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 provinces where sufficient data are available. 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 29 November 2021 (week 48 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 27 November). This analysis updates the report released on 22 November 2021. In this report, R is estimated from the data on laboratory-confirmed COVID-19 cases and hospital admissions associated with public sector healthcare facilities and laboratory services. There may be non-overlapping sources of bias for the two 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 [vi]. Note: COVID-19 is the name of the disease and SARS-CoV-2 is the name of the virus.

Highlights

- Following a period of relatively stable R estimates, R has increased substantially during November, with national public-sector R estimates based on cases near 2.5 at the end of the estimation period.
- Province-level R estimates increased during November, with R based on cases close to or above 1 in all provinces by the end of the estimation period. In Gauteng, R based on cases reached values above 2.5 – comparable to the highest values reached in the first three waves.
- R estimates based on admissions increased during November, though less rapidly than R estimates based on cases. R estimates based on admissions were close to 1 in all provinces except the Western Cape and Mpumalanga at the end of the estimation period.

Methods

Daily R estimation

We used data from the first confirmed case in March 2020 until 29 November 2021, based on the national DATCOV dataset on hospitalized cases, and the laboratory-confirmed case line list maintained by the National Institute for Communicable Diseases (NICD). Due to incorrectly-entered reference dates for substantial numbers of antigen tests, only cases confirmed via polymerase chain reaction (PCR) testing were used in this analysis. Due to substantial differences in testing patterns between the public and private sectors (Figure 1), with public-sector test volumes remaining stable relative to private sector test volumes, this report focusses on R estimates based on admissions to public-sector hospitals and case confirmations from public-sector testing services. 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 53% of hospitalized cases, and 55% of fatal cases in the DATCOV dataset. 65 cases (0.02%) 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 based on lab-confirmed cases: 27 November 2021). 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. 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 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 2.5th-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 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 October 2021 (for more details regarding the COVID-19 lockdowns in South Africa, please refer to the South African government website [vi]).

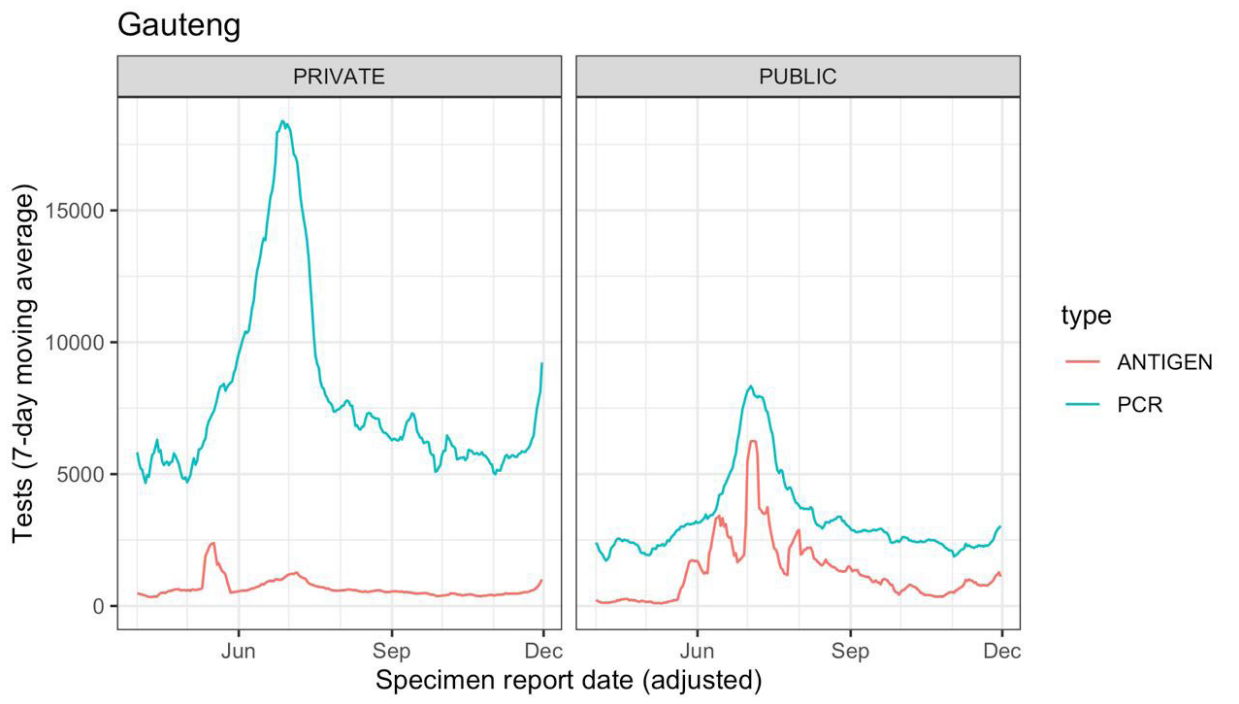

Figure 1. Reporting volumes for SARS-COV-2 diagnostic tests performed by private (left) and public (right) sector service providers in Gauteng province, South Africa. Specimen reporting dates for antigen tests were adjusted to account for substantial delays and incorrectly entered data during recent months.

Results

Nationally, public-sector R estimates remained relatively stable below 1 until early November, when R estimates based on cases and admissions began to increase. R estimates based on cases continued to increase through late-November, reaching a value above 2 at the end of the estimation period (Figure 2 and Table 1).

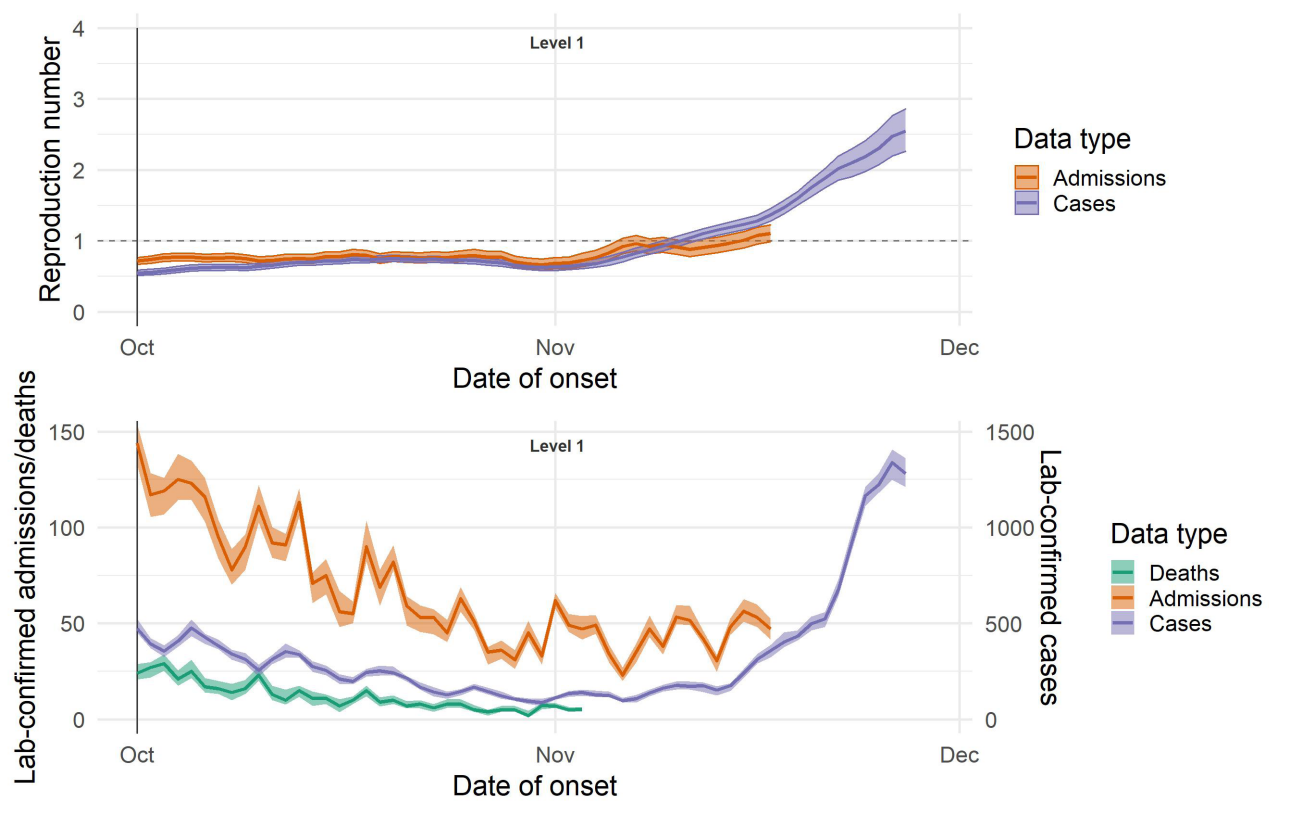

Figure 2. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, South Africa (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date 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.

Trends at the province level were similar, with public-sector R increasing during November (Figures 3-11). R based on cases was above one in all provinces at the end of the estimation period, and R estimates based on admissions were close one at the end of their respective estimation periods, except in the Western Cape and Mpumalanga, where R estimates based on admissions were below 1 (Table 1).

Table 1. Daily R estimates by province for R based on public-sector lab-confirmed cases and hospital admissions. Each cell contains median values with 95% confidence intervals.

	Cases	Cases	Admissions
	27 November 2021	17 November 2021	17 November 2021
National	2.55 (2.26,2.86)	1.37 (1.28,1.45)	1.10 (0.99,1.23)
Western Cape	1.63 (1.40,1.90)	0.77 (0.64,0.92)	0.69 (0.52,0.91)
Gauteng	3.06 (2.59,3.62)	2.49 (2.20,2.87)	1.87 (1.52,2.25)
Eastern Cape	1.25 (0.98,1.57)	1.11 (0.81,1.44)	0.67 (0.36,1.12)
KwaZulu-Natal	1.49 (1.27,1.69)	0.85 (0.71,1.01)	1.10 (0.79,1.43)
Free State	1.09 (0.89,1.34)	0.93 (0.75,1.13)	0.92 (0.65,1.27)
Northern Cape	1.02 (0.82,1.28)	1.03 (0.85,1.24)	1.34 (0.87,1.96)
North West	1.96 (1.69,2.25)	1.26 (1.02,1.54)	1.48 (1.03,2.07)
Mpumalanga	1.67 (1.40,1.96)	1.04 (0.82,1.27)	0.55 (0.29,0.94)
Limpopo	2.18 (1.52,2.92)	1.23 (0.65,2.08)	1.25 (0.68,2.09)

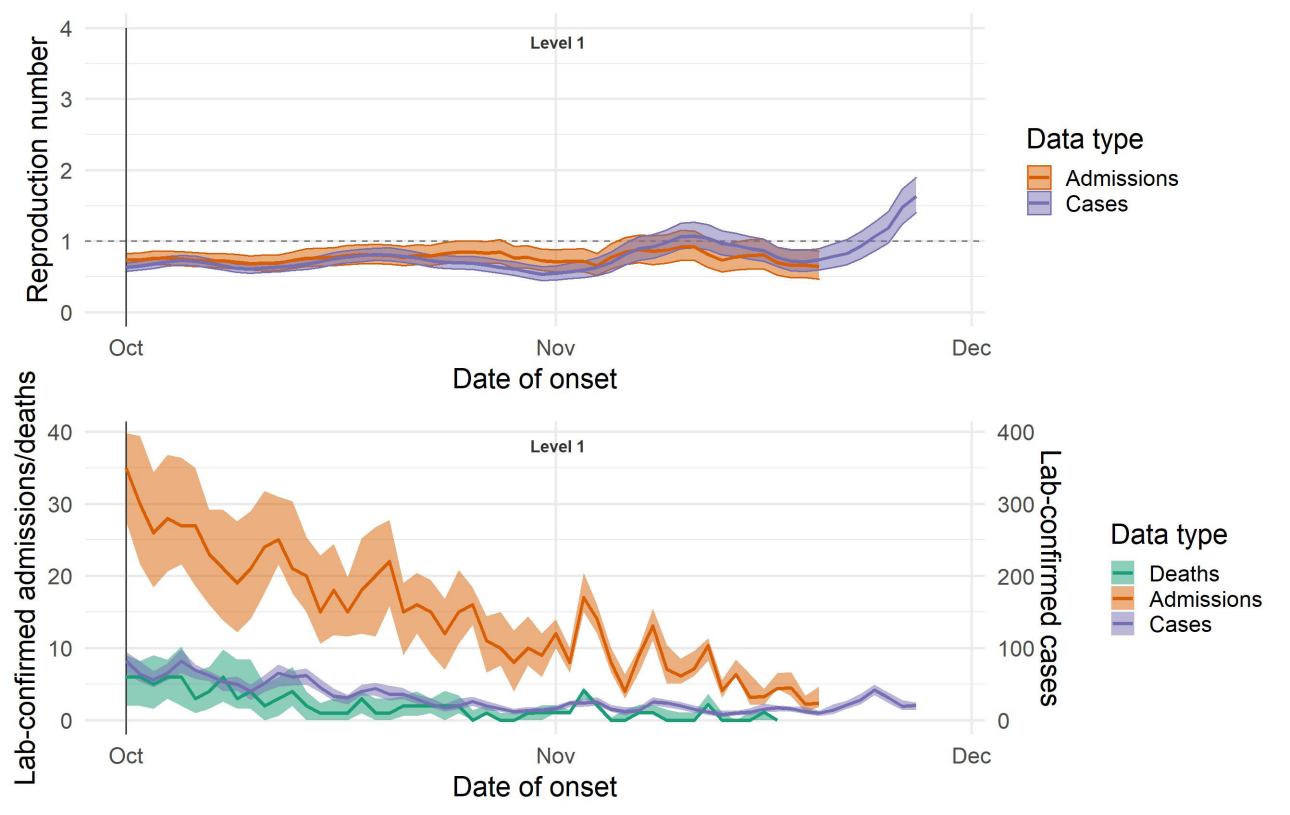

Figure 3. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Western Cape (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date 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.

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

WEEK **48** 2021

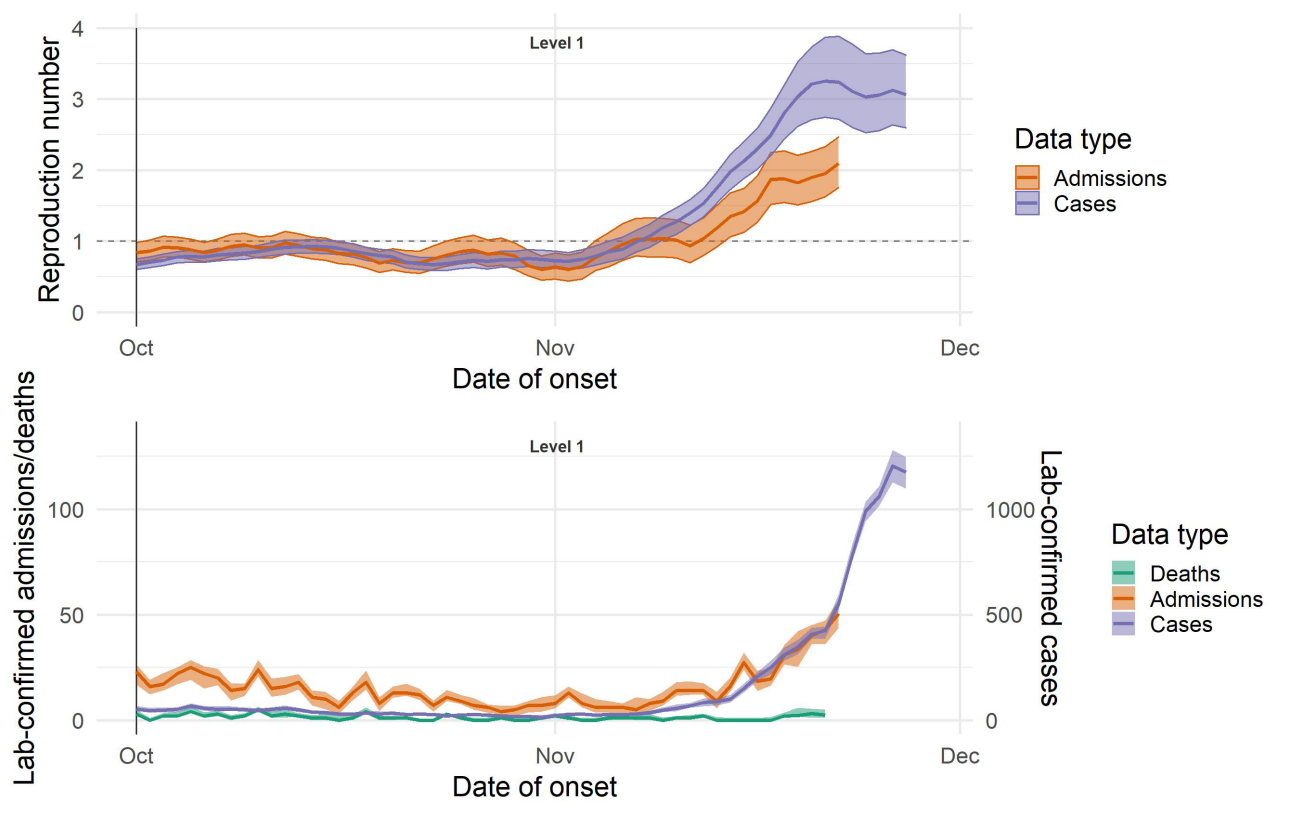

Figure 4. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals Gauteng (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with 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.

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

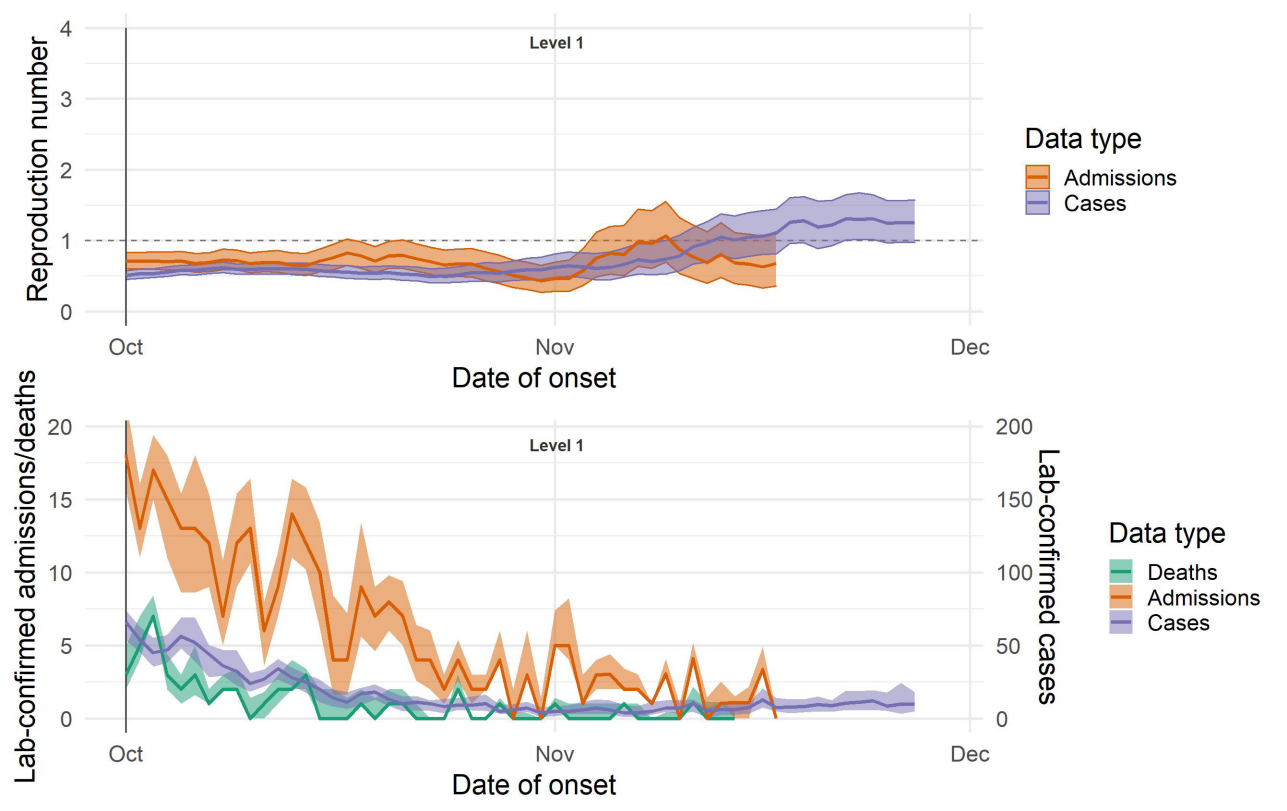

Figure 5. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Eastern Cape (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date 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.

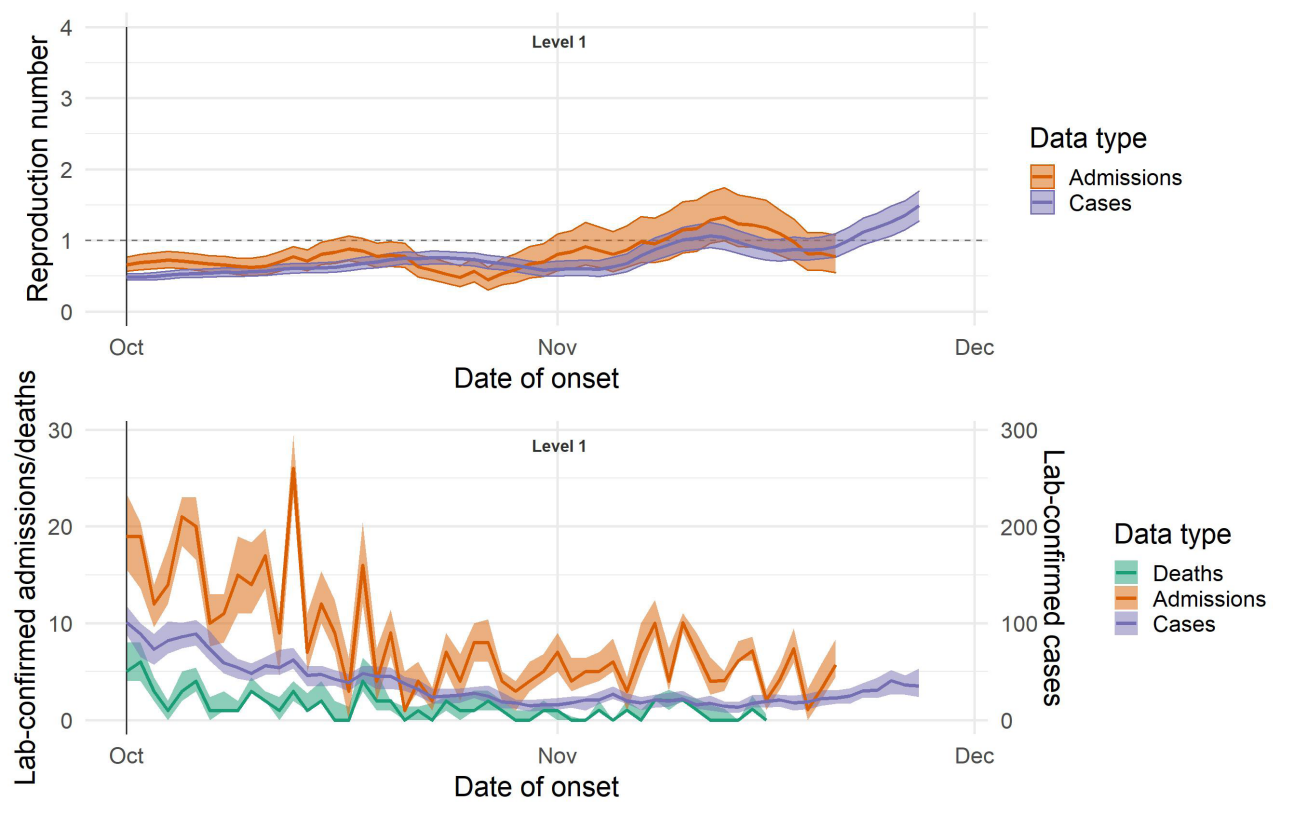

Figure 6. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, KwaZulu-Natal (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date 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.

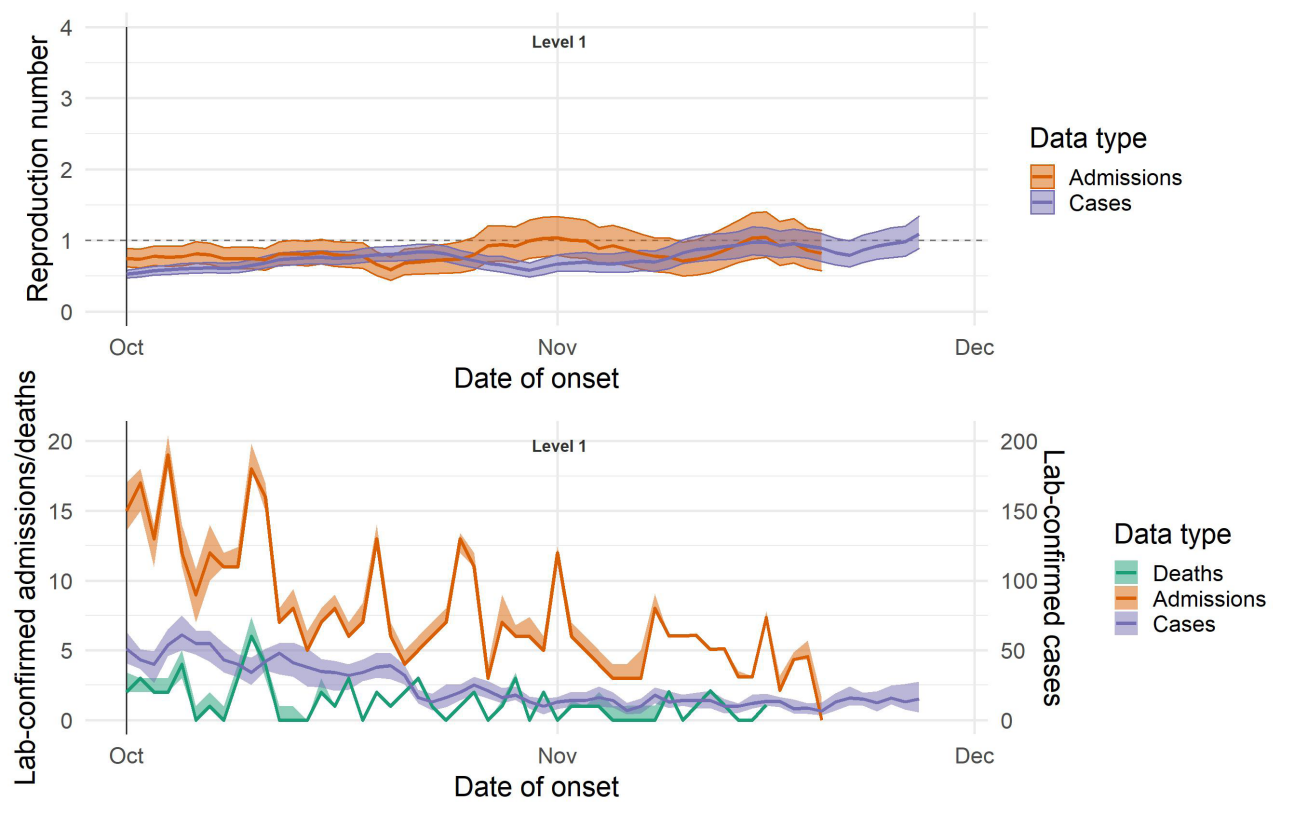

Figure 7. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Free State (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases. hospital admissions, and hospital-associated deaths by onset date with 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 ®, with 95% confidence intervals, Limpopo (last date included in the estimation: 29 November 2021). 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.

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

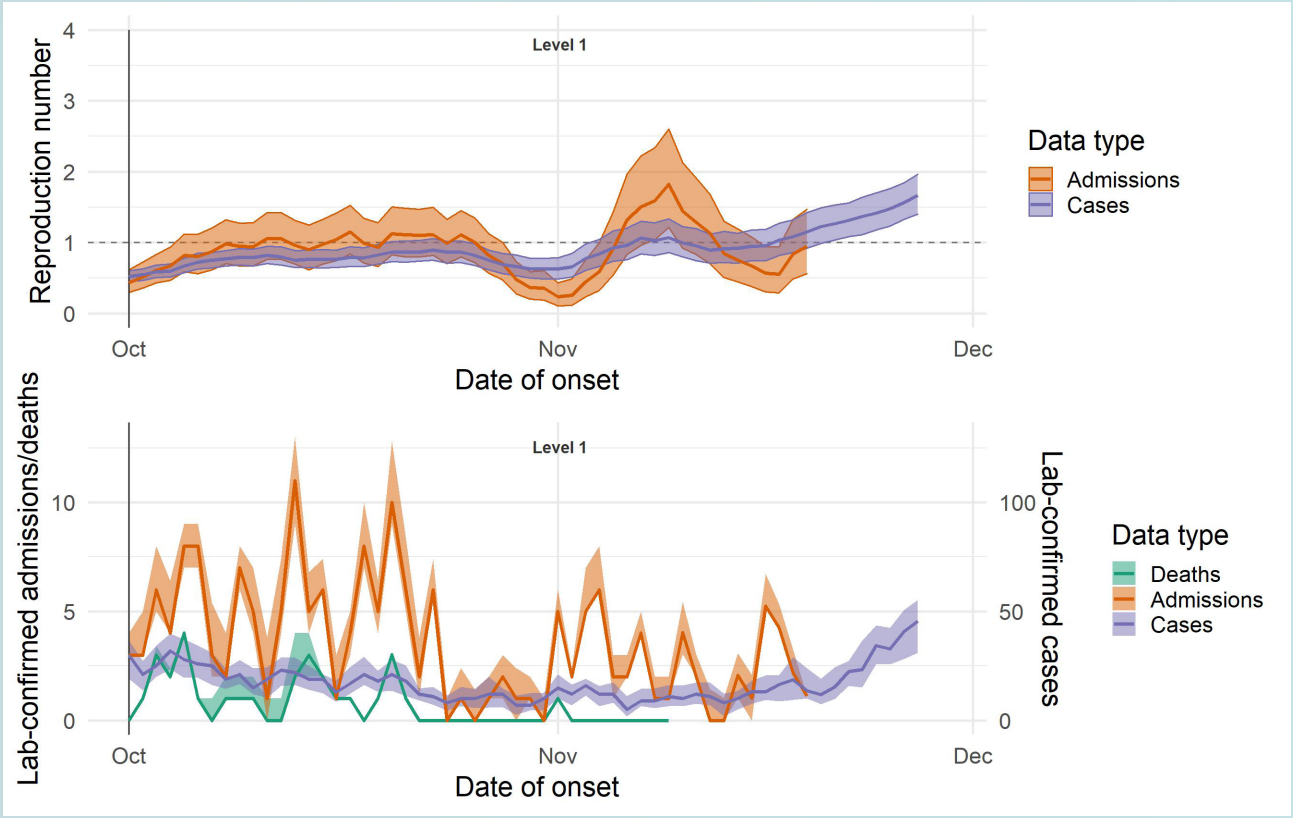

Figure 9. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, Mpumalanga (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date with 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.

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

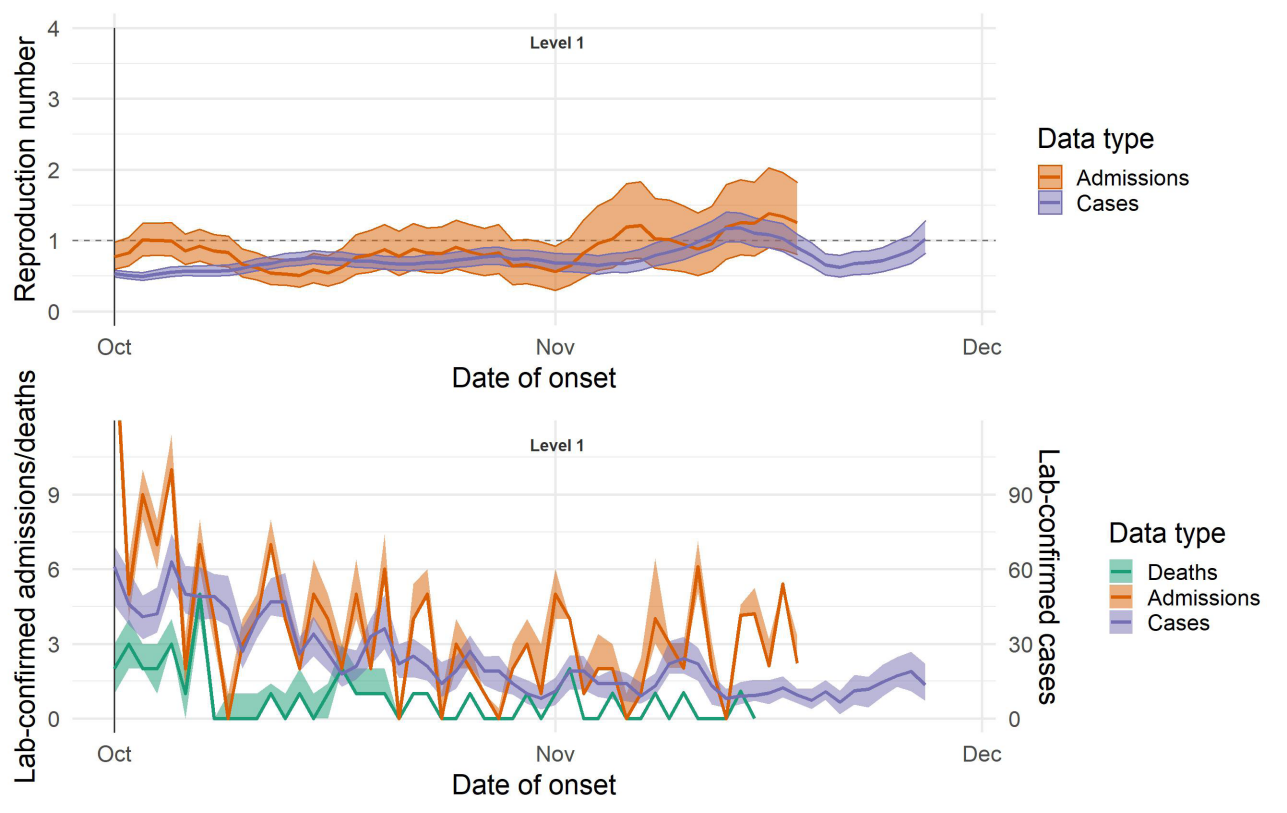

Figure 10. Upper panel: Estimated daily reproduction number ®, with 95% confidence intervals, Northern Cape (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date 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.

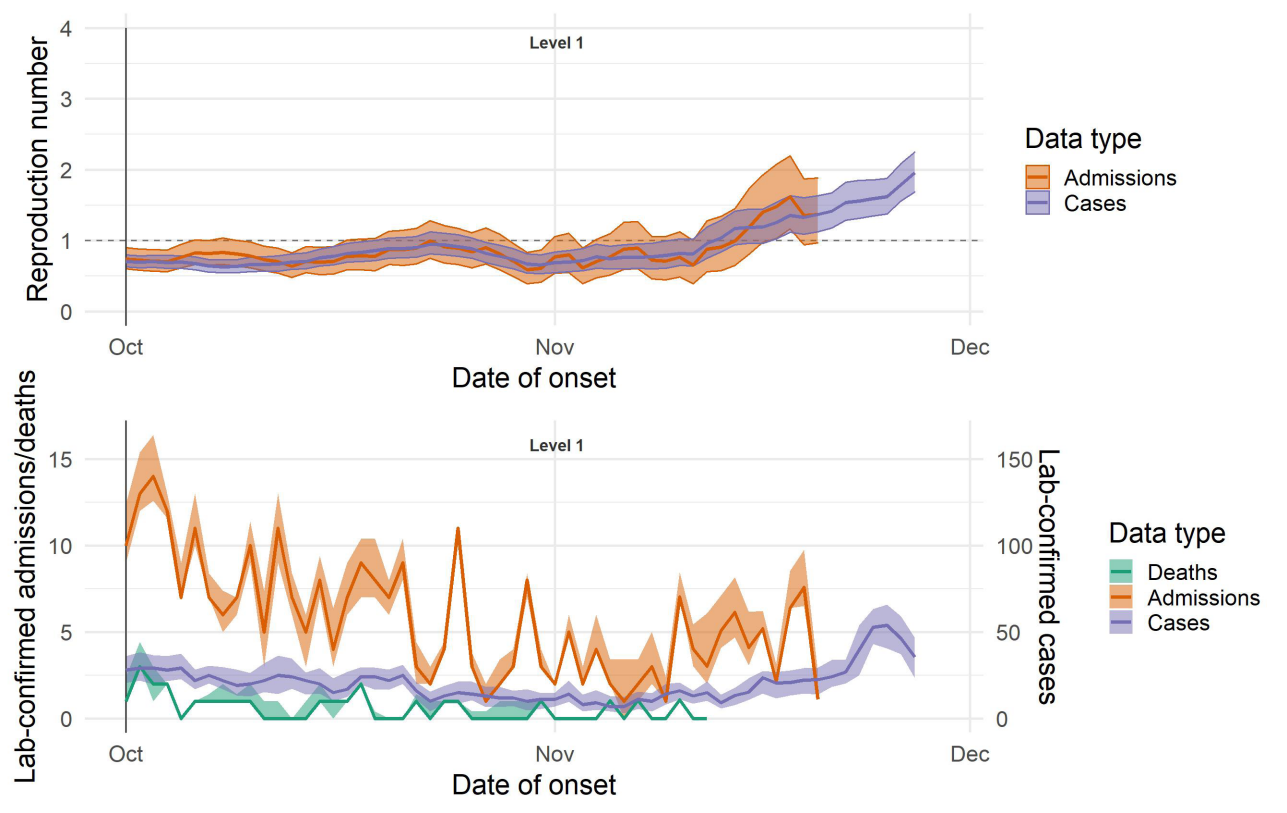

Figure 11. Upper panel: Estimated daily reproduction number (R), with 95% confidence intervals, North West (last date included in the estimation: 29 November 2021). Lower panel: estimated number of laboratory-confirmed COVID-19 cases, hospital admissions, and hospital-associated deaths by onset date 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.

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, and by increasing vaccination coverage.

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, and the Omicron (B.1.1.529) variant which is 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 two 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, the serial interval distribution is estimated based on a relatively small dataset from an on-going study.

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

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#

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