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

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

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 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 22 March 2021 (week 12 of 2021). 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 (R estimated up to 15 March). This analysis updates the report released on 20 Feb 2021. In this report, R is estimated from the data on laboratory-confirmed COVID-19 cases, hospital admissions, and, on a national level, hospital-based deaths. There may be non-overlapping sources of bias for the three data sources, which motivates a comparison of R estimates. R estimates are described for each of the lockdown levels implemented by the South African government – for more information regarding the timing and nature of lockdowns see the South African government website [vii]. Note: COVID-19 is the name of the disease and SARS-CoV-2 is the name of the virus.

Highlights

- Nationally, the average R remained initially steady during the adjusted level 3 lockdown, then dropped steadily for the first three weeks of January 2021. R stabilized briefly around the end of January, then gradually increased through the beginning of level 1 lockdown in March with a value close to one at the end of the estimation period.
- The trends in all provinces were generally similar to the national trends with R near to 1 at the end of the estimation period, except in the in Northern Cape, where R exceeds 1.
- The increasing trend in R indicates increasing transmission and may in part be driven by relaxation of controls to prevent the spread of COVID-19. It is essential that recommended measures to control the spread of COVID-19, including physical distancing, hand hygiene, good ventilation, adherence to venue capacity limits, and wearing of masks, are consistently implemented.
- This analysis has important limitations. Changes in the ascertainment rate of COVID-19 cases and deaths, the proportion of cases admitted to hospital, the delay between symptom onset and reporting, and other factors may change over time, potentially affecting R estimation. In addition, the relatively low numbers of deaths recorded between waves results in high levels of uncertainty and large fluctuation in R estimates based on daily deaths. Furthermore, a number of factors may have altered mortality outcomes over time, including treatment changes, pressure on the hospital system, and potential differences in severity between earlier circulating viruses and the 501Y.V2 / B.1.351 variant that dominated the second wave. Combined, these factors may lead to perturbations in the time series data that are unrelated to transmission. 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 and differences in healthcare provision.

Methods

Daily R estimation

We used data from the first confirmed case in March 2020 until 22 March 2021, based on the national DATCOV dataset on hospitalized cases and in-hospital deaths, and the laboratory-confirmed case line list maintained by the National Institute for Communicable Diseases (NICD). The laboratory-confirmed cases 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 53% of hospitalized cases, and 56% of fatal cases in the DATCOV dataset. 94 cases (0.04%) in the DATCOV database were missing both admission date and date of symptom onset and were excluded from the analyses based on hospital admissions and deaths. 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: 15 March 2021). 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 two imputation procedures were: health sector where sample collection/hospital admission occurred (private or public), age group, month of case report/hospital admission, outcome (for admissions), day of hospital admission (for admissions), 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. 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 (and 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 the imputed datasets [i,ii].

Previous versions of this report have included descriptions of trends in daily R values during lockdown levels 5 through 1, and parts of the adjusted level 3 lockdown. The current report focuses on more recent trends in daily R values, starting at the beginning of the adjusted level 3 lockdown on 29 December 2020 (for more details regarding the COVID-19 lockdowns in South Africa, please refer to the South African government website [vii]).

Results

The daily number of laboratory-confirmed COVID-19 cases steadily increased until mid-July 2020, following which daily numbers of new cases decreased steadily, flattening around mid-September through to early November, when they began to increase (Figure 1). From December through to early January, the daily number of laboratory-confirmed COVID-19 cases increased rapidly, followed by a rapid decrease through to the end of February 2021, after which case counts remained steady.

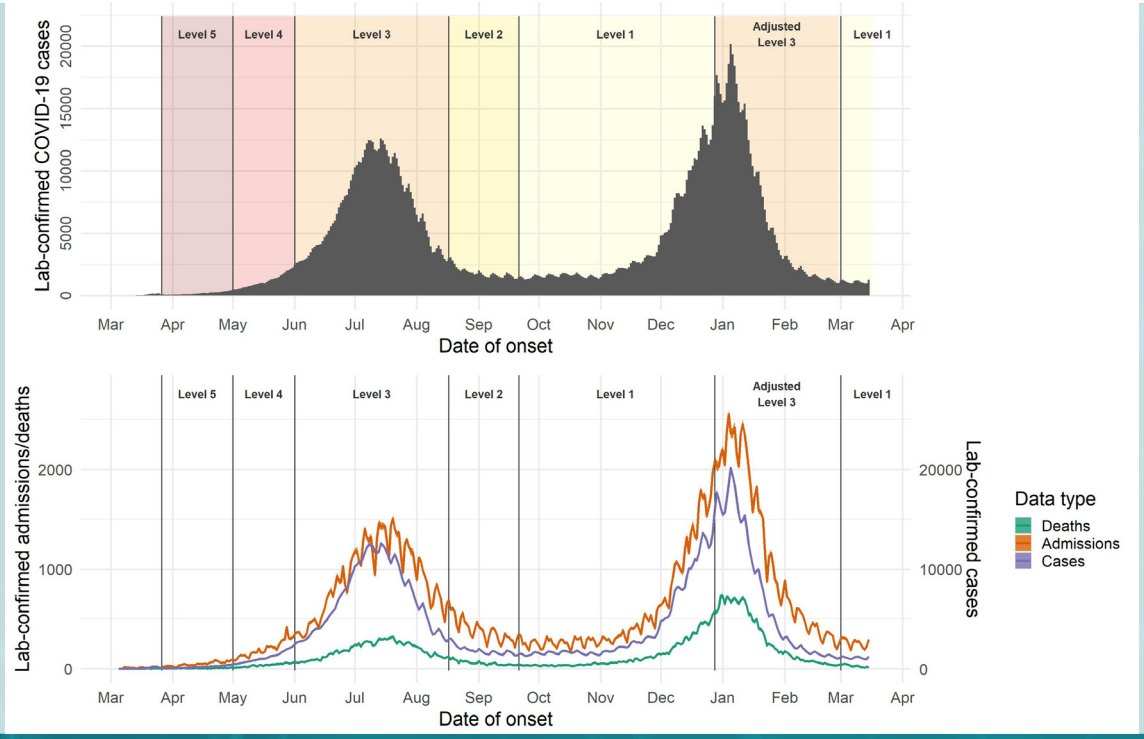

Figure 1. Daily number of laboratory-confirmed COVID-19 cases (above) and daily numbers of laboratory-confirmed COVID-19 cases, hospital admissions, and deaths (below), by date of symptom onset (missing data imputed; medians and 95% quantiles of imputed time series are shown), South Africa (last date included: 22 March 2021). Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

During the adjusted level 3 lockdown, and the beginning of the level 1 lockdown in March 2021, estimates of R using three different data sources, daily numbers of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-based deaths were generally similar (Figure 2). Nationally, the average R remained initially steady during the adjusted level 3 lockdown, then dropped steadily for the first three weeks of January 2021. R stabilized briefly around the end of January, then gradually increased through the beginning of level 1 lockdown in March, with a value close to 1 at the end of the estimation period (Figure 3).

Trends at the province level were generally similar to those at a national level, with R in the Northern Cape exceeding 1 at the end of the estimation period (Figures 4-10 and Table 1). In some provinces R stabilized close to 1 during the level 1 lockdown in March.

Table 1. Daily R estimates nationally and by province for 15th March 202, based on lab-confirmed cases and hospital admissions. Each cell contains median values with 95% confidence intervals.

	Cases	Admissions
National	0.99 (0.97,1.02)	0.94 (0.88,1.01)
Western Cape	0.86 (0.79,0.96)	0.90 (0.80,1.00)
Gauteng	0.97 (0.93,1.02)	0.86 (0.75,1.00)
Eastern Cape	0.90 (0.76,1.09)	0.89 (0.68,1.14)
KwaZulu-Natal	0.88 (0.78,0.96)	1.04 (0.85,1.21)
Free State	1.05 (0.97,1.16)	1.00 (0.83,1.19)
Northern Cape	1.40 (1.09,1.67)	1.48 (1.04,1.97)
North West	0.97 (0.90,1.06)	1.12 (0.98,1.28)
Mpumalanga	1.03 (0.96,1.10)	0.89 (0.75,1.05)
Limpopo	0.92 (0.79,1.06)	1.08 (0.84,1.37)

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

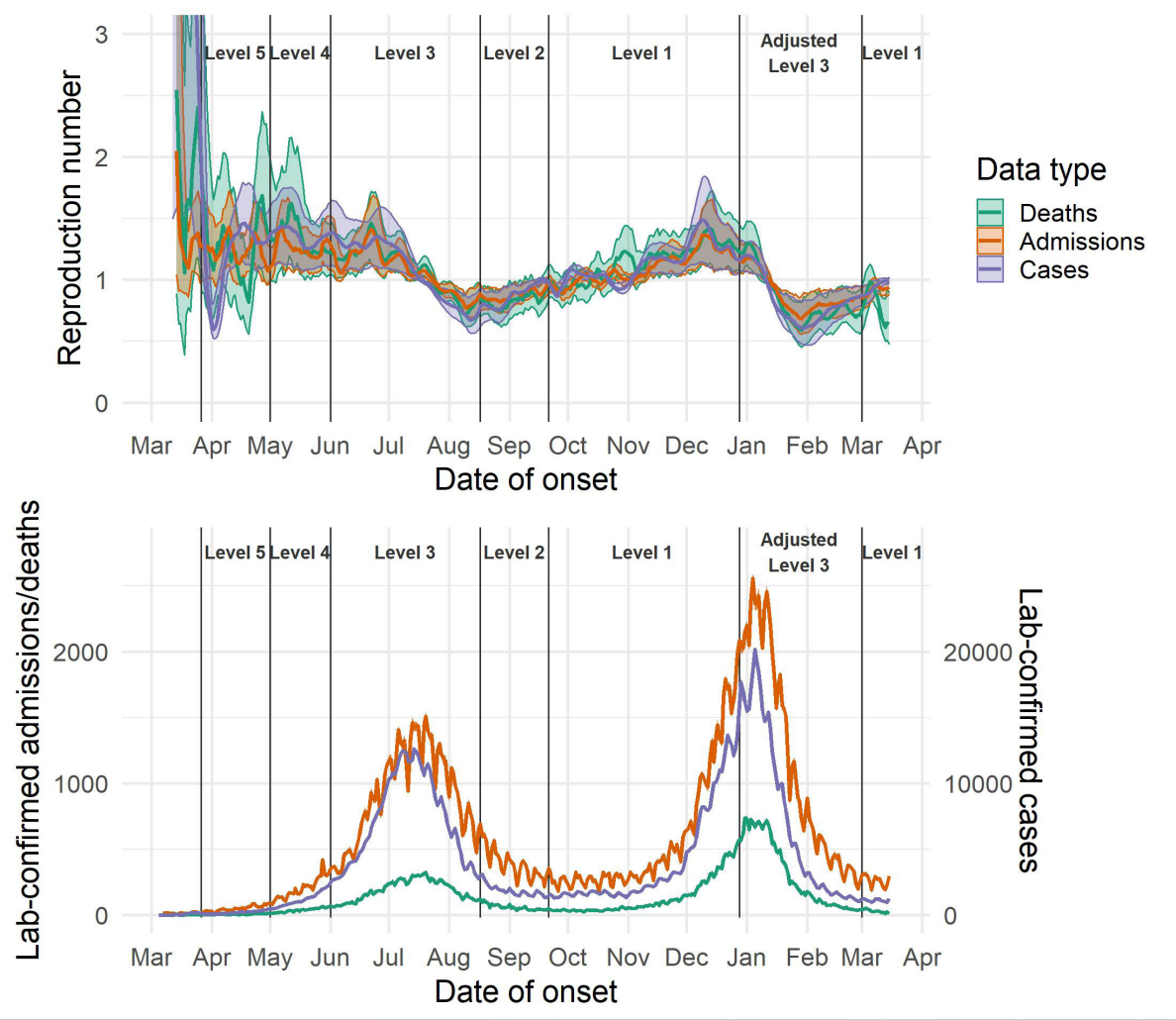

Figure 2. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, South Africa (last date included in the estimation: 15 March). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and deaths, by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

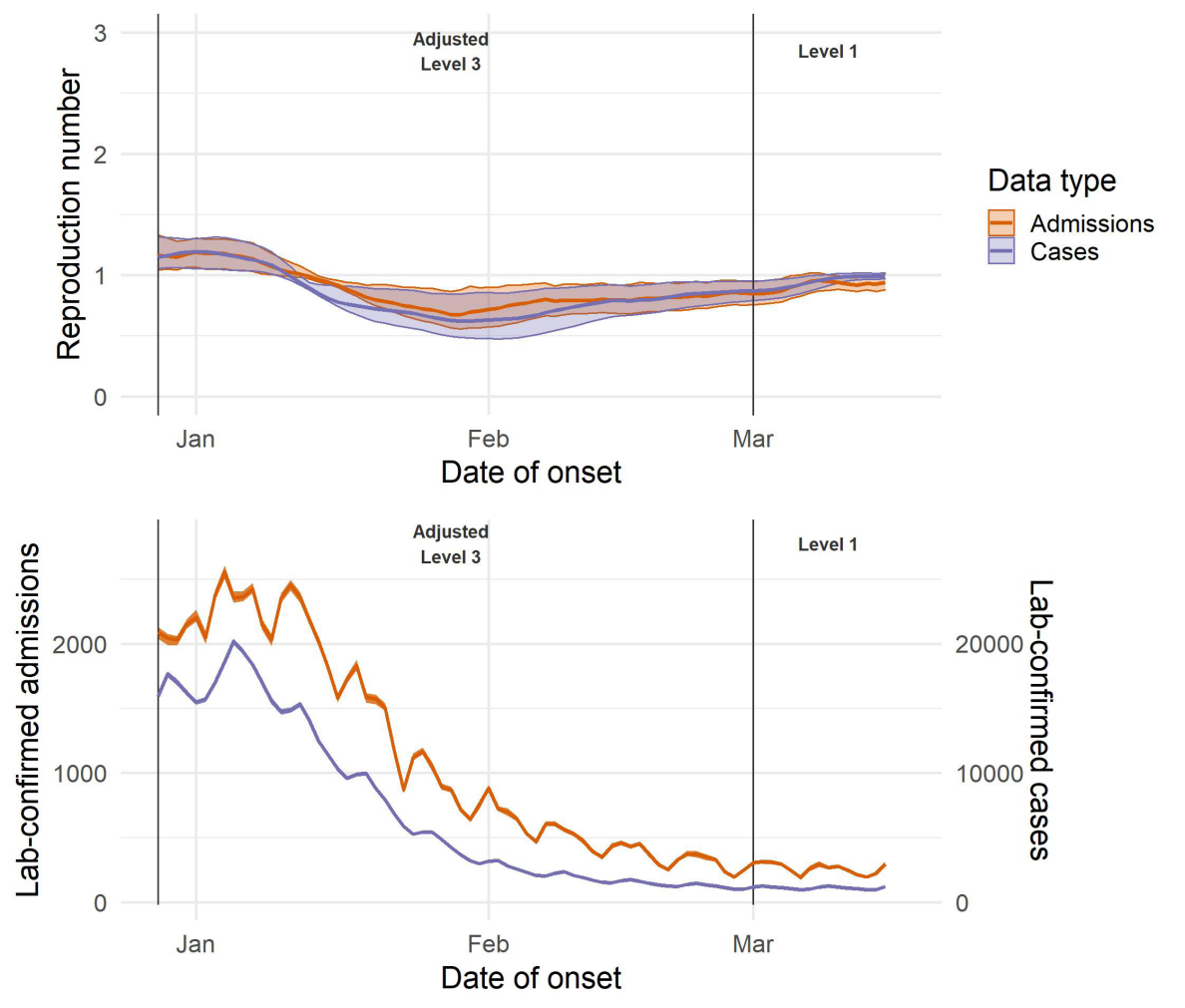

Figure 3. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, South Africa (last date included in the estimation: 15 March). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

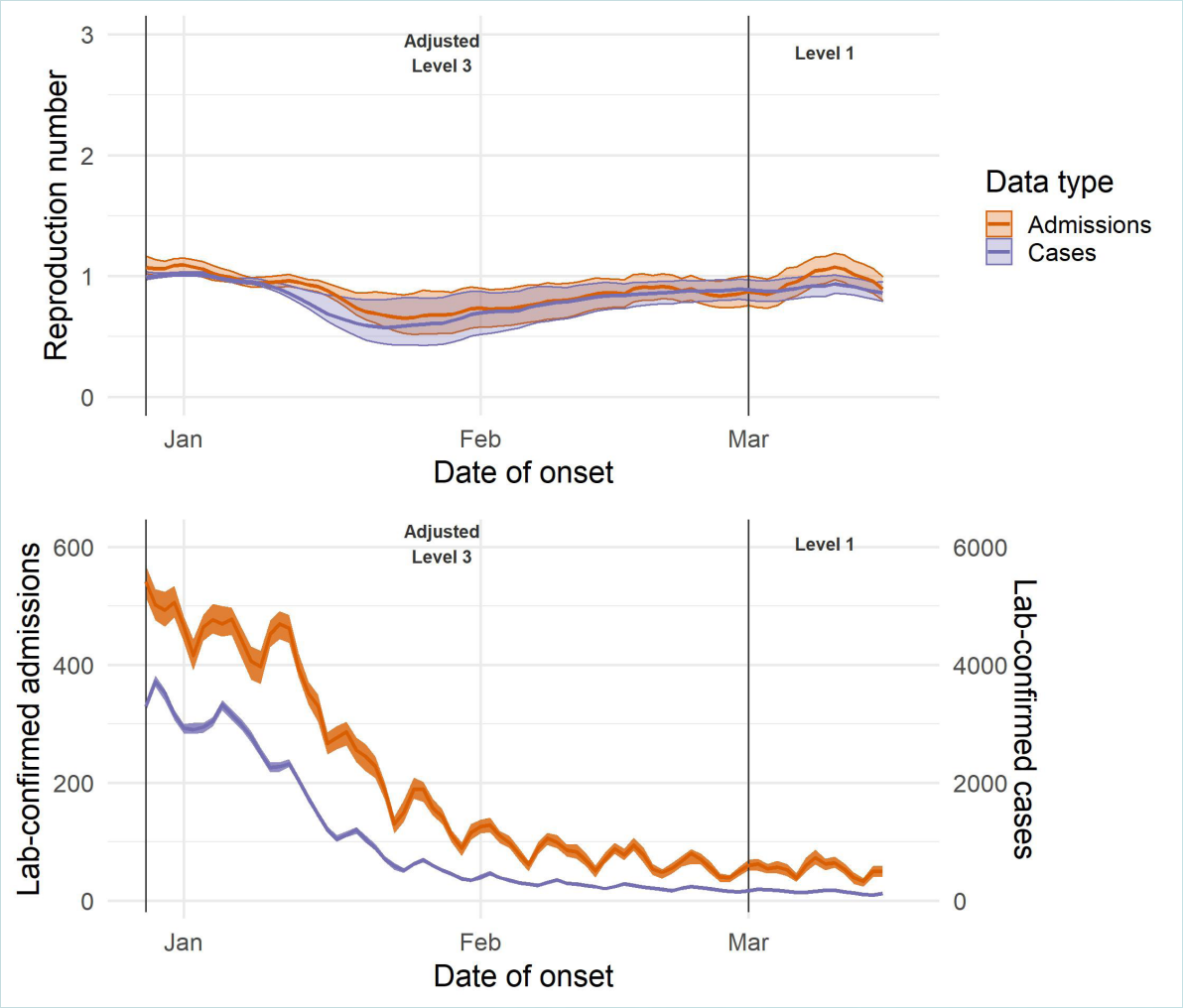

Figure 4. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Western Cape (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

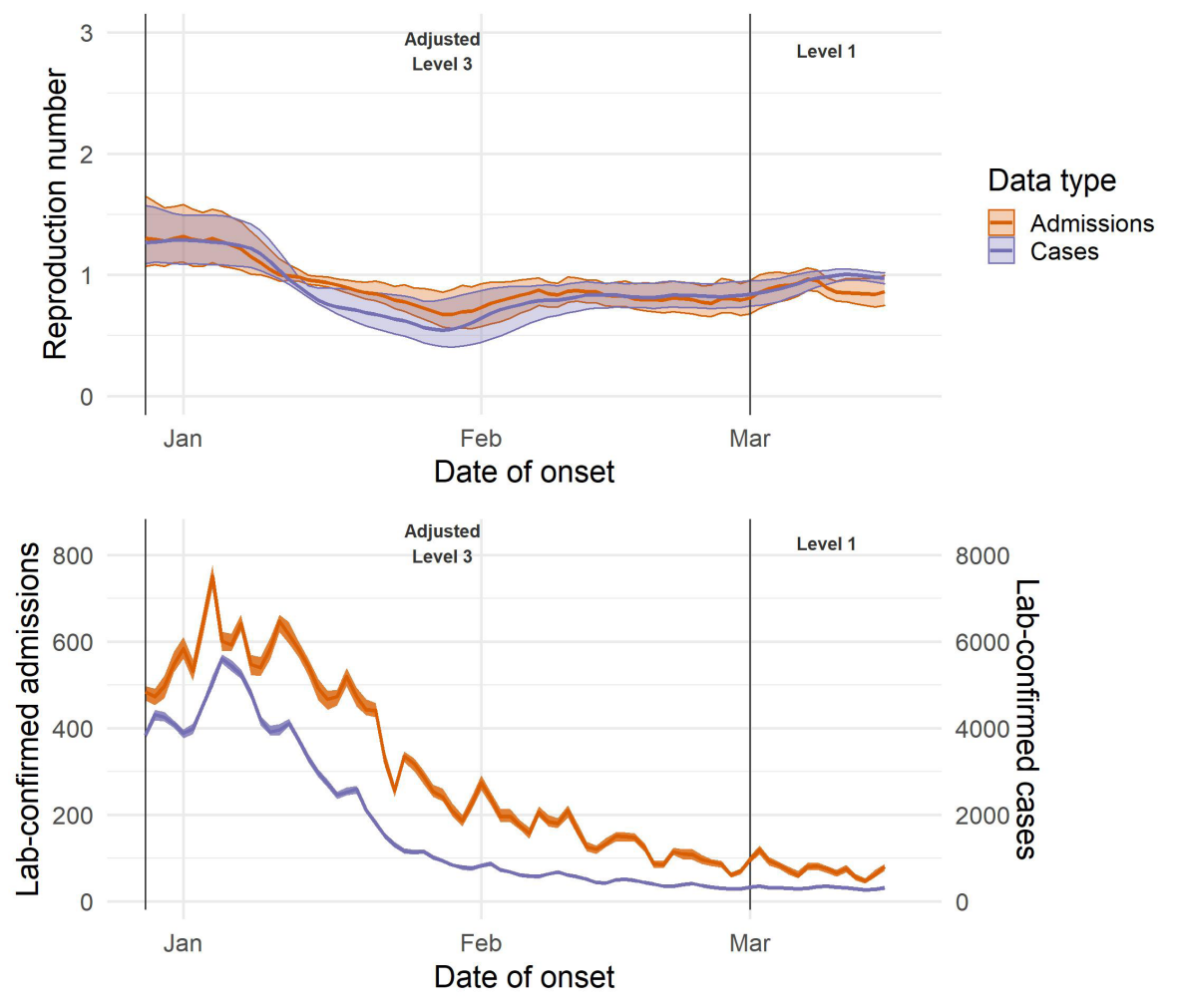

Figure 5. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals Gauteng (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

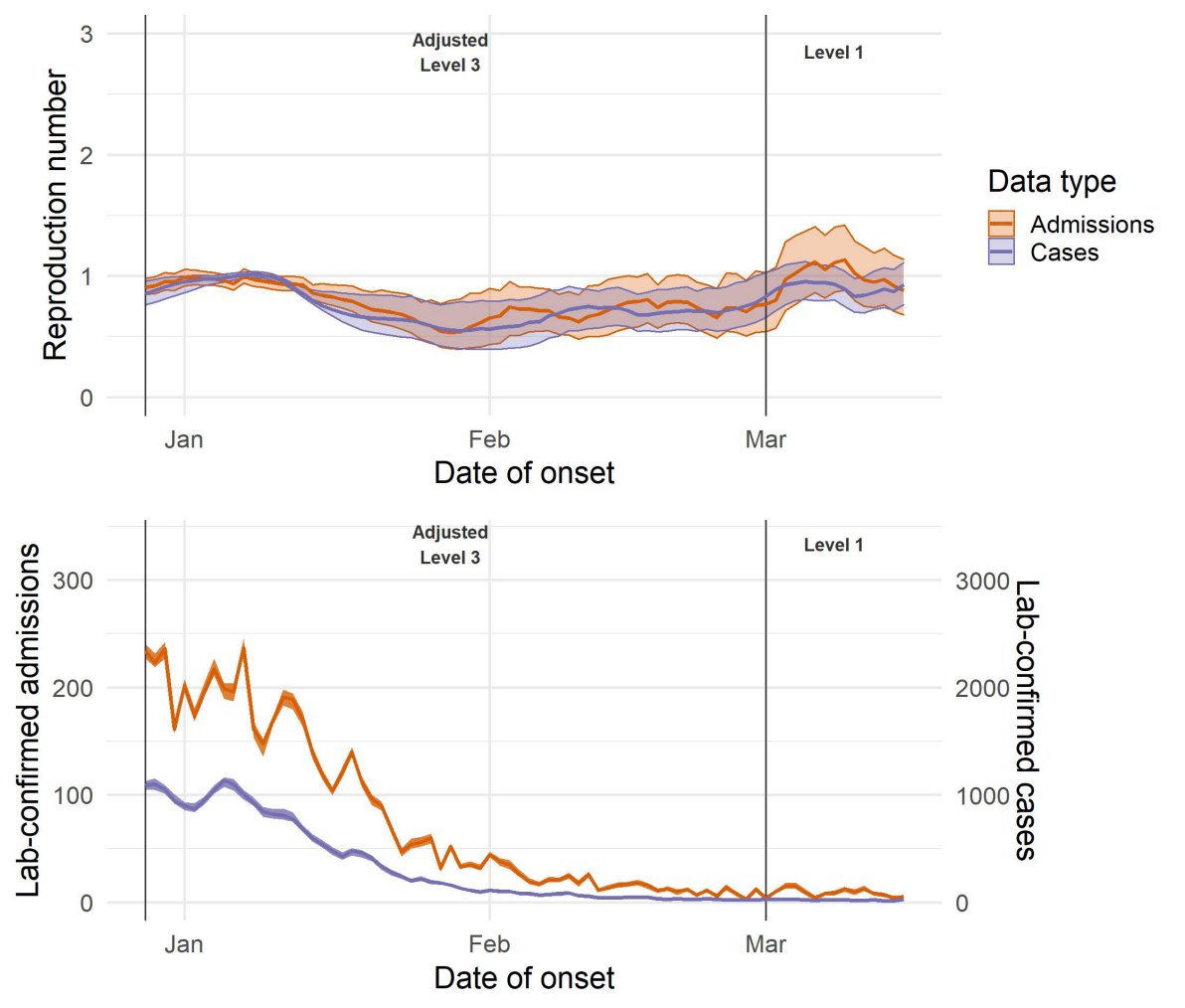

Figure 6. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Eastern Cape (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

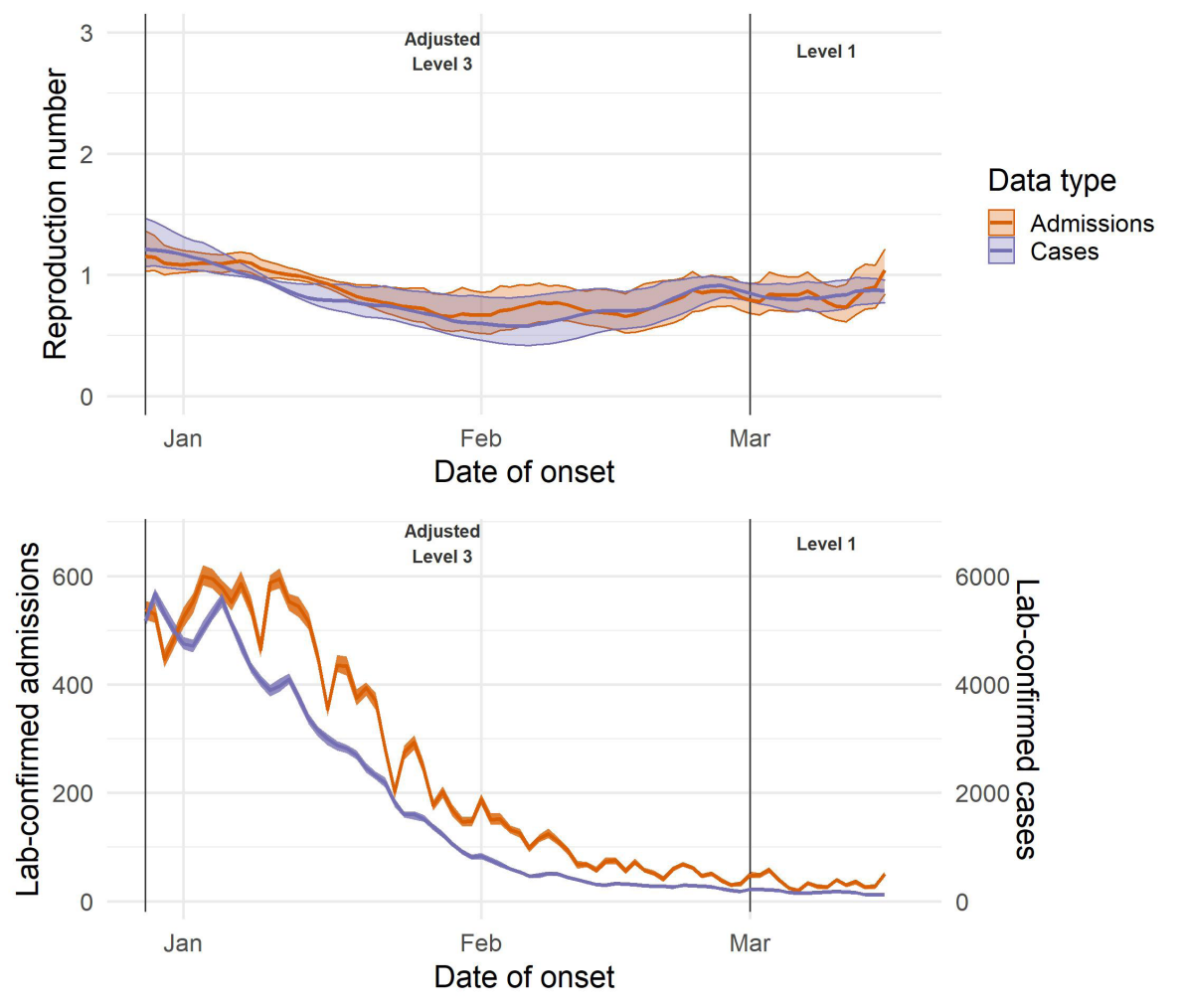

Figure 7. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, KwaZulu-Natal (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

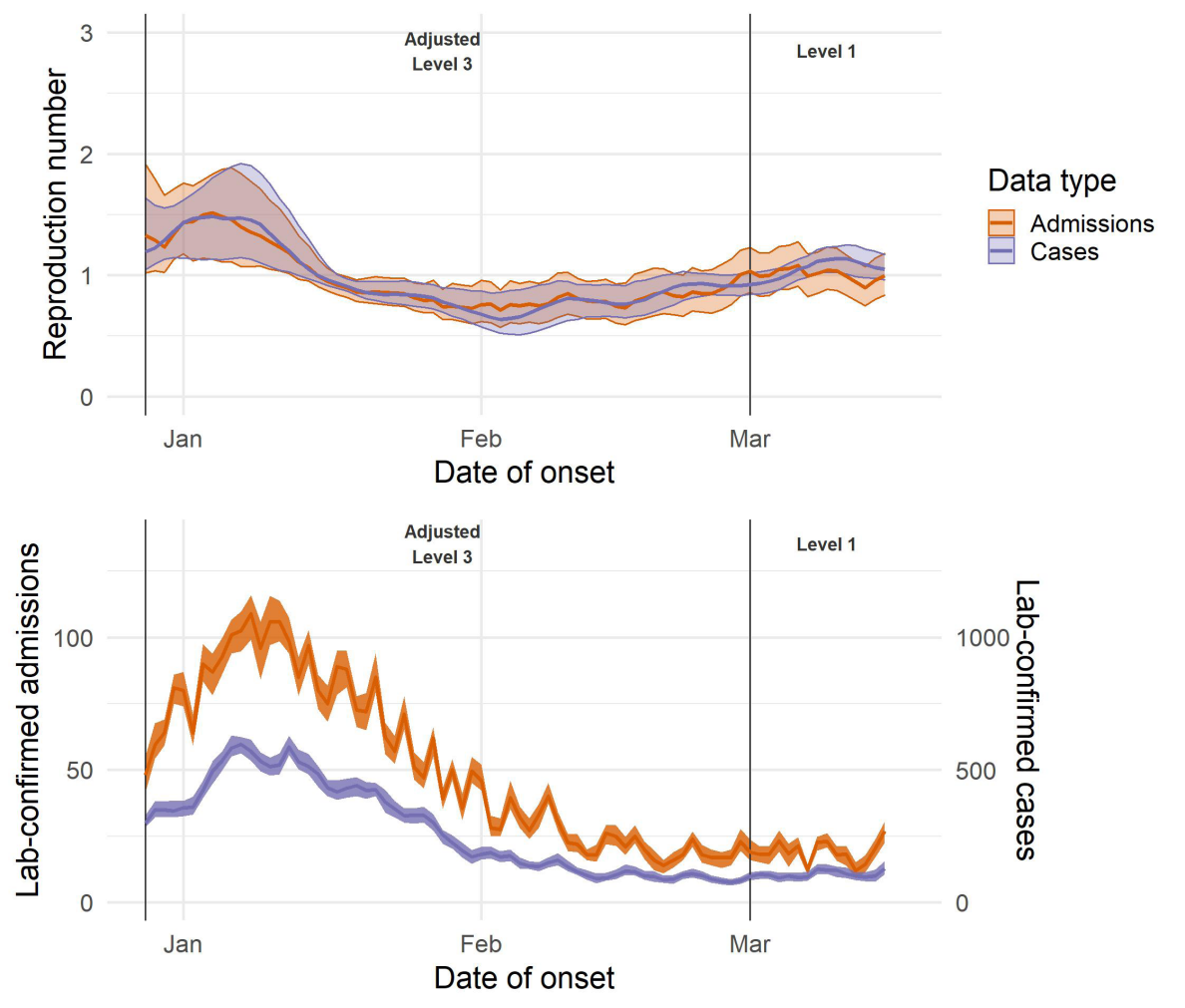

Figure 8. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Free State (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

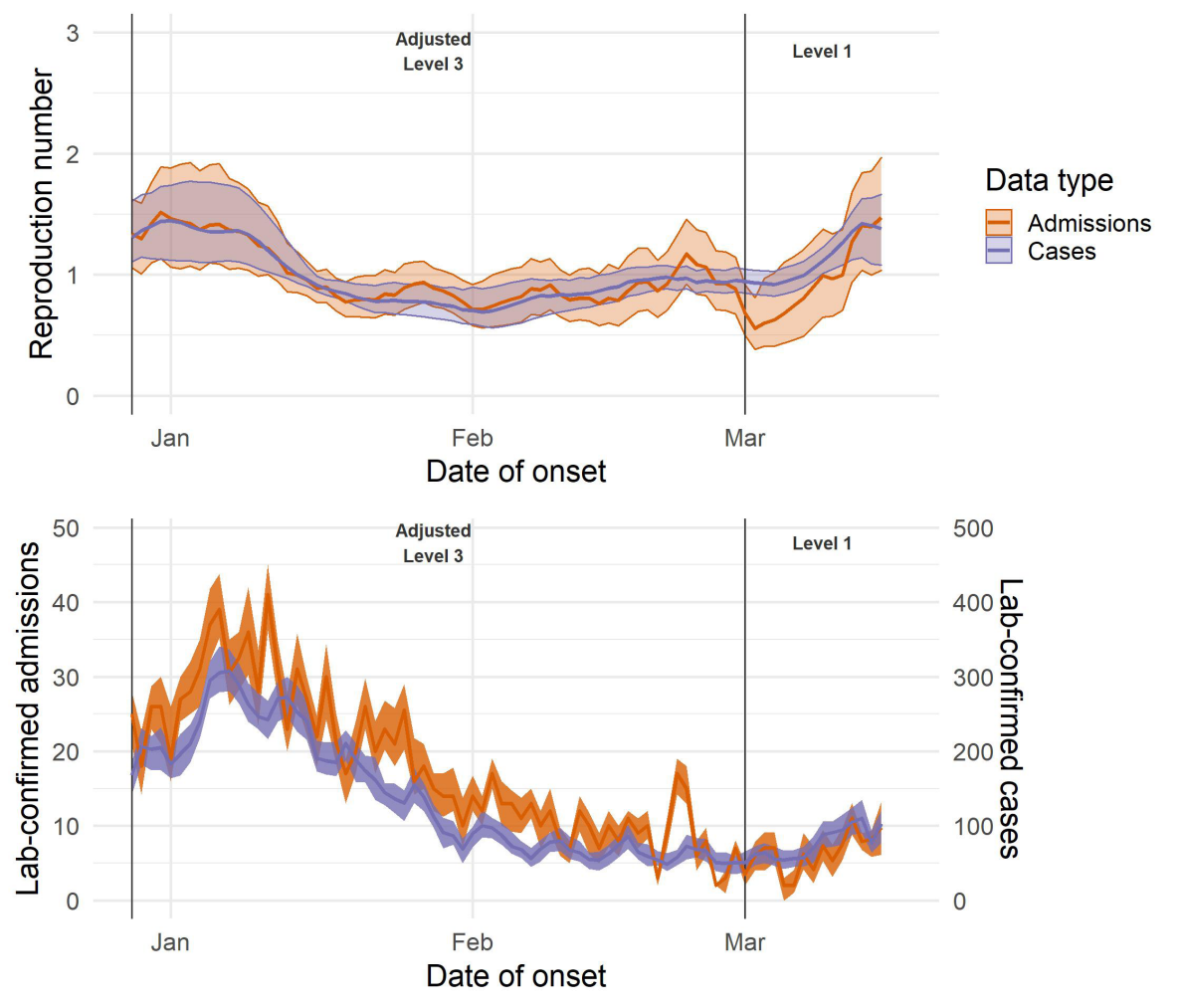

Figure 9. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Northern Cape (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

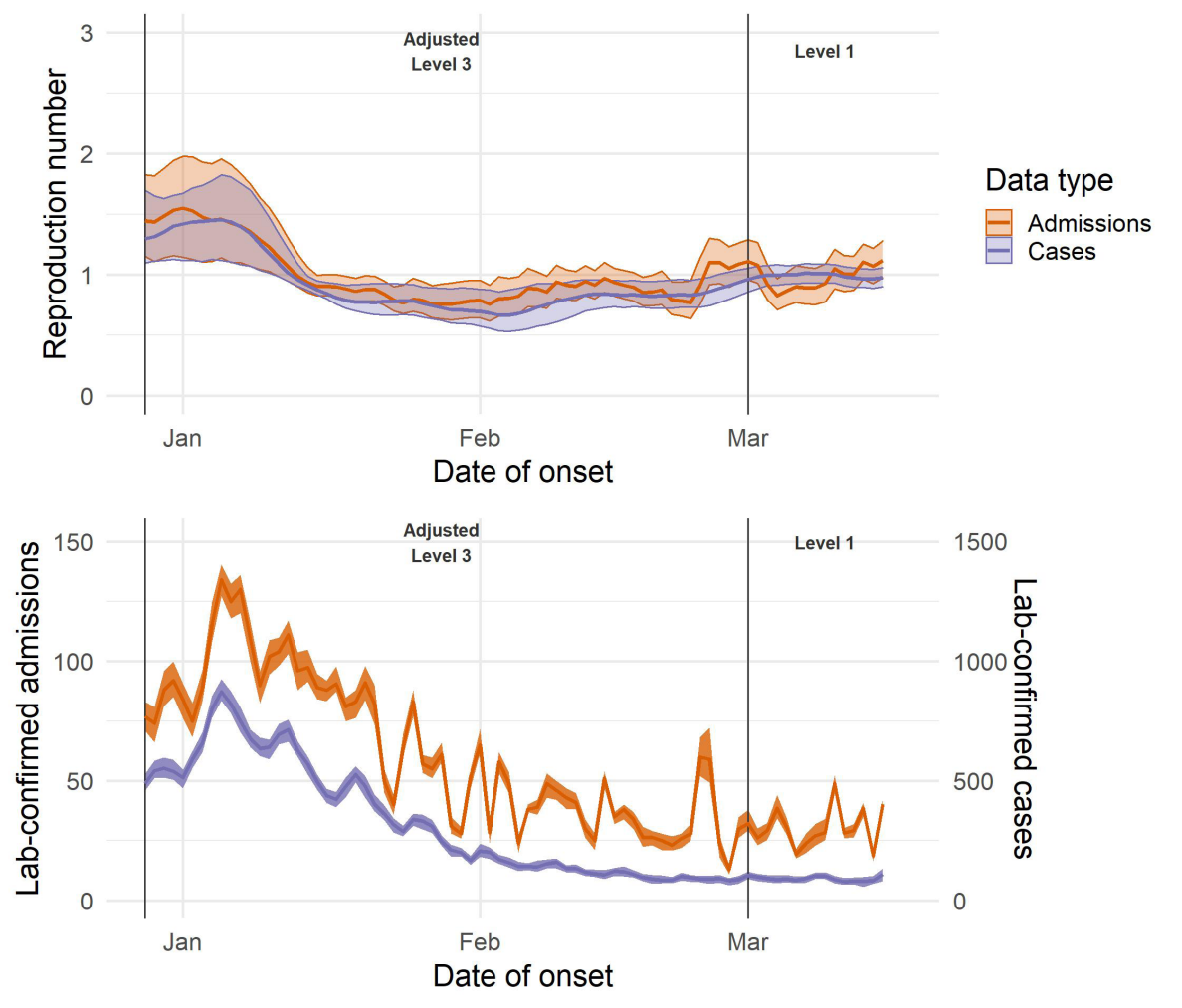

Figure 10. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, North West (last date included in the estimation: 15 March 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases and hospital admissions by onset date with missing data imputed. The medians and 95% ranges for the imputed datasets are shown. Daily numbers of confirmed laboratory-confirmed COVID-19 cases in the bottom panel are shown on the right-hand y axis.

Discussion

Nationally the average R remained initially steady during the adjusted level 3 lockdown, then dropped steadily for the first three weeks of January 2021. R stabilized briefly around the end of January, then gradually increased through the beginning of level 1 lockdown in early March with a value close to one at the end of the estimation period. The trends in all provinces were generally similar to the national trends with R near 1 at the end of the estimation period, except in Northern Cape where R exceeds 1. Some of the observed increases could be as a result of clusters of infections and numbers may reduce following implementation of control measures. The increasing trend in R indicates increasing transmission and may in part be driven by relaxation of controls to prevent the spread of COVID-19. It is essential that recommended measures to control the spread of COVID-19 are consistently implemented.

In general, the average daily R was similar between provinces and trends were similar to those nationally. Furthermore, in most cases, the three different endpoints used (cases, hospital admissions, and deaths) led to similar results when the numbers of cases, admissions, and deaths were sufficient.

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).

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, as well as by shifting care seeking behavior during the epidemic. The increased use of antigen tests through time, which have lower sensitivity than the more-commonly used PCR tests, may result in lower ascertainment rates, particularly during the second wave. 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. In addition, the relatively low numbers of deaths recorded between waves results in high levels of uncertainty and large fluctuation in R estimates based on daily deaths. We do not present death-based R estimates in province-level analyses due to instability of the estimates. Furthermore, a number of factors may have altered mortality outcomes over time, including the introduction of dexamethasone treatment in mid-June, the use of oxygen administration via high flow nasal cannula, changes in quality of healthcare provided if health systems are overwhelmed, and potential differences in severity between earlier circulating viruses and the 501Y.V2 / B.1.351 variant that dominated the second wave. Combined, these factors may lead to perturbations in the time series data that are unrelated to transmission. Comparing R estimates from the 3 data sources may help in assessing the severity of some of these biases, as indicated by inconsistent results across analyses of the three data sources. In addition to limitations in the ability of the available time series data to reflect underlying transmission, 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].

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

vii 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#