**This paper has not yet undergone external peer review**

# Impact of wastewater testing results on preliminary estimates of COVID-19 community transmission as of 18 August 2021

**Report delivered 18th August**

**This version compiled: September 15, 2021**

**Author list alphabetical**

**James Gilmour[1,2], Emily Harvey[1,3,*], Frank Mackenzie[1,2], Oliver Maclaren[4], Dion O'Neale[1,2,*], Frankie Patten-Elliott[1,2,4], Steven Turnbull[1,2], and David Wu[4]**

[1]Te Pūnaha Matatini, University of Auckland, NZ
[2]Physics Department, University of Auckland, NZ
[3]M.E. Research, Takapuna, Auckland, NZ
[4]Engineering Science, University of Auckland, NZ
[*]Corresponding author e-mail addresses: emily@me.co.nz; d.oneale@auckland.ac.nz

## EXECUTIVE SUMMARY

On 17th August 2021 it was announced that a case of community transmission of COVID-19 with no clear link to the border or MIQ had been identified in Devonport, Auckland. In the subsequent hours a number of further cases were identified, some with links to the first detected case. It was also reported that testing for SARS-CoV-2 in the wastewater, for a catchment including Devonport, on August 11th had returned a negative result. Here we filter the ensemble of stochastic simulations from the previous report of August 17th[1] to consider the constraints from this new information.
**We find that:**
– Those simulations with at least 9 or 10 cases at the time of first detection are detected at close to the same time as outbreaks from the full ensemble of sizes; a median of around 14.5 days c.f. 13.5 days for all outbreaks. – Those simulations with **at least 9 or 10 cases at time of first detection lead to prediction of larger outbreak sizes at detection; a median of around 33 active cases** at detection c.f. around 26 for all outbreaks.

– For current estimates of wastewater testing sensitivity, and given 9 cases at time of initial detection, a negative test at the Rosedale catchment, 1 week prior to detection, is consistent with 100% of simulated outbreaks. That is, **the negative wastewater test gives no further power to discern between simulation runs.**

– **Knowledge of 2 or more undetected cases, 10 days prior to the first detected case, in addition to 10 or more cases at time of first detection**, slightly increases the expected time to detection (median of 17 days) and **appreciably further increases the expected median size of outbreak at detection (48 active cases, or 54 cumulative cases)**.

## Introduction

This note builds on the results reported in[1]. We use a detailed individual-based network contagion model that explicitly represents ∼ 5 million individuals along with the the contexts in which they interact. This network model includes stochasticity, spatial information, and individual demographic information, along with multiple distinct 'transmission contexts' including dwellings, workplaces, schools, and more generally in the community. Details of both the network and the contagion process are broadly similar to those in[2,3].

**Key include assumptions:**

- The simulations consider the case of *community* transmission, with no known link to a border worker or to an individual with mandated testing.
- Each simulation was seeded by setting the state to infected (specifically to 'Exposed') for a single, randomly selected, individual in Auckland.
- We do not explicitly consider super-spreading behaviour at the individual level (individuals with higher viral load), though individuals have different distributions of contacts and transmission rates vary by context.
- Pre-detection testing rates and behaviour were the best estimate of Alert Level 1 (AL1) in Auckland, as detailed below.
- Our vaccine coverage parameters are currently relatively approximate, as detailed below.

## Characterisation and parameterisation

Here we outline further important aspects of our parameterisation for our modelling of the new community case.

### Vaccination coverage

The scenario we consider has vaccination levels of 15% for those aged 15–29, 16.1% for those aged 30–59, and 15.4% for those aged 60+. This results in approximately 13% of the total population being fully vaccinated. This coverage is slightly different to the Auckland Metro Area vaccination status as at 27th July 2021 (3 weeks prior to August 17th — this being the date at which a person would need to have received their second dose, in order to have an immune response comparable to the vaccine effects modeled). In particular these vaccination levels are slightly too high for 15–29 year olds, and too low for over 60s.

### Vaccine efficacy

In the previous report[1] we considered 3 levels of vaccine efficacy that we referred to as 'Baseline', 'Lower' and 'Higher'. For Baseline, the vaccine leads to a 70% reduction in infection, and for those who do get infected, a 50% reduction in onward transmission. For Lower, the vaccine leads to a 50% reduction in infection, and for those who do get infected, a 40% reduction in transmission. For Higher, the vaccine leads to a 90% reduction in infection, and for those who do get infected, a 50% reduction in transmission.

Since the time to detection and size of outbreak at detection in the previous report was broadly similar across these three levels of vaccine efficacy, we here consider the results from these three settings together. This increases the number of simulations available to filter on and allows for a rough consideration of the effect of person-to-person variation in vaccine efficacy that would be seen in a real-world setting, but which is otherwise not explicitly modeled in these simulations.

### Transmission rate

For the delta variant, we consider that the individual-to-individual transmissibility rate parameter ($\beta$) is double that for the wild-type. We also assume the risk of severe consequences (hospitalisation) given infection is double that of wild-type. We do not report $R_{\text{eff}}$ estimates in these results. It is worth noting that, as our model uses an explicit network connectivity structure between individuals rather than assuming people are 'well-mixed' (fully connected), $R_{\text{eff}}$ is an emergent property of these simulations, rather than an input parameter. A doubling of the individual-to-individual $\beta$ parameter in a network model will not necessarily translate into a doubling in the overall spread measure $R_{\text{eff}}$, unless the network is fully connected.

### Testing

We assume that approximately 10% of cases with symptoms would seek a test, and of those, 90% would test positive, giving a 9% testing positive percentage. We assume that it would be a mean of 5 days from symptom onset to returning a positive test result (exponential distribution).

### Number of simulations and initial seeding

We ran 100 simulations for each scenario of level of vaccination and vaccine effectiveness. Each simulation was seeded by setting the state to infected (specifically to 'Exposed') for a single, randomly selected, individual in Auckland.

## Filtering simulations to match real-time data

Since we are interested in those simulation trajectories that most closely match the the data from real-world observations on August 18th, we filter our simulation results to include only those runs that correspond to one of two scenarios. Both of these scenarios are informed by data from a number of sources including data from *episurv* and from media updates. Data from episurv in particular can be unreliable for recently confirmed cases and can change as more information, such as date of symptom onset, becomes available through contact tracing interviews. The scenarios we consider are:

**Scenario 1:** Accurate as at 2pm, 18th August. Episurv showed 9 confirmed cases. Media reports indicated that community waste water testing on August 11th had failed to detect SARS-CoV-2 in the Rosedale wastewater catchment which serves an estimated 240,000 people on the Northshore, including Devonport. ESR give an estimated sensitivity for wastewater testing of around 10 cases per 100,000.

A sensitivity threshold of 10 in 100,000 implies that for this scenario we should filter for simulations that have at least 9 cases at the time of detection, but 25 or fewer cases, one week prior to detection of the first case. This latter limit correspond to approximate wastewater detection thresholds of approximately 10 cases per 100,000. We also consider a variation on this scenario where we filter for simulations with 10 or fewer cases, one week prior to detection. This corresponds to the situation of a more sensitive wastewater test (or to a population of individuals with significantly higher viral shedding than anticipated by testers) of slightly better than 5 in 100,000.

**Scenario 2:** Accurate as at 6:30pm, 18th August. Episurv reported 10 confirmed cases; 2 of these cases had symptom onset dates reported for earlier than the 5th and the 9th of August. Since cases will typically be infectious for 2–3 days prior to symptom onset we treat these earlier onset dates as suggesting that there were at least 2 community cases of COVID-19 where the individuals were actively infected (and infectious), on August 7th, 10 days prior to the first detected case. Symptom onset dates in episurv data should be treated as being uncertain, particularly for recent updates to historical cases. For this scenario we filter for simulations that had 10 or more cases at first detection and 2 or more cases 10 days prior to first detection.

It is worth noting that while scenarios 1 and 2 are not highly parsimonious (one selects for the *absence* of cases, 7 days prior to detection while the other selects for the *presence* of cases 10 days prior to detection) they are not incompatible. That is, it is possible for an outbreak to have 2 or more cases, 10 days prior to detection, but fewer than 25 cases 7 days prior to detection.

## Contagion simulation results

The total number of simulations for which a case was detected that was not the initial seed case was 546 (out of an unfiltered total of 600).

### *Scenario 1:*

Of the 546 simulations with detected outbreaks, 434 (i.e. 79% of the 546 simulations) had at least 9 cases at the time of initial detection. Of these, all simulations had 25 or fewer cases one week prior to detection of the first case. This suggests that for a wastewater detection threshold of 10 per 100,000 a negative wastewater test result one week prior to detection of the first case would give no additional information that might allow us to distinguish between the 434 simulations with at least 9 cases at the time of initial detection. That is, **we would expect to not detect COVID-19 via wastewater testing in 100% of the simulated cases in this scenario**.

If the wastewater testing sensitivity was approximately doubled to 5 per 100,000, the prior information of a negative wastewater test would be marginally more informative, though it would still be consistent with 396 (91%) of the simulations with 9 or more cases at detection. A negative wastewater test is still by far the most likely case consistent with this scenario.

***Scenario 2:***
Of the 546 simulations with detected outbreaks, 428 (i.e. 78% of the 546 simulations) had at least 10 cases at the time of initial detection. Of these, 230 (42% of the 546 simulations with detected outbreaks) also had at least 2 cases, 10 days prior to the time of initial detection.

In Table 1 we list the median [lower quartile, upper quartile] values for the time to detection (days), and cumulative cases at detection, and active cases at detection. These metrics are calculated for the two scenarios above, with only the increased sensitivity wastewater testing version of Scenario 1 used (i.e. detecting 10 cases for the 240,000 person Rosedale catchment).

| Scenario | $n$ | Time to detection | Cumulative cases at detection | Active cases at detection |
|---|---|---|---|---|
| No filtering | 546 | 13.6 [9.8, 17.3] | 28 [11.25, 50.75] | 26 [10, 47] |
| $\geq$ 9 cases at detection | 434 | 14.6 [11.3, 18.1] | 35 [22, 62] | 32.5 [20, 56] |
| ... and $\leq$ 10 cases 1 week prior | 396 | 14.3 [11.1, 17.5] | 33 [20, 52] | 31 [19, 48] |
| $\geq$ 10 cases at detection | 428 | 14.7 [11.5, 18.1] | 36 [22.75, 64] | 33 [20, 56.25] |
| ... and $\geq$ 2 cases, 10 days prior | 230 | 17.2 [14.9, 20.1] | 54 [34, 83.5] | 48 [30.25, 75.75] |

**Table 1.** Time to detection and size of outbreak at detection for both scenarios in the case of spread of the delta variant with 13% of the population fully vaccinated. Metrics are median [LQ,UQ] unless otherwise stated. Here $n$ is the number of the 600 simulations that reached detection and satisfied the further conditions for the scenario.

When compared with the full (no filtering) suite of simulations, imposing the conditions of $\geq$ 9 or $\geq$ 10 cases at initial detection leads to simulations with a similar time to detection (median of 14.6 or 14.7 days c.f. 13.6 days for no filtering) but with appreciably more active (and cumulative) cases at detection (median of 32.5 or 33 c.f. median of 26 case for no filtering). This is to be expected as filtering for 9 or 10 cases at detection removes the 20% of runs with the smallest outbreak sizes at detection.

Further filtering for the **scenario 1** condition compatible with a negative wastewater test, one week prior to detection, (i.e. $\leq$ 25 cases) has, as expected, has very little effect. Excluding a small number of cases that were relatively large, 1 week prior to detection, has the effect of slightly decreasing the median size of outbreaks at detection. This decrease, though not unexpected, is also unlikely to be statistically significant. We do not perform any explicit significance test.

Adding the second condition to the filtering for **scenario 2** ($\geq$ 2 cases, 10 days prior to detection) selects for simulations with a longer time to detection (17.2 c.f. 14.7 days) and for larger outbreaks at detection (48 cf. 33 active cases at detection). It does this by selecting for outbreaks with a longer pre-detection period (the median time to detection for these outbreaks is comparable to the upper quartile time to detection for the unfiltered cases). With a generation time of around 5 days, imposing the condition of having non-zero cases 10 days prior to detection has a similar effect to selecting for only those simulations that are detected on the third generation of infections. This longer period of undetected spread is associated with an appreciable increase in the size of outbreaks at detection; e.g. a median of 48 active cases at detection c.f. 33 active cases at detection for those simulations satisfying only the $\geq$ 10 cases at detection condition or 26 active cases at detection for the full ensemble of simulations.

## References


1. Gilmour, J. *et al.* Preliminary modelling of a new community case of COVID-19 as of 17 August 2021. Tech. Rep., Te Pūnaha Matatini (2021).
2. Harvey, E. *et al.* Network-based simulations of re-emergence and spread of COVID-19 in Aotearoa New Zealand. Tech. Rep., Te Pūnaha Matatini (2020).
3. Harvey, E. *et al.* Network modelling of elimination strategy pillars: Prepare for it; stamp it out. Tech. Rep., Te Pūnaha Matatini (2020).