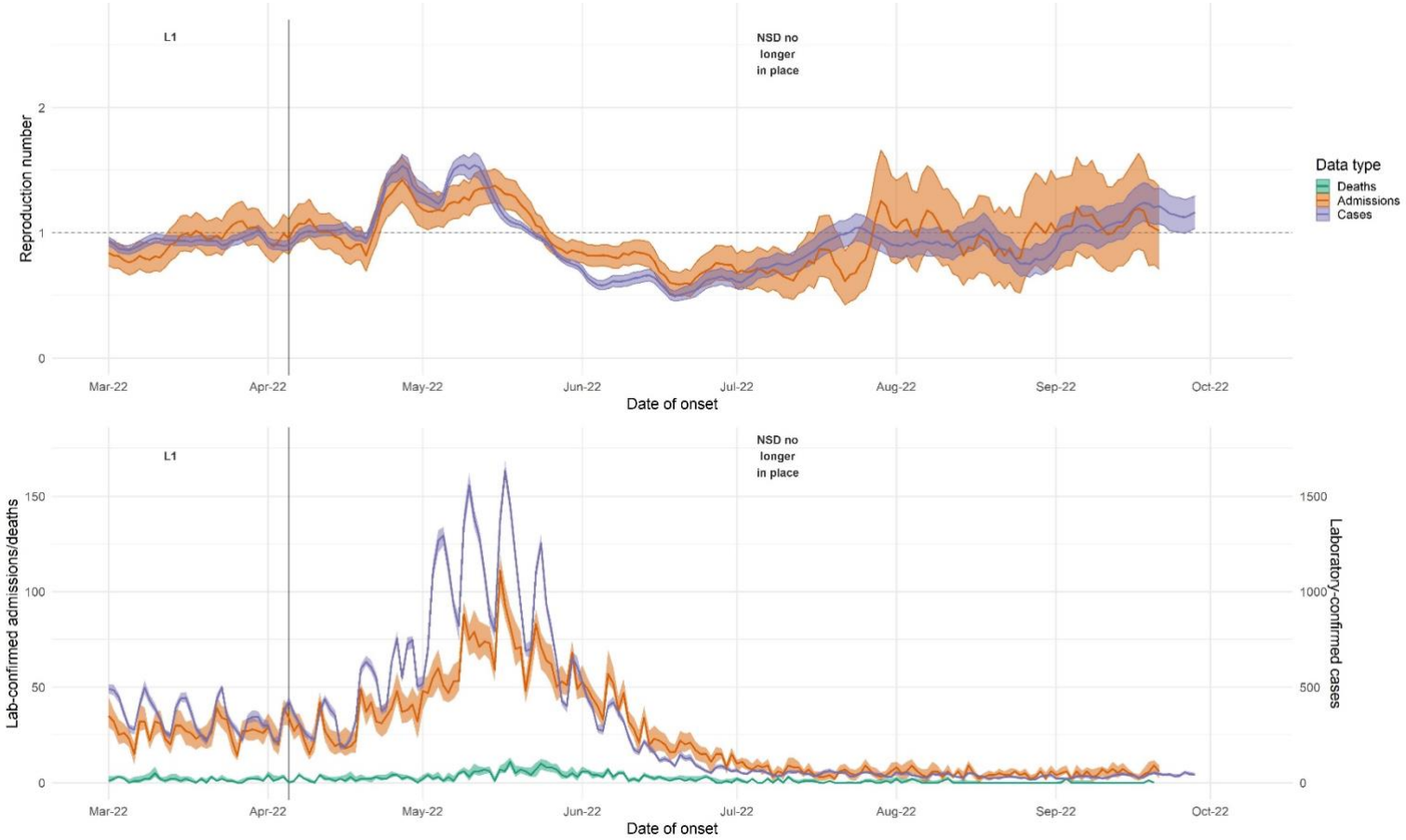

Figure 2. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, Western Cape (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

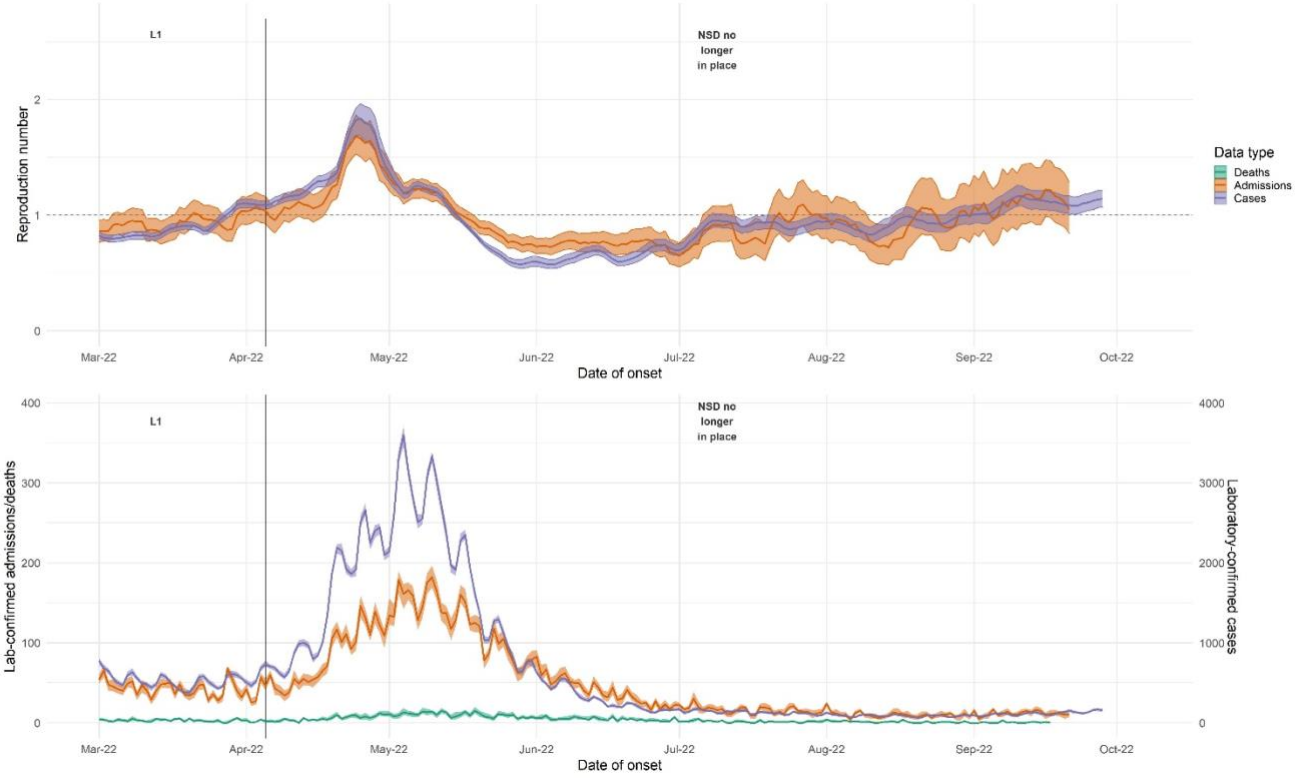

Figure 3. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals Gauteng (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

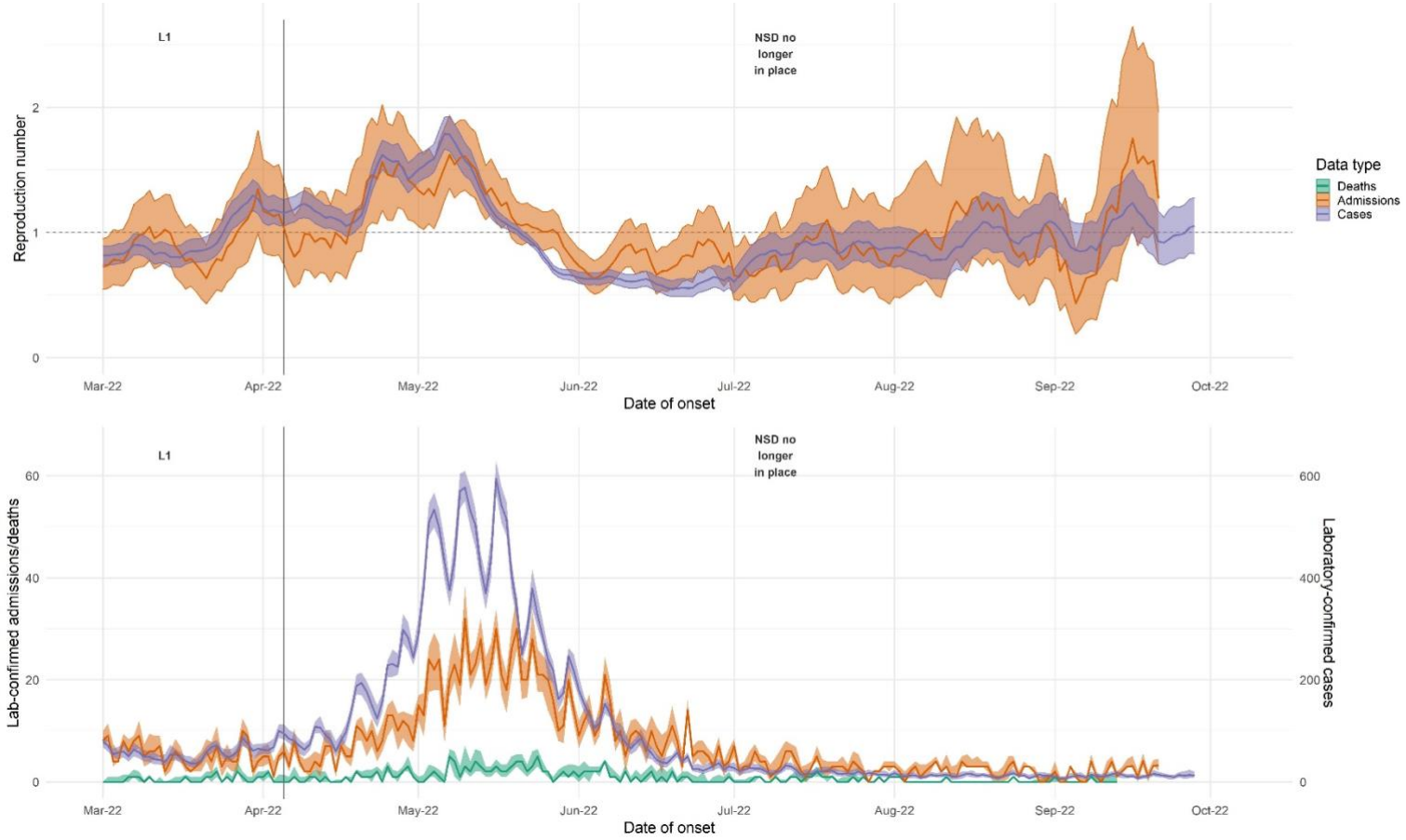

Figure 4. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, Eastern Cape (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

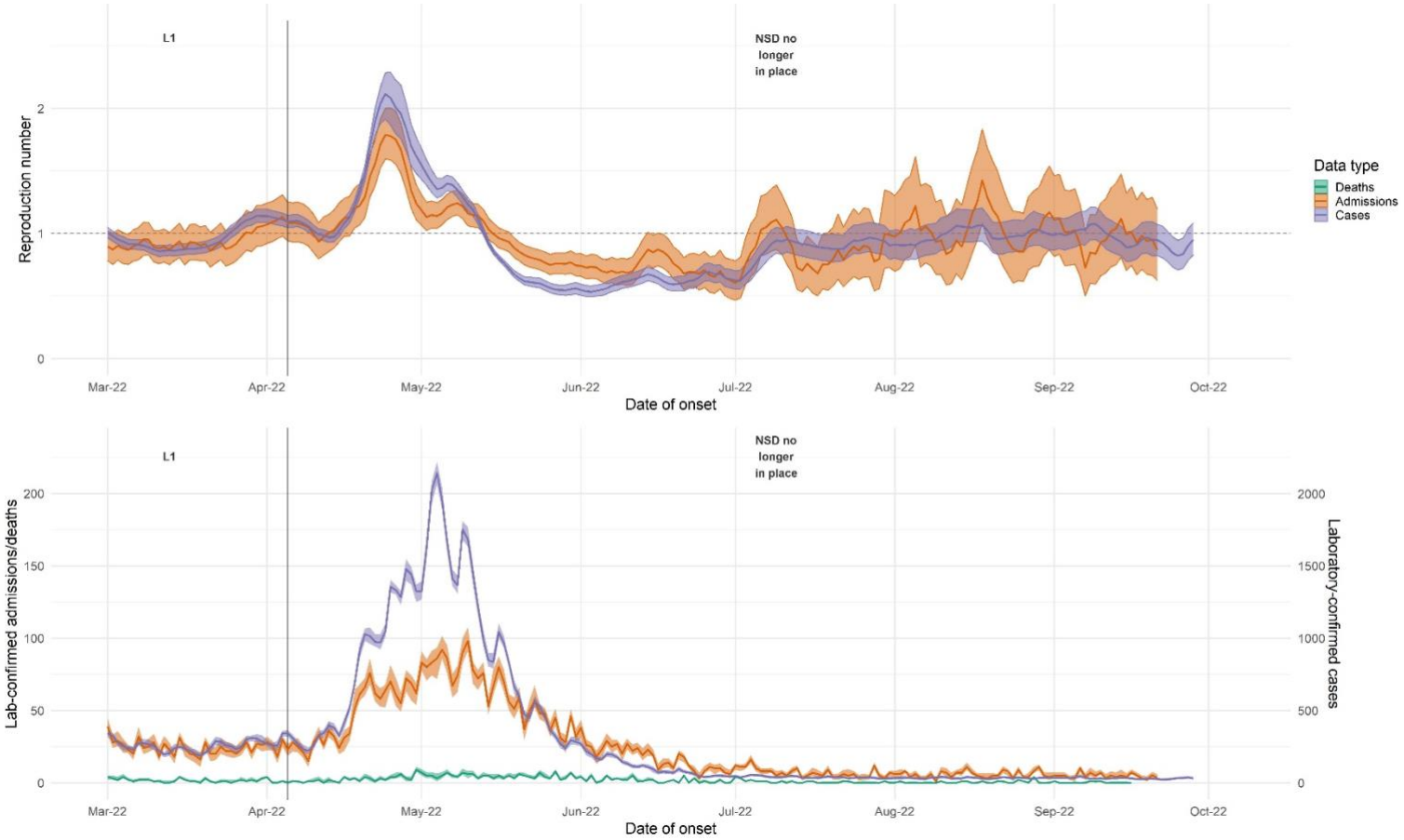

Figure 5. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, KwaZulu-Natal (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

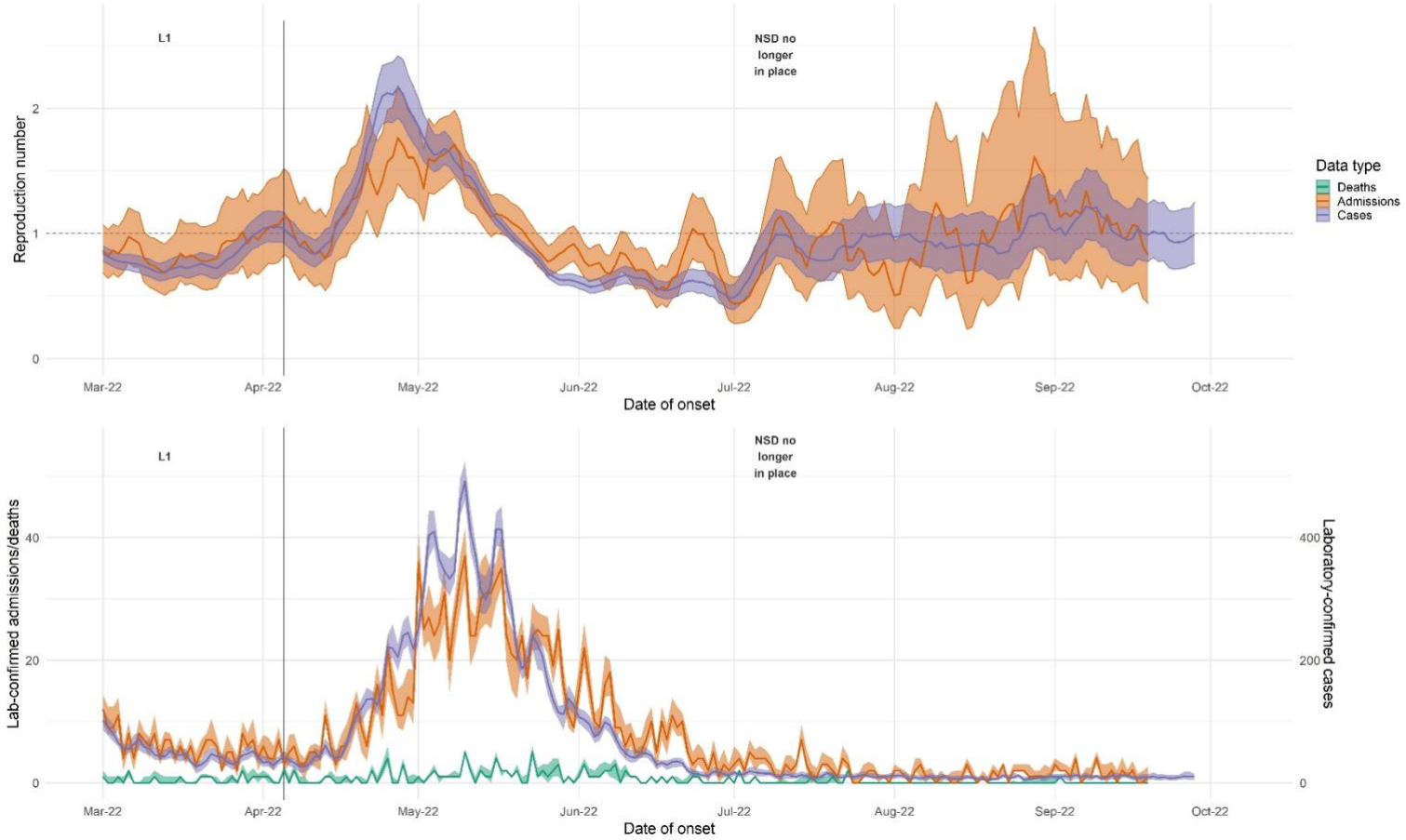

Figure 6. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, Free State (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases. hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

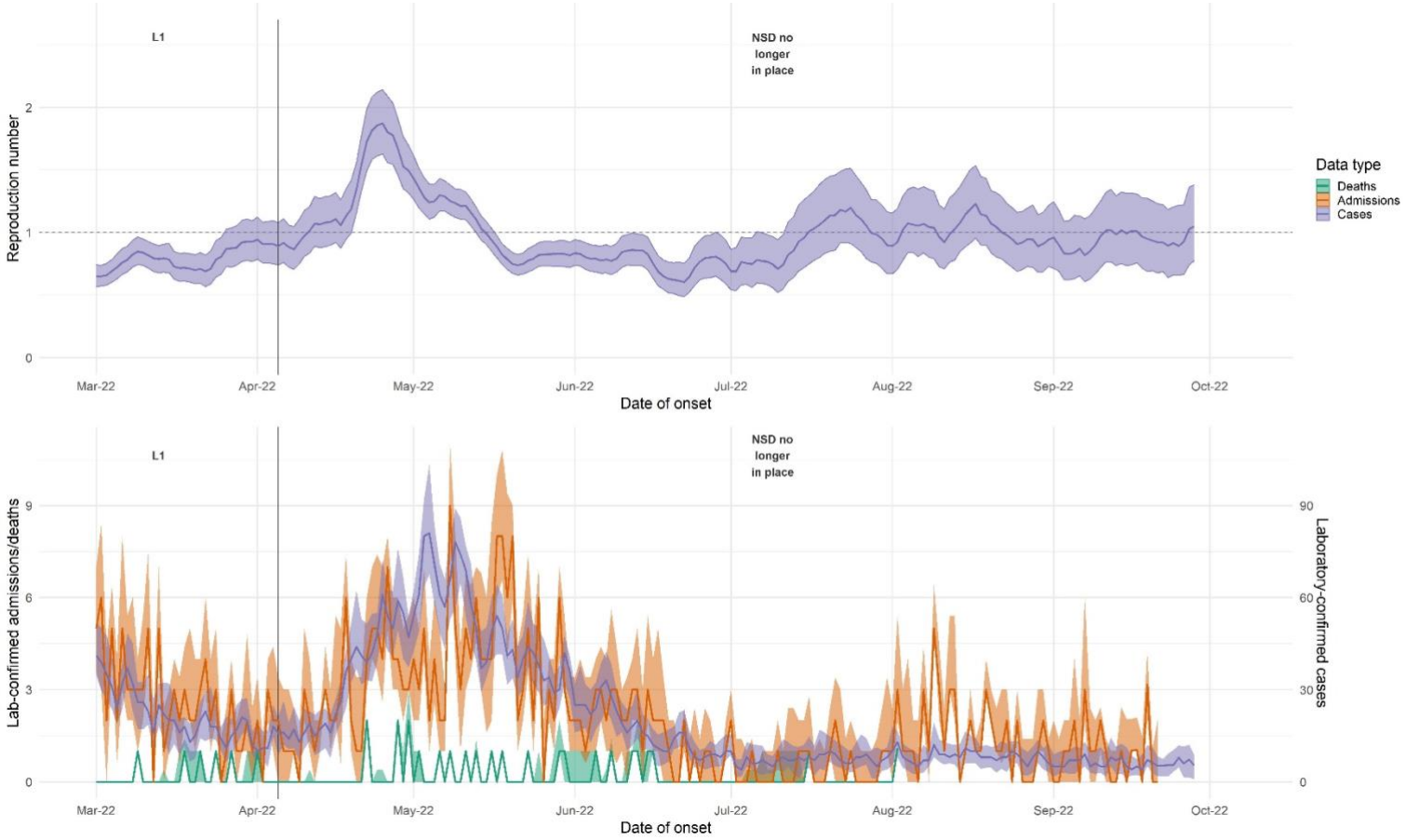

Figure 7. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, Limpopo (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases. Hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

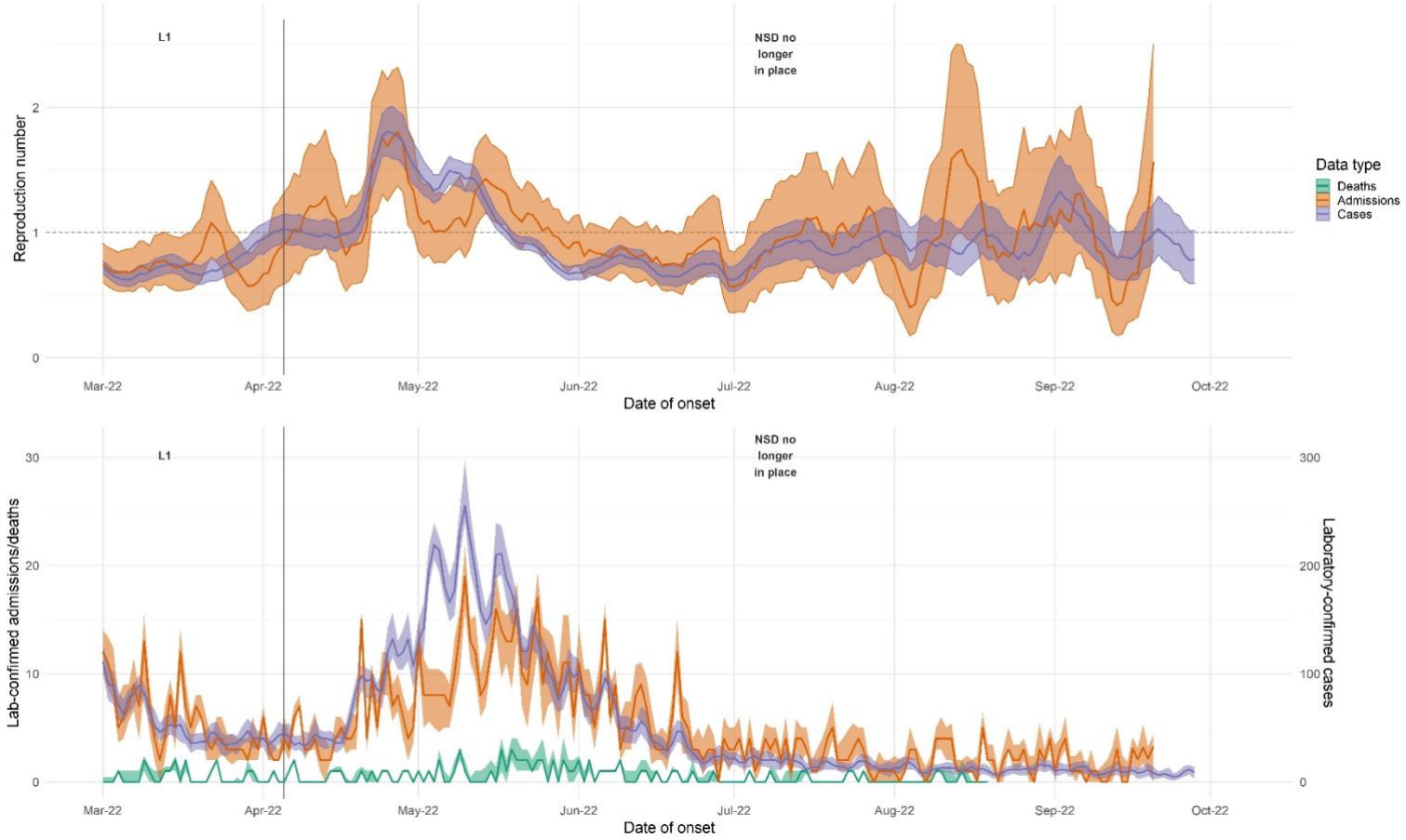

Figure 8. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, Mpumalanga (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

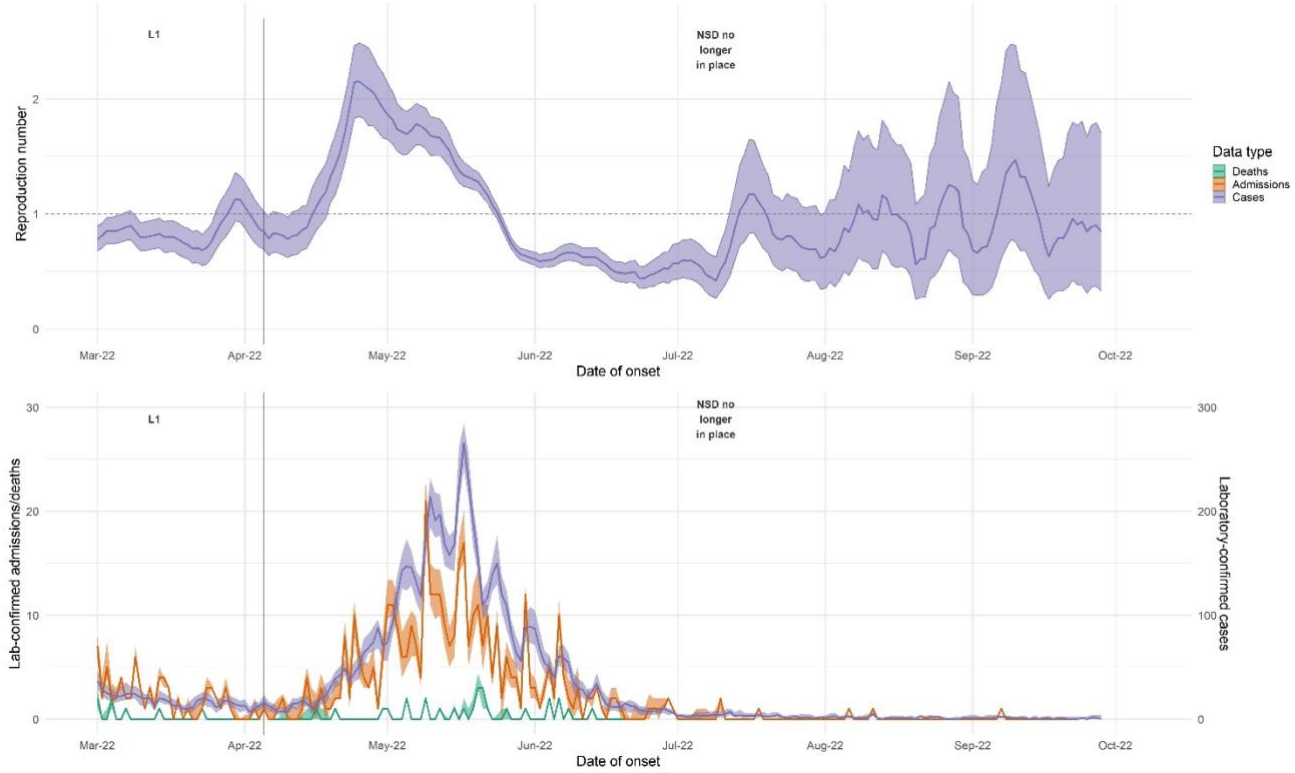

Figure 9. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, Northern Cape (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

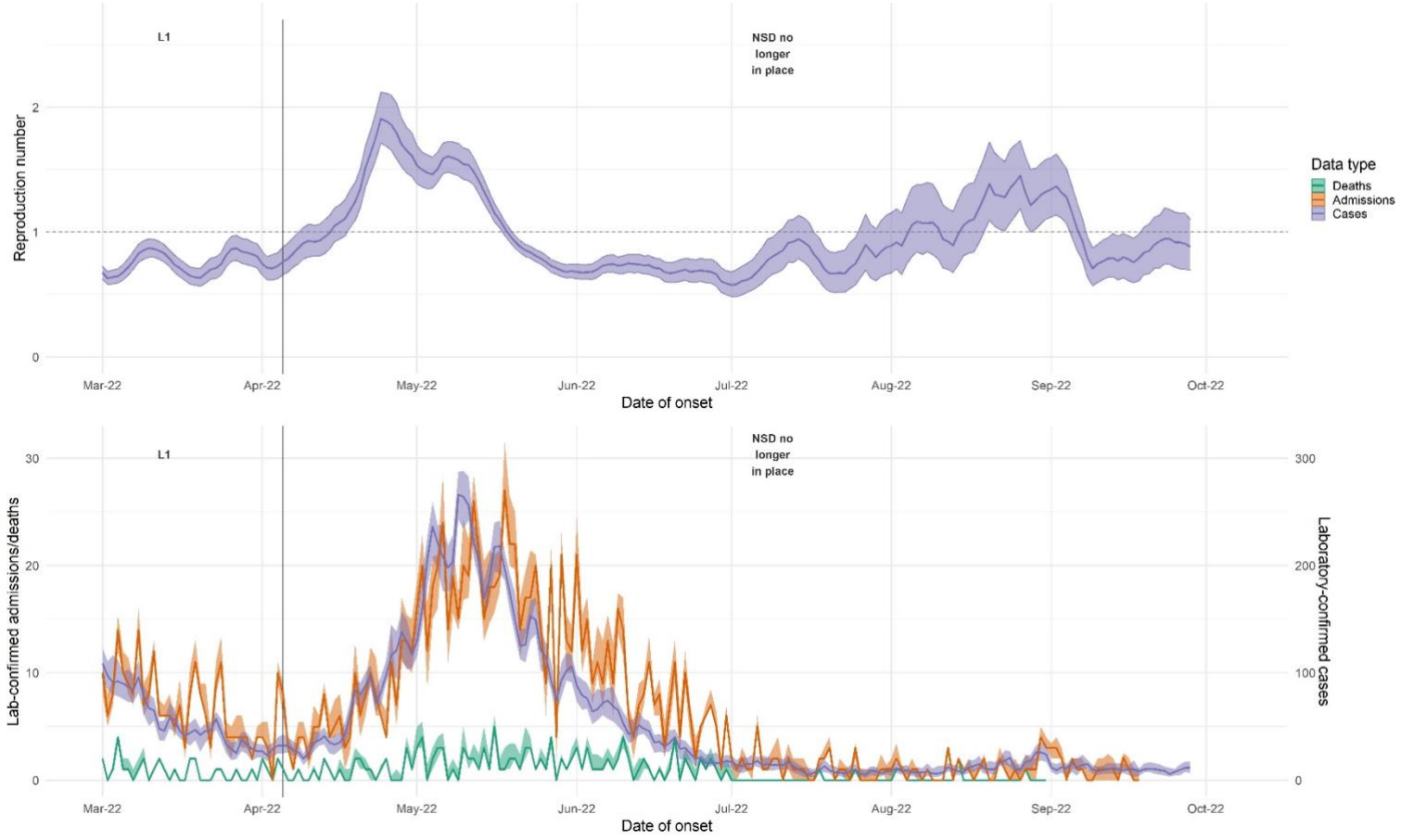

Figure 10. Upper panel: Estimated daily reproduction number (R), with 95% credible intervals, North West (last date included in the estimation: 28 September 2022). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

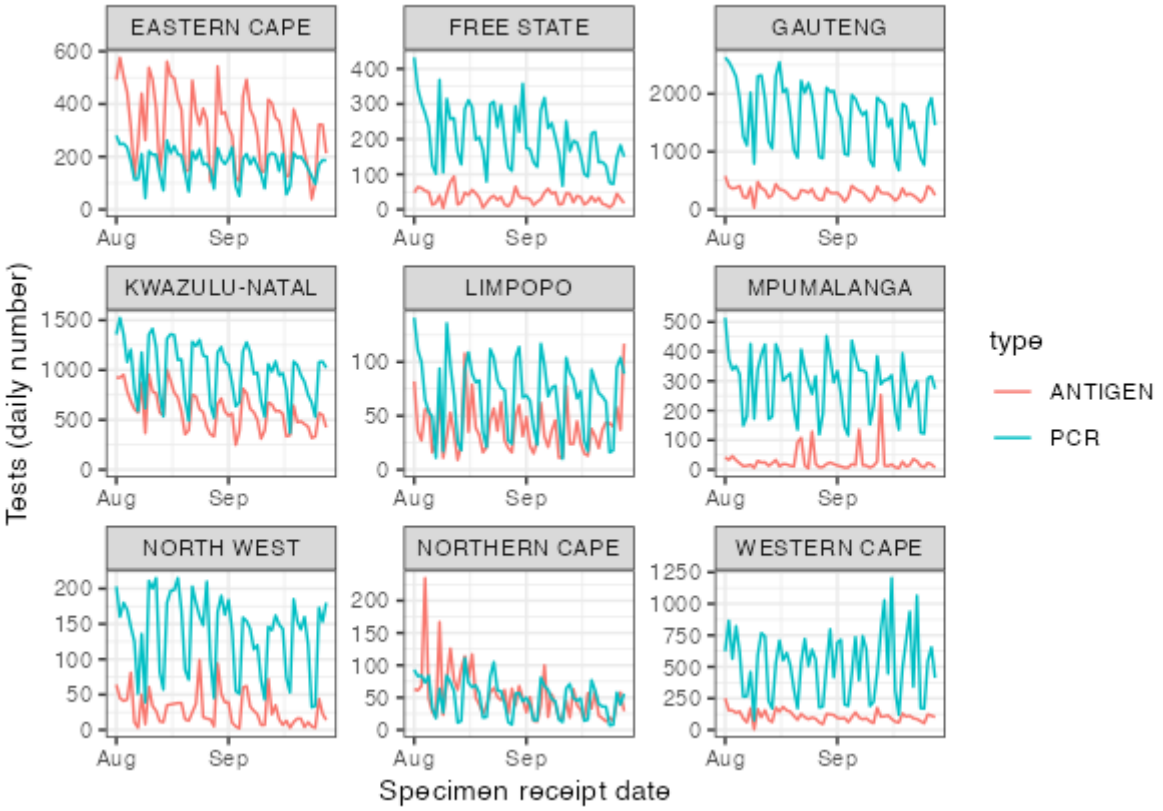

Figure 11. Daily volumes of antigen (red) and polymerase chain reaction (PCR, blue) tests by specimen receipt date, per province, between 1st August and 28th September 2022.

Limitations

The main limitation of this analysis is that the ascertainment rate of COVID-19 cases and deaths, along with the proportion of cases which are admitted to hospital, may change over time, potentially affecting R estimation. These effects are likely driven in part by changes in the criteria for testing and hospital admission, by shifting care seeking behavior during the epidemic, by increasing vaccination coverage, and by differences in severity e.g. by variant. In addition, small numbers of deaths in province-level analyses, as well as low and decreasing test volumes (see figure 11) lead to large credible intervals.

Along with the ascertainment rate, the delay between symptom onset and reporting of case/admission/death may change over time, which would affect the accuracy of the adjustment for right-censoring the end of the time series. Furthermore, a number of factors may have altered severity outcomes over time, including increasing vaccination coverage, changes in quality of healthcare provided if health systems are overwhelmed, and potential differences between earlier circulating viruses, the Beta (501Y.V2 / B.1.351) variant that dominated the second wave, the Delta (B.1.617.2) variant which was dominant during the third wave, the Omicron (B.1.1.529) variant

which was dominant during the fourth wave, and the Omicron BA.4 and BA.5 sub-variants currently dominant [viii]. Combined, these factors may lead to perturbations in the time series data that are unrelated to transmission. Comparing R estimates from the three data endpoints may help in assessing the severity of some of these biases, as indicated by inconsistent results across analyses of the three data endpoints. In addition to limitations in the ability of the available time series data to reflect underlying transmission, the serial interval distribution is estimated based on a relatively small data set.

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 Government of South Africa. (2021). Regulations and guidelines – coronavirus COVID-19. Retrieved from https://www.gov.za/covid-19/resources/regulations-and-guidelines-coronavirus-covid-19#

vii Cohen, C., Kleynhans, J., von Gottberg, A., McMorrow, M., Wolter, N., Bhiman, J., ... & Tempia, S. (2021). SARS-CoV-2 incidence, transmission and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020-2021. *medRxiv*. https://www.medrxiv.org/content/10.1101/2021.07.20.21260855v1.full.pdf

viii Network for Genomic Surveillance in South Africa (NGS-SA). (2021). SARS-CoV-2 Sequencing Update, 1 December 2021. National Institute for Communicable Diseases (NICD) of the National Health Laboratory (NHLS). South Africa. https://www.nicd.ac.za/wp-content/uploads/2021/12/Update-of-SA-sequencing-data-from-GISAID-1-Dec-Final.pdf

iv Pulliam, JRC van Schalkwyk, C, Govender, N, von Gottberg, A, Cohen, C, Groome, MJ, Dushoff, J, Mlisana, K, Moultrie, H (2021). Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. medRxiv 2021.11.11.21266068. https://doi.org/10.1101/2021.11.11.21266068