Estimating the case detection rate and temporal variation in transmission of COVID-19 in Australia

Technical Report 14th April 2020

David J. Price[1,2], Freya M. Shearer[2], Michael Meehan[3], Emma McBryde[3], Nick Golding[4], Jodie McVernon[1], James M. McCaw[1,2,5]

1. Victorian Infectious Diseases Laboratory Epidemiology Unit at The Peter Doherty Institute for Infection and Immunity; The University of Melbourne and Royal Melbourne Hospital, Melbourne, Australia
2. Modelling and Simulation Unit, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia
3. Australian Institute of Tropical Health & Medicine, James Cook University, Townsville, Australia
4. Telethon Kids Institute and Curtin University, Perth, Australia
5. School of Mathematics and Statistics, The University of Melbourne, Melbourne, Australia

Key messages

- We adapted and applied the method developed by colleagues at the London School of Hygiene and Tropical Medicine that uses the Case Fatality Rate in a region (adjusted for cases with known outcomes) to provide estimates of the *symptomatic* case detection rate in Australia. We note that LSHTM added Australia to their analysis on 1 April. The present authors have since updated the analysis, including the ability to estimate a time-dependent detection rate, at national level and for each state/territory.

- As of 9th April 2020, our estimate of the symptomatic case detection rate for Australia is 93% (95% CI 77–100%). The corresponding estimates for each state/territory are all greater than 80% (Figures 1 and 2).

- Analyses were performed to identify temporal changes in the *effective reproduction number* (R_{eff}) during the early course of the COVID-19 pandemic in each Australian state/territory.

- These analyses produced broadly consistent results showing that the effective reproduction number is likely less than 1 in NSW, VIC, QLD, SA, and WA as of 5 April 2020 (Figures 3–5). It should be noted that these estimates are averaged across the whole of each jurisdiction, and may reflect $R_{eff} >> 1$ in a number of localised settings and $R_{eff} << 1$ elsewhere.

- R_{eff} is estimated to be above 1 in TAS, which should be interpreted with caution given the small *cumulative* number of cases and the large *relative* increase in cases recently reported (32 cases reported between 10 and 12 April).

Estimating the symptomatic case detection rate

Symptomatic case detection rates were estimated by assuming a baseline clinical case fatality rate (CFR) of 1.38% (based on a large study conducted in China [1]). Regional CFR estimates are then compared with this value (1.38%/CFR) to estimate the proportion of symptomatic cases that have been identified in that region. Importantly, this method corrects for delays between case confirmation and death. Note that this method does not account for regional differences in age-structure or differential risks of severe outcomes across age groups compared to China (where the baseline CFR was estimated). The method is under continual development and revised estimates will continue to be provided. Figure 1 shows the most recent estimate for which data are deemed reliable (24 March). Figure 2 presents the time-dependent symptomatic case detection rate.

Figure 1: Estimates of the symptomatic case detection rate (light blue dots = mean; dark blue line = 50% CI; light blue lines = 95% CI) for each Australian state/territory using publicly available data up to 24 March 2020.

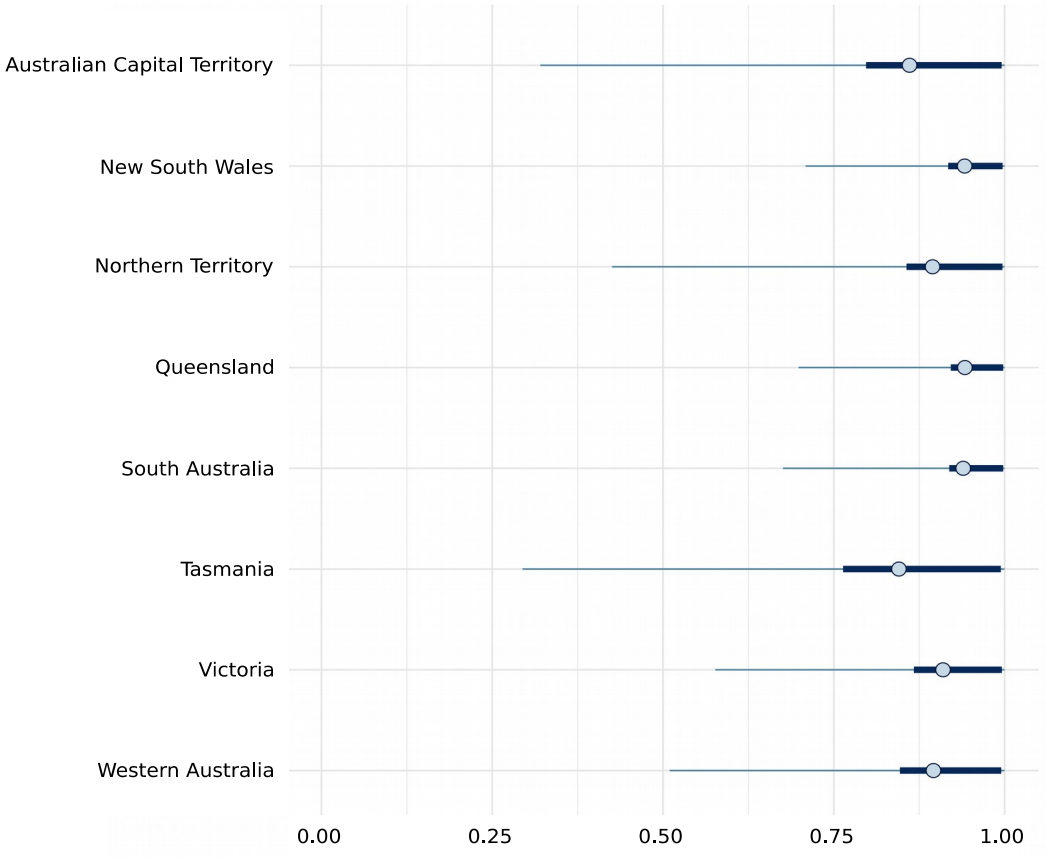
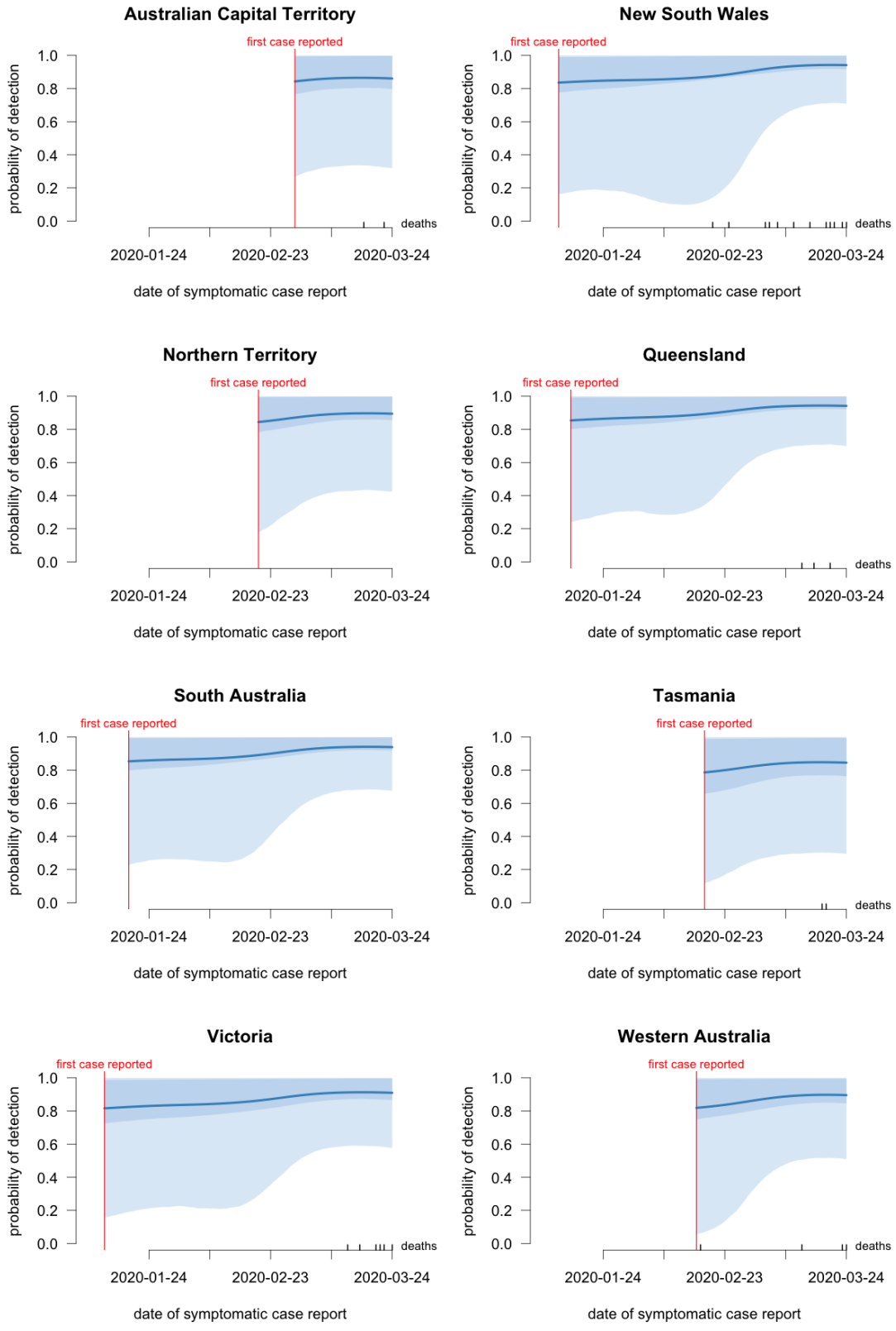


Figure 2: Time series estimates of the symptomatic case detection rate for each Australian state/territory using publicly available data up to 24 March 2020 (dark blue line = mean; shaded ribbons = 50% and 95% confidence intervals). The red vertical line indicates the date of the first report case and the timings of reported deaths are indicated by black ticks on the x-axis.



Estimating the temporal variation in the effective reproduction number in each jurisdiction

The effective reproduction number at any point in time provides a data-informed model-based estimate of the rate of change in case incidence. If $R_{eff} < 1$, then the epidemic is estimated to be in decline. If $R_{eff} > 1$, the epidemic is estimated to be growing.

Analysis 1: Using the statistical method developed by Abbott et al (2020) of LSHTM

Figure 3 presents estimates for the effective reproduction number for five Australian states.

Analysis 2: Exploring the time-varying effective reproduction number for different levels of relative infectiousness of local cases to imported cases

Figure 4 presents an analysis, using a slightly simplified method compared to that of Abbott et al (2020), but allowing for a difference in the infectiousness of local cases versus imported cases.

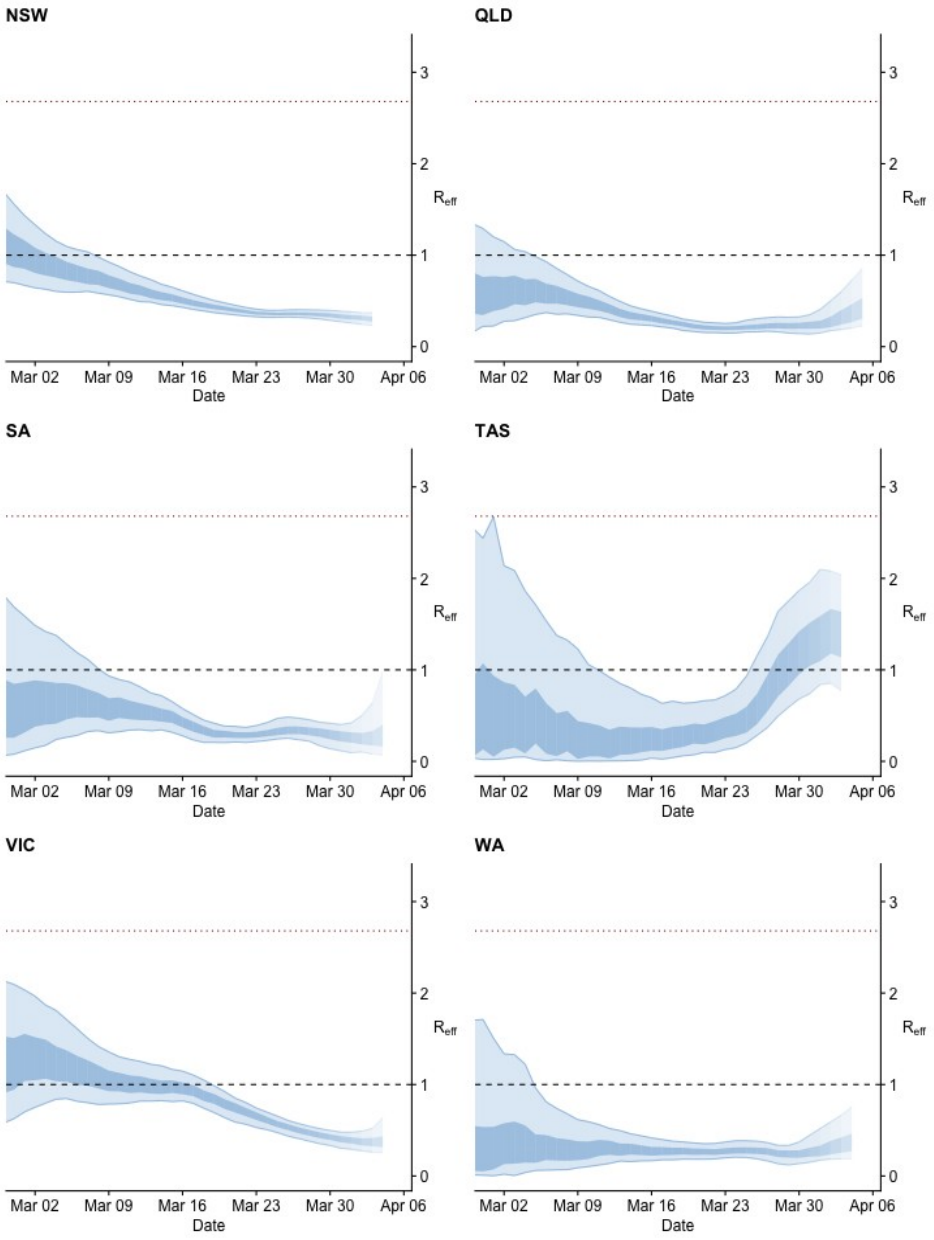
Analysis 3: Using statistical method developed by White and Pagano (2008)

Figure 5 provides an analysis using the older White and Pagano (2008) method.

Interpretation

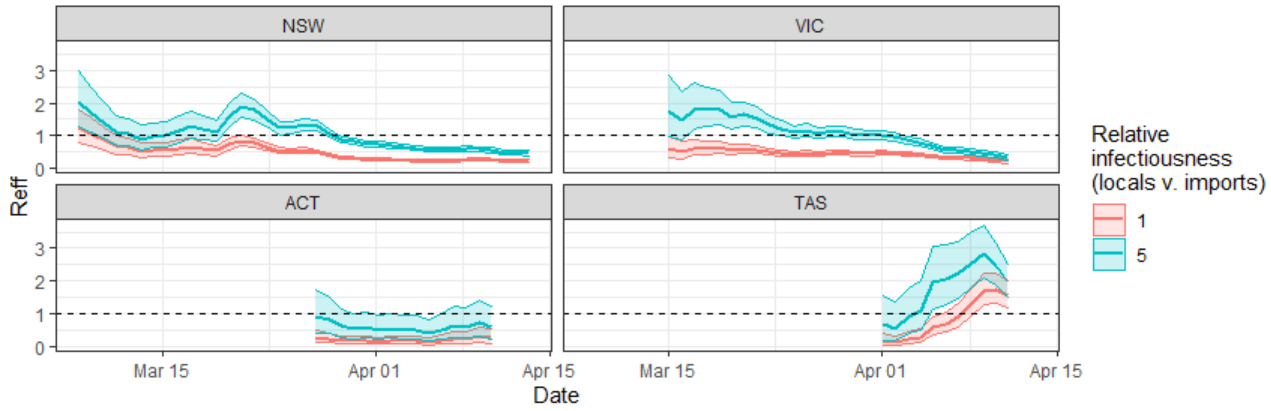
All three methods indicate that the COVID-19 epidemic in Australia is currently being suppressed sufficiently strongly to achieve *Reff<1*. If current measures were sustained indefinitely, and in the absence of imported cases or localised clusters, local elimination may be achievable. However, this effort could take many months, and would be dependent on the initial number of cases in each jurisdiction.

Figure 3: Time-varying estimate of the effective reproduction number of COVID-19 (light blue ribbon = 90% credible interval; dark blue ribbon = 50% credible interval) up to 5 April based on data up to and including 13 April, for each Australian state/territory with sufficient local transmission (excludes ACT, NT). Confidence in the estimated values is indicated by shading with reduced shading corresponding to reduced confidence. The black dotted line indicates the target value of 1 for the effective reproduction number required for control. The red dotted line indicates the reproduction number estimated for the early epidemic phase in Wuhan, China in the absence of public health interventions and assuming that the population was completely susceptible to infection (2.68).



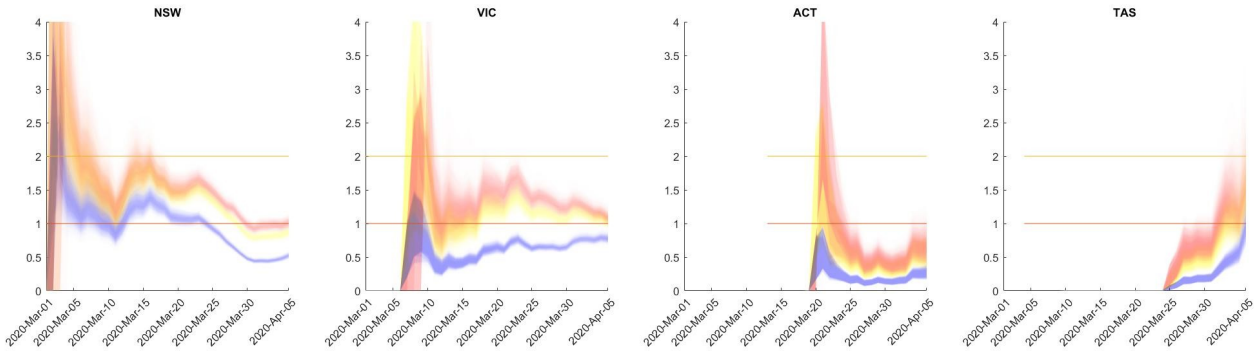
Note: Results for Tasmania should be interpreted with caution. Only a small number of cases have been reported in Tasmania to date, and so the results are very sensitive to small changes in case numbers since they are *relatively* large changes overall.

Figure 4: Time-varying estimates of the effective reproduction number of COVID-19 for each Australian state/territory (with sufficient publicly available data). Estimates were made using data up to and including 13 April for NSW, TAS and ACT, and up to and including 12 April for VIC. The relative infectiousness of locally acquired cases to imported cases are shown in red (equally infectious), green (local cases are five times more infectious). 95% credible intervals are indicated by the shaded ribbons. The black dotted line indicates the target value of 1 for the effective reproduction number required for control.



Note: Results for Tasmania should be interpreted with caution. Only a small number of cases have been reported in Tasmania to date, and so the results are very sensitive to small changes in case numbers since they are *relatively* large changes overall.

Figure 5: Time-varying estimates of the effective reproduction number of COVID-19 for each Australian state/territory with sufficient publicly available data. Estimates include data up to and including 7 April for NSW, TAS and ACT, and up to and including 6 April for VIC. 95% credible intervals are shown. The relative infectiousness of locally acquired cases to imported cases are shown in purple (equally infectious), yellow (local cases are three times more infectious) and orange (local cases are five times more infectious). The solid orange line indicates the target value of 1 for the effective reproduction number required for control.



Note: All cases listed as “under investigation” for import status were treated as locally acquired cases in this analysis. Tasmania currently has a low number of cases which were recently reported, and thus a high proportion remain “under investigation” for import status. As such, the apparent escalating epidemic in Tasmania should be interpreted with caution.

Overview of methodology

Background

Disease transmissibility can be characterised by the *effective* reproduction number (R_{eff}) — the average number of secondary infections caused by an infected individual in the presence of public health interventions (and with no assumption of a fully-susceptible population). If control efforts are able to bring R_{eff} below 1, we will see a decline in the number of new cases on average. Monitoring changes in R_{eff} over time is critical for assessing the impact of our overall response strategy to COVID-19.

Analysis 1: Estimates of time-varying effective reproduction number using a statistical method developed by Abbott et al (2020) of LSHTM

We use COVID-19 case counts from the COVID-19 NNDSS database, stratified by import status (imported from overseas or locally acquired), for each Australian state/territory up to and including 13 April 2020, to estimate R_{eff} over time from 1 March up to 5 April (Figure 3).

We use a statistical method developed by colleagues at the London School of Hygiene and Tropical Medicine (LSHTM), recently adapted for COVID-19, which builds on their extensive experience and peer-reviewed work in this area. This work is currently being developed and shared rapidly with the international community to inform situational assessment.

This method estimates R_{eff} by using a 7-day moving average window to smooth the curve and reduce the impact of localised events (*e.g.*, local outbreaks) that may cause large fluctuations. Importantly, the method accounts for delays in reporting (*i.e.*, the time from symptom onset to reporting) which is critical for incorporating the most recent data in the analysis (*i.e.*, for inferring when an observed drop in the number of reported cases reflects an actual drop in case numbers).

Note that up to 20% of reported cases in the national database do not have a reported import status (see Figure S1 in the Appendix). For the purpose of this analysis, we have assumed that all cases with unknown or unconfirmed source of acquisition are locally acquired.

Further, the estimated time-varying R_{eff} value is based on cases that have been identified as a result of local transmission, whereas imported cases are managed separately.

Analysis 2: Exploring the time-varying effective reproduction number for different levels of relative infectiousness of local cases to imported cases

We use publicly available COVID-19 case counts (www.covid19data.com.au), stratified by import status (imported from overseas or locally acquired), for each Australian state/territory up to and including 14 April, to estimate R_{eff} over time. We use a similar statistical approach to Analysis 1, which allows for varying levels of infectiousness between imported and local cases (Figure 5). Note that this approach does not account for delays in reporting and uses case notification dates rather than dates of symptom onset as the proxy for infection date.

It is possible that known imported cases may be less infectious than locally acquired cases, for example due to quarantine recommendations for incoming travellers. This analysis explores the impact of various levels of relative infectiousness of locally acquired cases to imported cases on the time-varying effective reproduction number.

Note that for VIC, data were only available up to and including 12 April. For all states/territories, cases classified as "under investigation" for import status were excluded from the analysis.

Note that both Analyses 1 and 2 assume a constant detection proportion throughout the time period.

Analysis 3: Estimates of time-varying effective reproduction number using a statistical method developed by White and Pagano (2008)

Once again, we use publicly available COVID-19 case counts (www.covid19data.com.au), stratified by import status (imported from overseas or locally acquired) to estimate R_{eff} over time for each Australian state/territory. Data up to and including 7 April was included for NSW, TAS and ACT, and up to and including 6 April for VIC.

We use a statistical method developed by White and Pagano (2008) [6], adapted to account for the contribution of imported cases to transmission. Note that this method does not account for delays in reporting and uses case notification dates rather than dates of symptom onset as the proxy for infection date.

Similar to Analysis 2, we explore the impact of different levels of relative infectiousness of locally acquired cases to imported cases on the time-varying effective reproduction number. As with Analysis 1, we have assumed that all cases with unknown or unconfirmed source of acquisition are locally acquired.

Technical Appendix

Estimating the symptomatic case detection rate

This analysis is based on the method developed by Russell et al (2020) of the London School of Hygiene and Tropical Medicine, Centre for Mathematical Modelling of Infectious Diseases novel coronavirus working group. A member of the local Australian team is a direct contributor to this project.

Full details of their statistical analysis and code base are available via their website (below).

https://cmmid.github.io/topics/covid19/severity/global_cfr_estimates.html

Estimating the temporal variation in the effective reproduction number in each jurisdiction

Analysis 1: Estimates of time-varying effective reproduction number using a statistical method developed by Abbott et al (2020) of LSHTM

Our analysis is based on the method developed by Abbott et al (2020) of the London School of Hygiene and Tropical Medicine, Centre for Mathematical Modelling of Infectious Diseases novel coronavirus working group.

Full details of their statistical analysis and code base are available via their website (below) and described in the key references at the end of this document [2,3].

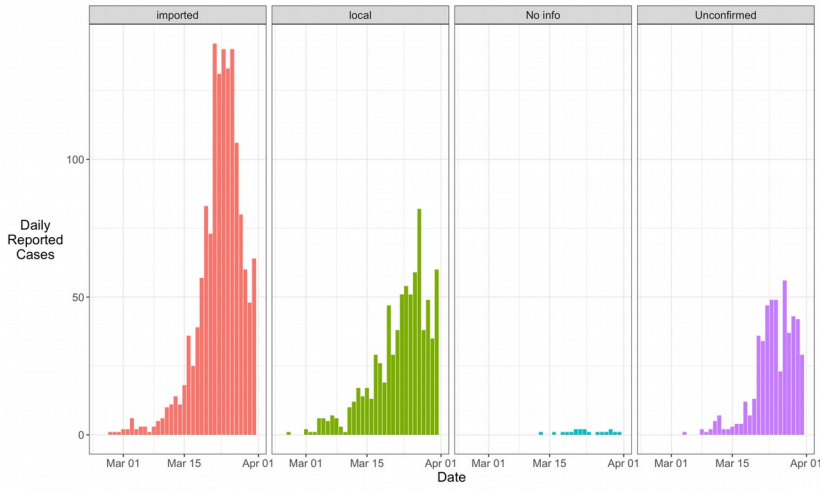
https://epiforecasts.io/covid/

We provide a brief overview of the method below, focusing on how the analysis was adapted to the Australian context.

Data

We used line-lists of reported cases for each state/territory from the COVID-19 NNDSS database, accessed via an agreement with the Commonwealth of Australia Department of Health. The line-lists contain the date when the case first exhibited symptoms, date when the case notification was received by the jurisdictional health department and where the case was acquired.

Figure S1: NNDSS number of reported cases (aggregated across states/territories) by import status classification. Note that all "no info" and "unconfirmed" cases were assumed to be local for the purposes of this analysis.



Adjusting for reporting delays
A *pre-hoc* statistical analysis was conducted in order to estimate a distribution of the reporting delays from the line-lists of cases, using the code base provided by Abbott et al 2020. The estimated reporting delay is assumed to remain constant over time. These reporting delays are used to: i) infer the time of symptom onset for those without this information, and; ii) infer how many cases in recent days are yet to be recorded. Adjusting for reporting delays is critical for inferring when a drop in reported cases reflects a true drop in cases.

Onset dates are estimated for individuals without one recorded (~8%). Onset dates are typically missing for most recent data, and reporting delays may be changing throughout the epidemic. These data may not be representative of the underlying delay distribution.

Trends identified using this approach are robust to under-reporting assuming it is constant, but absolute values may be biased by reporting rates. Pronounced changes in reporting rates may also impact the trends identified. However, evidence shown in Figures 1 and 2 suggest that Australia is capturing a very high proportion of symptomatic cases.

Estimating the effective reproduction number over time
Briefly, the effective reproduction number was estimated for each day from 1 March 2020 up to 5 April 2020 using line list data – date of symptom onset, date of report, and import status – for each state. The method assumes the serial interval (i.e., time between symptom onset for an index and secondary case) is uncertain, with a mean of 4.7 days (95% CrI: 3.7, 6.0) and a standard deviation of 2.9 days (95% CrI: 1.9, 4.9), as estimated from early outbreak data in Wuhan [4]. Combining the incidence over time with the uncertain distribution of serial intervals allows us to estimate R_{eff} over time. The underlying statistical methodology is extensively detailed in [2].
A prior distribution was specified for R_{eff}, with mean 2.6 (informed by [5]) and a broad standard deviation of 2 so as to allow for a range of R_{eff} values.

R_{eff} is estimated using a 7-day moving average window in order to smooth the curve and reduce the impact of localised events (*e.g.*, local outbreaks) causing large variations.

Note that up to 20% of reported cases in the national database do not have a reported import status (see Figure S1). For the purpose of this analysis, we have assumed that all cases with unknown or unconfirmed source of acquisition are locally acquired.

Accounting for imported cases
A large proportion of cases reported in Australia from January until now were imported from overseas. It is critical to account for two distinct populations in the case notification data — imported and locally acquired — in order to perform robust analyses of transmission in the early stages of this outbreak.

The estimated time-varying R_{eff} value is based on cases that have been identified as a result of local transmission, whereas imported cases are managed separately.

Analysis 2: Exploring the time-varying effective reproduction number for different levels of relative infectiousness of local cases to imported cases

We use publicly available COVID-19 case counts (www.covid19data.com.au), stratified by import status (imported from overseas or locally acquired), for each Australian state/territory up to and including 14 April, to estimate R_{eff} over time. We use a similar statistical approach to Analysis 1, which allows for varying levels of infectiousness between imported and local cases (Figure 5). Note

that this approach does not account for delays in reporting and uses case notification dates rather than dates of symptom onset as the proxy for infection date.

It is possible that known imported cases may be less infectious than locally acquired cases, for example due to quarantine recommendations for incoming travellers. This analysis explores the impact of various levels of relative infectiousness of locally acquired cases to imported cases on the time-varying effective reproduction number. Note that for VIC, data were only available up to and including 6 April. For all states/territories, cases classified as "under investigation" for import status were excluded from the analysis.

Estimates were made from when the first local case was reported in each jurisdiction, with the initial outbreak size assumed to be the sum of imported cases reported over the previous five days (*i.e.*, the approximate serial interval). Early estimates are highly sensitive to these initial conditions.

We assumed a Gamma-distributed serial interval with mean = 4.7 and standard deviation = 2.9. These are the same assumptions used in Analysis 1.

Analysis 3: Estimates of time-varying effective reproduction number using a statistical method developed by White and Pagano (2008)

Once again, we use publicly available COVID-19 case counts (www.covid19data.com.au), stratified by import status (imported from overseas or locally acquired) to estimate R_{eff} over time for each Australian state/territory. Data up to and including 7 April was included for NSW, TAS and ACT, and up to and including 6 April for VIC.

We use a statistical method developed by White and Pagano (2008) [6], adapted to account for the contribution of imported cases to transmission. Note that this method does not account for delays in reporting and uses case notification dates rather than dates of symptom onset as the proxy for infection date.

Similar to Analysis 2, we explore the impact of different levels of relative infectiousness of locally acquired cases to imported cases on the time-varying effective reproduction number. As with Analysis 1, we have assumed that all cases with unknown or unconfirmed source of acquisition are locally acquired.

Full details of the statistical analysis are described in the key reference at the end of this document [6].

References

1. Verity R, Okell LC, Dorigatti I *et al*. Estimates of the severity of covid-19 disease. *medRxiv* 2020
2. Cori A, Ferguson NM, Fraser C *et al*. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology* 2013; **178**: 1505–12. DOI: 10.1093/aje/kwt133
3. Wallinga J, Teunis P. Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. *American Journal of Epidemiology* 2004; **160**: 509–16. DOI: 10.1093/aje/kwh255
4. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis* 2020; **4**(93):284-286. DOI: 10.1016/j.ijid.2020.02.060
5. Imai N, Cori A, *et al*. Report 3: Transmissibility of 2019-nCoV *Imperial College London COVID-19 Response Team;* 25th January 2020
6. White LF, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Stat Med*. 2008; 27(16): 2999–3016. DOI: 10.1002/sim.3136