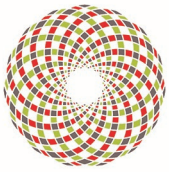

Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Note: This paper has not yet undergone external peer review

Network modelling of elimination strategy pillars: Prepare for it, Stamp it out

Report provided to officials December 9th, 2020;
this version published after internal peer review July 20th, 2021

Author list alphabetical

Emily Harvey[1,2,*], Oliver Maclaren[3], Dion O'Neale[1,4,*], Frankie Patten-Elliott[3], Steven Turnbull[1,4,5], and David Wu[3]

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

EXECUTIVE SUMMARY

We use an Aotearoa-specific network contagion model to simulate the spread of COVID-19 in the community in two different regimes that correspond to the **Prepare for it** and the **Stamp it out** pillars of Aotearoa New Zealand's COVID-19 elimination strategy. Within each of these regimes we consider a range of different scenarios that represent different behavioural and non-pharmaceutical interventions (NPIs), all aimed at achieving elimination. For spread **prior to detection** (Prepare for it), we consider the effects of testing and behavioural changes including: lower pre-detection transmission rates relative to baseline (e.g. more mask use and general awareness), shorter delays in returning test results, and higher proportions of symptomatic cases seeking tests. For these pre-detection scenarios we report the **size of outbreak at detection and the time to detection** from the exposure of the seed case. For **post-detection** simulations (Stamp it out), we consider interventions including: greater or lower post-detection transmission rates relative to baseline, the proportion of close or casual community contacts being tested, and the rate at which close contact tracing is carried out. For these scenarios we report **time to complete isolation and total (cumulative) outbreak size**. We focus here on **Alert Level 3**, as we find that simulations for Alert Level 2.5 have uncontrolled growth with R_{eff}>1. We consider AL2.5, and other scenarios which may not be sufficient to achieve elimination, in a companion report.

Key points:

- For the testing proportions and time from symptom presentation to positive test result currently observed at Alert Level 1 (AL1), we would expect that for **a community outbreak event with no direct link to the border the median time to detection of the first case would be 19 days** [LQ 13 days, UQ 26 days]. The outbreak size at detection most likely to be associated with this would be on the **order of 20–30 cases**.

- **Pre-detection** (Prepare for it), the most effective behavioural change for **reducing both the time to detection and the initial outbreak size** is an **increase in the proportion of symptomatic individuals seeking testing**.

- **Post-detection** (Stamp it out), we find that **for an outbreak of size 11–20 total** (both known and unknown) cases at first detection, **elimination, defined as complete isolation of an outbreak, would be obtained after a median of 42 days** [LQ 29 days, UQ 58 days] at **Alert Level 3** (parameters comparable to AL3 in the Auckland August outbreak). **For an initial outbreak of 21–50 total cases, the median duration at AL3 needed for complete isolation is 65 days** [LQ 50 days, UQ 80 days].
- There is **significant variation in total outbreak size and time spent at AL3 until complete isolation**. These depend not only on the size of the initial outbreak, but also on details of the contagion tree as cases spread through different parts of a heterogeneous interaction network.
- There is **large overlap between the outcomes from all Alert Level 3 interventions, suggesting that strong transmission controls of AL3** are the dominant factor in stamping out an outbreak.
- This is corroborated by our findings that **Alert Level 2.5 is insufficient to lead to elimination**, with $R_{eff}>1$. We consider AL2.5 and a broader range of scenarios and sensitivity studies in companion reports.
- For comparison, we note that the parameterisation of ('optimistic') AL2.5 used for the **branching process model** in a companion Te Pūnaha Matatini report[1] is likely more representative of our AL3. Hence, **the total time at AL3 and AL2.5 in that report should be compared with the duration at AL3 in this report**.
- The AL3 interventions we consider, in particular improvements in contact tracing (QR codes, Bluetooth, faster tracing), **will likely have different effects when applied in different scenarios or at different Alert Levels, however**. The impacts of these at different alert levels are not simulated in the present report, but will be in future reports.

Introduction

Aotearoa New Zealand's government has received praise for its science-informed approach to COVID-19, with computer simulation and modelling playing an important supporting role for decision making. Following the first reported case of COVID-19 in Aotearoa on the 28th of February, 2020, early simulation models[2,3] helped inform an **elimination strategy** which saw the country enter a nationwide lockdown and close its international borders to everyone except citizens and residents. In contrast to many nations, New Zealand has successfully eliminated COVID-19 (at least twice), and maintains its elimination strategy. This strategy has evolved since initial incursions, and is now based on three key pillars: **keep it out**, **prepare for it**, and **stamp it out**. Department of the Prime Minister and Cabinet (DPMC) has requested modelling work to support their decision making in the context of this elimination strategy. The goal of the requested modelling work is to provide quantitative assessments of particular risk measures DPMC has identified under scenarios associated with each pillar.

Since February 2020, NZ-specific COVID-19 models have grown in sophistication, incorporating additional uncertainties and heterogeneity. In particular, Te Pūnaha Matatini (TPM) researchers currently use two types of model to provide advice to Government: a **branching process model** and a **network contagion model**. The first of these incorporates stochasticity and some small amounts of heterogeneity; the second is significantly more complex, and is a model of all ≈ 5 million individuals in New Zealand and 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, and schools. It also includes an explicit representation of the contact tracing process. Each of these models has strengths and weaknesses, and are useful for addressing different policy questions.

Here we present simulations using our **network contagion model** addressing **two of the three pillars**: **Prepare for it**, and **Stamp it out**. The key risk measures for the scenarios associated with these two pillars are: **time from re-incursion to detection** and **size of outbreak at detection** (Prepare for it scenarios), and **time to elimination** (Stamp it out scenarios). A companion report, from Te Pūnaha Matatini, based on the branching process model addresses the situation of re-incursion linked to the border[1]. In both our Prepare for it and Stamp it out scenarios, we focus on the more concerning situation in which a **community outbreak is detected and the source is unknown**, rather than the case of a positive test in a known MIQ worker or person with an epidemiological link to an MIQ worker. Here we focus on scenarios in which we would expect to achieve elimination; our companion reports consider other cases such as the impacts of variants with increased transmission, weaker or delayed intervention, in which case the ability to achieve elimination is less clear *a priori*.

Description of model and key assumptions

We implement a stochastic model of infection dynamics on a detailed interaction network of all ≈ 5 million individuals in Aotearoa NZ. Each **individual** is represented by a node in this network; additional **group** nodes are used to represent the different **infection contexts** through which individuals can interact and transmit infection. Each individual has the demographic

characteristics of age, sex, ethnicity, and geographic location (Statistical Area 2 (SA2)) of usual residence. These are sourced from Census 2018 figures. Individuals are placed in **dwellings**, with other individuals, in the same geographic location (SA2) based on Census 2018 dwelling size and age structure within that SA2. Besides dwellings, many individuals have places of **work** (tax data from the Statistics NZ Integrated Data Infrastructure (IDI)) and/or **education** (Ministry of Education roll data), and all individuals participate in so-called **community** events which capture all interactions which are not with other people in their dwelling, or work/school (i.e socialising, shopping, etc). Currently the community layer includes a mean of 1.4 events per person, with Poisson distribution, with a mean size of 2.4 people following a heavy-tailed (power-law) distribution with an upper limit of 350. The initial set of community interactions is only within each TA, then long-range travel links within Aotearoa New Zealand, are added as approximately 5% of interactions, with the TA to TA weightings based on cellphone movement data[4].

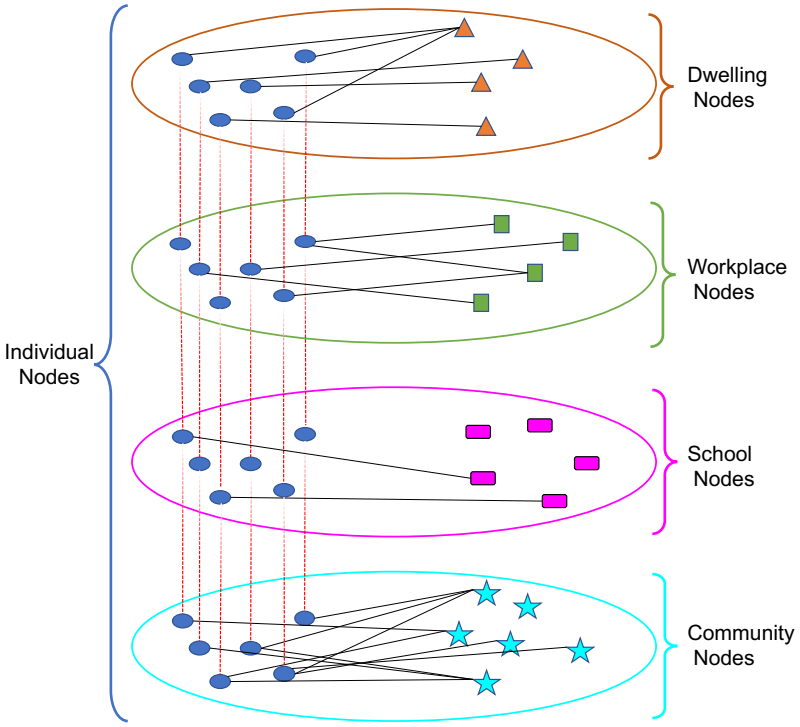

Figure 1. Diagram of the multilayer bipartite interaction network showing individual nodes on the left, which persist (are the same people) in all the layers of the network. The current version of the model uses four layers, one each for dwellings, workplaces, schools, and community.

Group nodes are further classified as involving **close** or **casual** (background) contact behaviour. In small dwelling, school, and workplace groups, we assume that all contacts are both close and casual contacts. But if groups are larger (e.g. a whole school or a large workplace), we create smaller groups within the large group which represent the smaller number of close contacts such as a class within a school or a team within a workplace. Within the large school and workplace groups, we assume only a casual contact level of interaction. For community events, however, we categorise these as either close or casual contact type interactions in advance, with approximately one third of community events being close.

Close contacts are assumed to be contact traceable, if 'known' (by the National Contact Tracing Solution team), and the proportion of close contacts who would be 'known' varies by context (e.g. all dwelling close contacts are known, but only a lower proportion of social/community close contacts are known) and is influenced by contact tracing effectiveness and technology assumptions. In our model, casual contacts are not directly traced, but if they know that their casual contact(s) have been confirmed as COVID-19 positive (through e.g. media reports, NZ COVID Tracer app alerts), then they increase their likelihood of seeking a COVID-19 test. The proportion of casual contacts who would know that they may have been exposed, and who subsequently, would seek a test varies by interaction context. This number is also influenced by the methods used by public health officials for notifying people and what the advice to casual contacts is. Our model also includes policy

effects, representing non-pharmaceutical interventions such as Alert Levels, which reduce the chance of transmission in various interaction contexts, and increase community (general public) test-seeking behaviour.

We use the Gillespie algorithm[5] to simulate our main contagion dynamics. This algorithm is a so-called exact algorithm for simulating realisations from a collection of independent *transition* processes with rates (or *hazards*) of the form:

$$\text{Probability of transition } i \text{ per unit time} = h_i(\text{system state}) \tag{1}$$

for $i = 1, ..., n$. In this context, a transition consists of one or more individuals changing their state; for example, a susceptible individual becoming exposed due to an encounter with an infected individual. This simulation approach was popularised in the stochastic chemical kinetics literature, but also has a long history of applications to population dynamics[6] and is now standard in the network contagion literature[7].

Although our main dynamics are Markovian and compatible with the standard Gillespie algorithm, we incorporate additional realism by allowing for delay processes, particularly for our quarantine and tracing dynamics. We do this by having Markovian 'initiation' dynamics but delayed 'completion' times, with arbitrary distributions allowed for completion times. Delay processes can be handled in an exact manner using an extension of the standard direct Gillespie method[8,9], and we follow this approach here.

Key assumptions

Full model details are described elsewhere[10], but key assumptions and parameters include:

- Disease progression for infected individuals proceeds through a sequence of states. Initially, exposed individuals are infected but not yet infectious, they transition to either pre-symptomatic or asymptomatic states (both infectious). Cases that will go on to develop symptoms (pre-symptomatic) are further split into being 'mild' or 'severe' cases, with this varying by age. Note that this case severity is based on eventual outcome, and is not specifically the symptom severity at onset. Severe cases can become hospitalised and can die while the remainder of the infected cases recover. The parameters controlling transitions between these states are based on international literature — primarily from Fraser et al.[11].
- There is negligible difference in the rate of recovery from exposure between different symptom presentations.
- The proportion of infections that are asymptomatic varies with age, and equates to about 16% over the whole population, in line with findings from PCR based studies with inclusive symptom case definitions[12,13]. Asymptomatic cases are assuming to have zero chance of being tested for COVID-19 unless they have been identified as a casual or close contact of a confirmed case.
- The split of (pre-symptomatic and) symptomatic cases into being 'mild' or 'severe' cases, also varies by age. Case severity determines primarily the probability of hospitalisation (zero for mild cases)[10], but also affects the likelihood of infected individuals to seek testing (severe cases are more likely to seek medical attention and thus have a higher testing probability).
- Infectiousness of individuals at any infectious (asymptomatic, pre-symptomatic, or symptomatic) stage of their infection is identical to any other individual in the same infectious stage of infection.
- Infectiousness of asymptomatic and pre-symptomatic individuals is identical to symptomatic individuals.
- We model the overall proportions of individuals seeking tests, without being contact traced, separately to the rates at which (positive) test results are returned. We do this by randomly dividing the population into various 'test-seeking' and 'non-test-seeking' groups at the beginning of each simulation. These groups depend on infection state and scenario settings, i.e. individuals are classified into multiple groups according to whether they are 'test-seeking when in state *X* under Alert Level *Y*'.
- Individuals will not isolate or change their behaviour until they either receive a positive test, or are contacted by contact tracing.
- Individuals will stay in self-isolation for 14 days, starting from when they are first contacted by contact tracing.
- Individuals that receive positive tests will stay in an MIQ facility until they recover.
- Infected individuals in self-isolation can infect others in their dwelling and have a small (1%) chance to infect others outside of isolation.

- Infected individuals in MIQ (confirmed cases) have a small (1%) chance to infect others outside of isolation/quarantine.
- Contact tracing traces a proportion of close contacts of a confirmed case. The proportion depends on the interaction group type, as well as scenario settings.
- Close contact tracing is modelled by a finite number of attempts at contact, each with a small chance of failure, dependent on the type of group (e.g. work vs dwelling) that connects the close contact with the confirmed case.
- Casual contacts of a confirmed case have a certain probability that they would both i) know they were a close contact, and then ii) would go to get a test; with some delay between the confirmed case notifying and a test being returned to the casual contact. The probability that a casual contact would seek a test depends on the interaction group type or context, as well as scenario settings.

Prepare for it

The Prepare for it simulations model the case of community spread of COVID-19 prior to the first detected and confirmed case of an outbreak. The contagion takes place in a setting that represents the current Alert Level 1 situation at the beginning of December 2020. A key consideration in the speed to detection of any new outbreak is the proportion of symptomatic individuals being tested — these vary from week to week and between different regions of Aotearoa. The testing proportions used in the baseline scenario presented here are representative of those currently seen in Auckland. While Auckland testing proportions are typically appreciably higher than those in other parts of the country, it is also most likely that any re-emergence of COVID-19in the community would also occur in Auckland. We also assume that transmission is at baseline levels in all interaction contexts, with the exception that we assume that there is some mask wearing and social distancing which will slightly reduce casual community transmission.

Scenarios considered

- **Baseline Alert Level 1**: For this scenario we assume that there is some mask wearing and social distancing which will reduce casual community transmission by 10% below the model's calibrated default values. We set the proportion of symptomatic cases who would seek a test to 10% for mild/moderate cases and 50% for severe cases*, based on levels of testing in Auckland estimated from FluTracking data[14]. We assume a test positivity rate of 80%[15], which equates to a probability of detection of 8% for mild/moderate cases, and 40% for severe. Finally, we assume the time from symptom onset to test result is exponentially distributed with a mean time to result of 5 days. NB: the proportion of cases detected will vary depending on age solely due to the higher proportion asymptomatic and lower proportion severe. For example, in the baseline case 0-14 year olds will have a case detection of 7% whereas over 60s will have a case detection of 16%. We know from FluTracking data[14] that testing proportions are much lower in younger age groups even for the same symptom presentation, so we suspect that we are overestimating testing in young people.
- **Lower transmission rates at Alert Level 1**: For this scenario we consider the impact of having reduced transmission at Alert Level 1 – these include social distancing and mask use. We implement this as 20% reductions in close contact transmission in workplace and community contexts (i.e. everything except dwelling and school close contacts), 40% reductions in casual contact transmission within the workplace context, and 44% reductions in casual contact transmission within the community context. NB: these reductions are relative to the model's calibrated default values, not relative to other scenarios.
- **Faster community testing**: For this scenario we increase the speed of testing so that the mean time from symptom onset to test result is only 2 days (following an exponential distribution). NB: This is a very best case lower bound - 3 days is the best we have seen so far, with closer to 4 days for confirmed cases during the August outbreak.
- **Higher proportion of symptomatic cases seeking tests**: Based on FluTracking data[14], for this scenario we increase the proportion of symptomatic cases who would seek a test to i) a 'medium' level of: 20% for mild/moderate cases and 70% for severe cases and ii) a 'high' level of: 30% for mild/moderate cases and 80% for severe cases.
- **Combinations**: The two combinations we consider here are faster testing (2 days) combined with both the 'medium' and 'high' proportions of symptomatic cases seeking testing.

*severe cases are those that would be expected to seek medical attention for breathing difficulties, pneumonia, etc.

Comparison with parameters used with the TPM branching process model

The TPM branching process model has also been used to investigate the 'Prepare for it' pillar[1]. Here we compare and contrast the different parameters used for scenarios and some underlying assumption differences.

Asymptomatic cases and infectiousness: The branching process model assumes an asymptomatic proportion of ≈30%, and a reduced infectiousness of 50% for asymptomatic and pre-symptomatic cases. In contrast, in the network model we have a much lower proportion asymptomatic (16% over the whole population, but varying with age) and assume the same infectiousness for all asymptomatic, pre-symptomatic, and symptomatic cases.

Testing speed: For time from symptom onset to confirmed test result, the TPM branching process report considers a mean of 6 days as the baseline and a mean of 3 days as their 'fast' testing speed. These are both one day slower than in the network model results presented here. Furthermore, the network model times to test are distributed exponentially due to the Markovian processes in our model, whereas the branching process model may use a different distribution for their times. This assumption can be relaxed in our approach by treating testing in terms of delayed processes, but this is not done here.

Probability of detection: The TPM branching process report uses the probability of detection (testing positive) for symptomatic cases of 30%, this equates to 37.5% of symptomatic cases seeking testing (assuming a high 80% test positivity rate[15]. They then use 50% symptomatic case detection (62.5% seeking a test) as their 'high' level of testing. However, they also assume double the proportion of cases are asymptomatic (30%) compared to in our model.

Combining all the different parameters and assumptions, the branching process report has overall (symptomatic and asymptomatic) case detection proportions of 21% for baseline and 35% for high, whereas we have case detection probabilities of baseline=10%, medium=18%, and high=25%.

Reduced transmission at AL1: The TPM branching process report uses an R_0 of 2.5 for baseline, and reduces this to 2 for their lower transmission scenario — this is a blanket transmission reduction of 20%. In the network model, however, we specify our reduction based on how it affects different types of interactions (close vs casual) in and different interaction contexts. Here we assume a 20% reduction in transmission in close contacts for workplaces and community, and a greater reduction of 40% in casual contacts transmission. The network model splits potential infections between close and casual contacts and between interaction contexts internally, so the actual transmission reduction will differ slightly person to person, rather than being a blanket reduction.

Simulation results

Here we present our results from the Prepare for it scenarios under Alert Level 1 with variations in testing and behavioural settings. Our key results are shown in Tables 1 and 2, as well as Figure 2 below.

We find that for the Baseline AL1 scenario, the time to detection of the first case in a community transmission outbreak is from two to just under four weeks (median=19 days; LQ=13, UQ=26), with a most common outbreak size (according to the size bands given) in the range of 21–50 total cases.

The most important parameter in reducing both the time to detection and the outbreak size appears to be the proportion of symptomatic cases which seek testing: when this is increased to 30% (a highly optimistic proportion, comparable to that observed in Auckland during periods of known community spread) the time to detection reduces from a median of 19 days in the Baseline to a median of 14 days with the increased testing proportion. With this higher testing proportion, 44% of outbreaks simulated were smaller than 10 total cases at detection. Faster community testing (time from symptom onset to return of rest results) also decreases the time to detection and the size of the resulting outbreak.

Scenario	Time to detection (days)	Number of simulations	Outbreak size at detection
Baseline AL1	19 [13, 26]	84	23 [10, 37]
Lower transmission levels at AL1	18 [13, 28]	84	13.5 [9, 24]
Faster community testing	16.5 [11, 23]	88	15 [6, 24]
Medium proportion symptomatic cases seeking testing	18 [12, 24]	85	12 [5, 23]
High proportion symptomatic cases seeking testing	14 [9, 21]	82	10 [4, 21]
Medium proportion and faster testing	13 [8, 20]	82	9.5 [3, 17]
High proportion and faster testing	10 [6, 17]	84	5 [3, 13]

Table 1. Time from re-incursion to detection and size of outbreak at detection All results are formatted as *median [lower quartile, upper quartile]* unless otherwise specified. Number of simulations vary, as not all incursions get detected before dying out.

Scenario	P (2-10)	P (11-20)	P (21-50)	P (51-100)	P (>100)
Baseline AL1	0.20	0.19	0.40	0.11	0.04
Lower transmission levels at AL1	0.27	0.38	0.15	0.15	0.00
Faster community testing	0.38	0.27	0.26	0.06	0.00
Medium proportion symptomatic cases seeking testing	0.38	0.28	0.21	0.07	0.00
High proportion symptomatic cases seeking testing	0.44	0.23	0.23	0.02	0.00
Medium proportion and faster testing	0.45	0.26	0.18	0.01	0.00
High proportion and faster testing	0.54	0.18	0.14	0.00	0.00

Table 2. Proportion of outbreaks in each size band at detection — excluding those with only a single case, i.e. where the detected case is also the seed case.

We also present these results in terms of confidence intervals and tests relative baseline in Figure 2, binning across all outbreak sizes for convenience. These also give a clear indication that the proportion of cases seeking testing is the parameter with the most important impact.

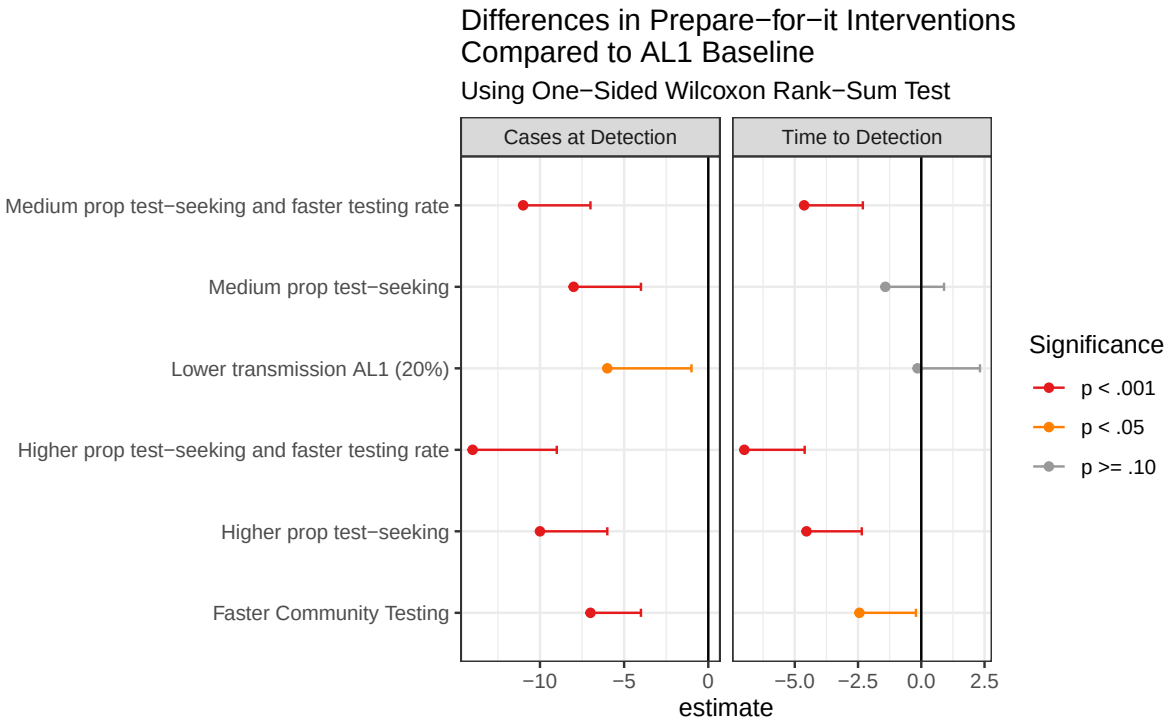

Figure 2. Time from re-incursion to detection and size of outbreak at detection. This figure shows the impact of various behavioural modifications at AL1 on our key metrics of time to detection and cases at detection. These are the same results as those in Table 1 and 2 above, but shown in terms of tests and interval estimates of shift relative to baseline.

Comparing our results with the TPM branching process model, we see similar sizes for the Baseline case, but fewer very large outbreaks for improved testing. We suspect that this may be because the branching process model explicitly includes a super-spreading term, when one person infects many all at one time. In the network model, we currently only have differences in infection driven by the network structure. This means that for the short time to detection runs, we expect to be less likely than them to get very large outbreaks.

Finally, we emphasise that the faster testing rate here is likely to be the maximum obtainable rate with current technology and current availability of testing centres, while the testing proportion could be increased (and has been during periods of heightened public awareness during outbreaks).

Stamp it out

The Prepare for it scenarios, in the previous section, give an indication of the nature of contagion that would occur prior to detection of community transmission of COVID-19. Upon detection, it is expected that there would be a centrally coordinated response from health officials with a goal of returning to a state of elimination of COVID-19 transmission in the community — the so-called **Stamp it out** pillar. For the outbreak sizes observed in the Prepare for it simulations, with no known link to the border, the most likely (and appropriate) response to detection of community transmission would be to raise Alert Levels to AL3, as occurred with the Auckland August 2020 outbreak. The elevated Alert Level can also be augmented with a range of additional NPIs that were not part of the initial Auckland August response, including behavioural and technological factors. During the Auckland August outbreak, only Auckland was moved to AL3, while the remainder of the country remained at a lower Alert Level. In these simulations, we do not consider regionally limited interventions. This is on the assumption that any regional boundaries would be chosen such that the AL3 intervention would extend sufficiently to apply to all infected individuals in the outbreak and that movement restrictions would be in place to ensure that the outbreak remained within the region where AL3 applied.

In order to create initial conditions for the Stamp it out simulations, we set simulations running using parameters for the 'Baseline' Alert Level 1 Prepare for it scenario until detection. This ensured that for a given initial outbreak in the Stamp it out simulations we start with not just an initial number of infections but also a full contagion tree compatible to allow for contact tracing. As a consequence, the same intervention is applied to a variety of initial conditions, and the number of simulations for outbreaks of different sizes at detection is not identical across interventions. However, all of these simulations do originate under more probable conditions for re-emergence of COVID-19 in the community, prior to detection. Distributions for the outbreak sizes corresponding to the initial conditions for the various AL3 interventions simulated are given in the Appendix of this report (Table 5 and Figure 6). This provides more realistic infection histories for each simulation, but also increases the run-to-run variability within and between different interventions. It also means that the number of simulations for outbreaks of different sizes at detection, and for different scenarios of AL3-like interventions, is not identical.

Alert Level 3 Interventions

Baseline AL3

In the baseline for Alert Level 3 we assume that there is widespread reduction in transmission through various control policies. We assume that schools are closed, with limited capacity for the children of essential workers and some transmission reduction measures in place. Workplaces would encourage working from home where possible, and measures are taken to keep workers safe, for example through physical distancing, increased hygiene measures, contactless interaction and use of PPE. In the community, we assume that gatherings over 10 people are not allowed, with appropriate considerations for weddings, funerals, and tangihanga. We also assume that there is widespread adoption of mask-wearing and distancing, as well as travel restrictions and general reduction in interactions. Overall we expect these measures to reduce the likelihood of transmission in schools by around 97%, workplaces by around 55% and in all other (non-dwelling) interactions by around 90%. These reductions are relative to the model's calibrated default values, not relative to other scenarios.

Once the first case has been detected, we assume there is a testing surge and 50% of mild symptomatic community cases (40% of cases test positive) and 80% of serious symptomatic community cases get tested (64% of cases test positive), with results back in a mean of 4 days post symptom onset, matching rates seen in the Auckland August outbreak. Individuals that test positive are isolated for 14 days. We assume they are placed in a managed isolation facility, where there is full compliance, and contact is only allowed between people that normally reside in the same dwelling. We model breaches and mismanagement with a small 1% chance that isolated individuals infect individuals outside their dwellings.

Close contacts are traced through the NCTS, using a Weibull Distribution with parameters: $scale = 3.04$, $shape = 2.66$, which is fit to metric S0003 in the National Contact Tracing Performance reports. It is assumed that all members of a confirmed case's dwelling, 95% of close workplace or school contacts, and 80% of community close contacts are known and reachable by contact tracers. Contact tracing attempts are then made with a 10% chance of failure per attempt for contacts that do not share a dwelling with the confirmed case. If an attempt fails, up to 5 more attempts are made at making contact, with an inter-attempt time modelled by an exponential distribution with rate = 4. Successfully contacted individuals also isolate for 14 days, behaving similarly to individuals who have tested positive.

Casual contacts are notified through media announcements, NZ COVID Tracer app 'exposure notifications', and some manual contact tracing where capacity allows. Casual contacts are advised to get a test regardless of symptoms, emulating the government's policy for casual-plus contacts. In the baseline we assume that 50% of casual contacts in dwellings[†], 50% of casual contacts in schools and workplaces, and 10% of casual contacts in the community would know they were a casual

[†] all dwelling contacts will be close contacts except in large shared dwellings (occupancy over 12). NB: apartment complexes are not large dwellings, as these are considered separate, and unlinked, dwellings in Census records.

contact and seek a test. For this percentage, the delay between the confirmed case notification and the casual contact seeking a test is between 2.5 and 7 days, with the times modelled as a scaled Beta distribution with parameters $a = 3, b = 5$.

Stronger/Weaker AL3

Here we use the same parameters for close and casual contact tracing processes as in Baseline AL3, but with 10% greater and 10% lower overall transmission reductions due to AL3 in school, workplace, and community contexts.

Contact tracing changes at Alert Level 3

The three areas of contact tracing enhancements of interest here are: 1. Wide use of COVID Tracer app; 2. Wide use of a Bluetooth tracing technology, and 3. More effective manual contact tracing. In order investigate the impact of these contact tracing changes within our model, we first translate these interventions into expected changes in our model parameters. This necessarily involves additional assumptions that may not reflect the actual impacts of these interventions in reality. We also emphasise that here we consider **the effect of contact tracing enhancements over and above an Alert Level 3 intervention scenario.** In general we expect that some technologies may have more or less impact depending on the alert level. For example, contact tracing may have less of an effect if people are confined to their homes already. That is, we do not expect the effects of interventions to simply add (or multiply) together, but instead expect them to depend on the other model settings.

AL3 + wide use of COVID Tracer app

There are multiple expected impacts on contact tracing of the wide use of the COVID Tracer app. Firstly, just downloading and registering in the app has given NCTS a more up to date database of contact details for people. This has been a big improvement, but is assumed to already be included in our Baseline AL3, as the bulk of this impact has already been felt. Secondly, the COVID Tracer app allows 'exposure event' notifications to be sent out to users who scanned in at a similar time as a confirmed case is known to have been there. This list of people would include both close and casual contacts. We assume that close contacts would have been identified anyway through other routes of investigation, and that the main impact of the COVID Tracer app would have here would be to speed up the contact for these. For casual contacts, the only other way people would know is through media announcements and press releases. This means that the use of the app would have a big impact on the chance that someone would know that they had been a casual contact of a confirmed case, and increases the likelihood of them getting tested. This effect would be strongest in community interactions, as most dwelling, workplace, and school casual contacts are known already through other means.

In terms of parameter changes in our model, **the only impact we explicitly test here is the likelihood of casual contacts in the community getting tested increasing to 25%**. We keep the same transmission reductions, the same proportion of close contacts known, and the same close contact tracing speed as in Baseline AL3. We explicitly do not include how much use of COVID Tracer app would change the speed of close contact tracing, as that is tested independently, and we are attempting to disentangle the different effects.

AL3 + wide use of Bluetooth tracing technology

As with the COVID Tracer app, the use of Bluetooth tracing technology would be expected to have a number of different impacts. The main impact is that it would enable more of the unknown close contacts to be 'known' than would be possible through other manual contact tracing routes. For example, the close contact type interaction through sitting next to a stranger on a bus, or standing next to them for a period of time in a checkout queue at the supermarket. Another key impact is that it could potentially speed up the rate at which close contacts would be identified and contacted, even for close contacts who would be known through other manual contact tracing investigation pathways.

In our present Bluetooth simulations **the only impact we explicitly test here is if the proportion of close contacts known increases to 95% for community close contacts (from 80% in the base case)**. We keep the same transmission reductions, the same likelihood of casual contacts getting tested, and the same close contact tracing speed as in Baseline AL3. Importantly, **here we do not include how much a Bluetooth tracing technology would change the speed of close contact tracing**. Instead, we consider the effect of faster close contact tracing independently. This allows us to consider the impact of faster tracing regardless of whether it comes about via Bluetooth or via some other mechanism.

AL3 + faster close contact tracing

For this scenario, we investigate the impact of faster close contact tracing. As noted above, **we do not differentiate between mechanisms for how this speed-up is achieved**. In particular, this could be due to technology advancements (COVID Tracer app, Bluetooth tracing technology, etc) as well as improvements in manual contact tracing processes. Improvements in contact tracing speed would most likely come as a combination of these. An optimistic estimate of the total impact of Bluetooth technology could then, for example, be given by a combination of this and the previous intervention.

In terms of parameter changes in our model, **the impact we explicitly test here is the close contact tracing speed increasing**. We implement this by adjusting the parameters of the Weibull Distribution to: $scale = 2.6$, $shape = 1.7$. This is a

reduction of the median time to notify a close contact from 2.65 days in the baseline case to 2.10 days in this scenario. We keep the same transmission reduction, the same proportion of close contacts known, and the same likelihood of casual contacts getting tested as in Baseline AL3.

Alert Level 2.5 Interventions with baseline contact tracing

We find that simulations for a strict Alert Level 2.5 has uncontrolled growth with $R_{\text{eff}} > 1$. This was estimated using the approximation $R_{\text{eff}} \approx r \times$generation time + 1, where r is the observed exponential epidemic growth rate[16,17]. We observe a median of 1787 [606, 2951] cases at 120 days from detection, hence we cannot calculate or report on time to elimination, and we do not include these results here. We have not yet considered simulations of Alert Level 2.5 with improved contact tracing.

Simulation results

Here we present results for interventions applied in order to 'stamp out' an outbreak. It is difficult to determine exactly when in each scenario the intervention would be considered successful ('elimination' considered as achieved) and the Alert Level 3 intervention would be lifted. Here we measure this in terms of the time to complete isolation of all cases (containment of the outbreak). This is the time at which there are no undetected cases still in the community (some cases can still be undetected, but as long as they are in self-isolation through being a known close contact, we consider this contained). This is not a directly observable time in reality, but can be informed by modelling studies such as the present work. We also report the final size of the outbreak, both in terms of known (confirmed) cases and all infections (known and unknown).

We find that for the Baseline scenario, the days to complete isolation is around 16 days [LQ = 4, UQ = 35] for outbreaks initially sized between 2 and 10 cases on discovery (see Table 3). This number is greater for outbreaks of greater sizes; 42 days for outbreaks of 11-20 cases, 65 days for outbreaks of 21-50 cases, and 88 days for outbreaks of 55 to 100 cases.

Across the scenarios tested, we found high levels of variability in the number of days taken for outbreaks to reach complete isolation (see Figure 3), and the cumulative total number of cases per outbreak (see Table 4 and Figure 4). This makes it difficult to disentangle the impact of the different scenarios. Considering our results in terms of tests and confidence intervals for differences relative to baseline in Figure 5, we see that there is no strong signal present. We do see clearly that larger initial outbreaks tend to take longer to control and result in a larger cumulative total number of cases across all scenarios, however.

Scenario	Days to complete isolation for different initial outbreak sizes			
	2-10	11-20	21-50	51-100
Baseline AL3	16 [4,35]	42 [29,58]	65 [50,80]	88 [79,112]
Stronger AL3	9 [4,21]	32 [23,53]	61 [46,81]	68 [64,96]
Weaker AL3	23 [6,29]	62 [30,91]	63 [51,99]	144 [102,164]
AL3 + faster tracing	23 [11,40]	55 [41,81]	61 [41,72]	81 [68,127]
AL3 + wide use of COVID Tracer app	28 [7,48]	41 [29,50]	77 [53,95]	125 [114,162]
AL3 + Bluetooth tracing technology	20 [6,35]	58 [26,108]	64 [48,77]	114 [99,129]

Table 3. Time until all cases are either known or in close contact isolation, for different outbreak sizes. Results shown are *median [lower quartile, upper quartile]* unless otherwise stated.

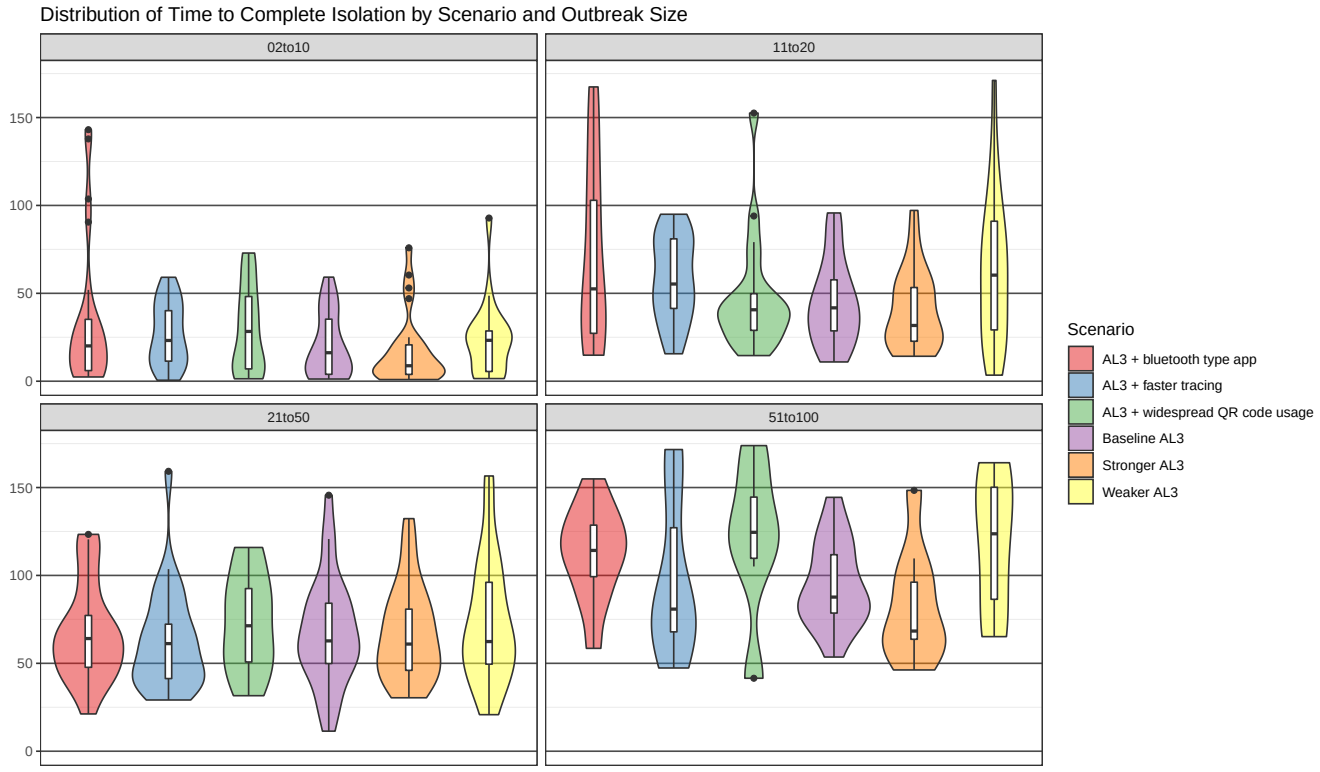

Figure 3. Distribution of Time to Complete Isolation by Scenario and Outbreak Size at Initial Detection. This plot shows the distributions of the times until all cases are either known (confirmed) or in isolation as a consequence of being a close contact of a confirmed case. Results are split by outbreak size at initial detection (four sub-figures) and across the scenarios of Baseline AL3, Stronger AL3, Weaker AL3, AL3 with faster tracing, AL3 with a Bluetooth tracing technology, and AL3 with wide use of COVID Tracer app. Density plots indicate the lower quartile, median, upper quartile, and distribution of times across the set of simulations for each initial outbreak size and intervention scenario.

Scenario	Cases	Size of outbreak at detection			
		2-10	11-20	21-50	51-100
Baseline AL3	Confirmed	8 [4,12]	26 [22,29]	74 [41,113]	145 [131,193]
	Total	14 [7,26]	48 [43,62]	133 [80,207]	281 [256,419]
Stronger AL3	Confirmed	5 [3,7]	23 [17,28]	65 [47,90]	147 [135,162]
	Total	10 [6,14]	39 [34,60]	135 [95,166]	259 [247,328]
Weaker AL3	Confirmed	10 [5,17]	35 [23,54]	79 [43,120]	156 [142,217]
	Total	17 [6,29]	60 [37,105]	158 [97,214]	322 [289,421]
AL3 + faster tracing	Confirmed	8 [4,14]	37 [27,44]	73 [52,89]	160 [148,185]
	Total	15 [10,29]	67 [54,101]	139 [100,168]	326 [267,383]
AL3 + wide use of COVID Tracer app	Confirmed	8 [4,15]	25 [20,35]	75 [64,107]	183 [146,224]
	Total	16 [6,26]	49 [36,64]	134 [111,218]	338 [275,419]
AL3 + Bluetooth tracing technology	Confirmed	7 [4,17]	37 [25,62]	80 [60,130]	237 [163,298]
	Total	10 [4,25]	65 [45,95]	121 [101,187]	353 [261,460]

Table 4. Final size of outbreak (confirmed cases and total cases) for different initial outbreak sizes at detection and interventions. Results shown are *median [lower quartile, upper quartile]* unless otherwise stated.

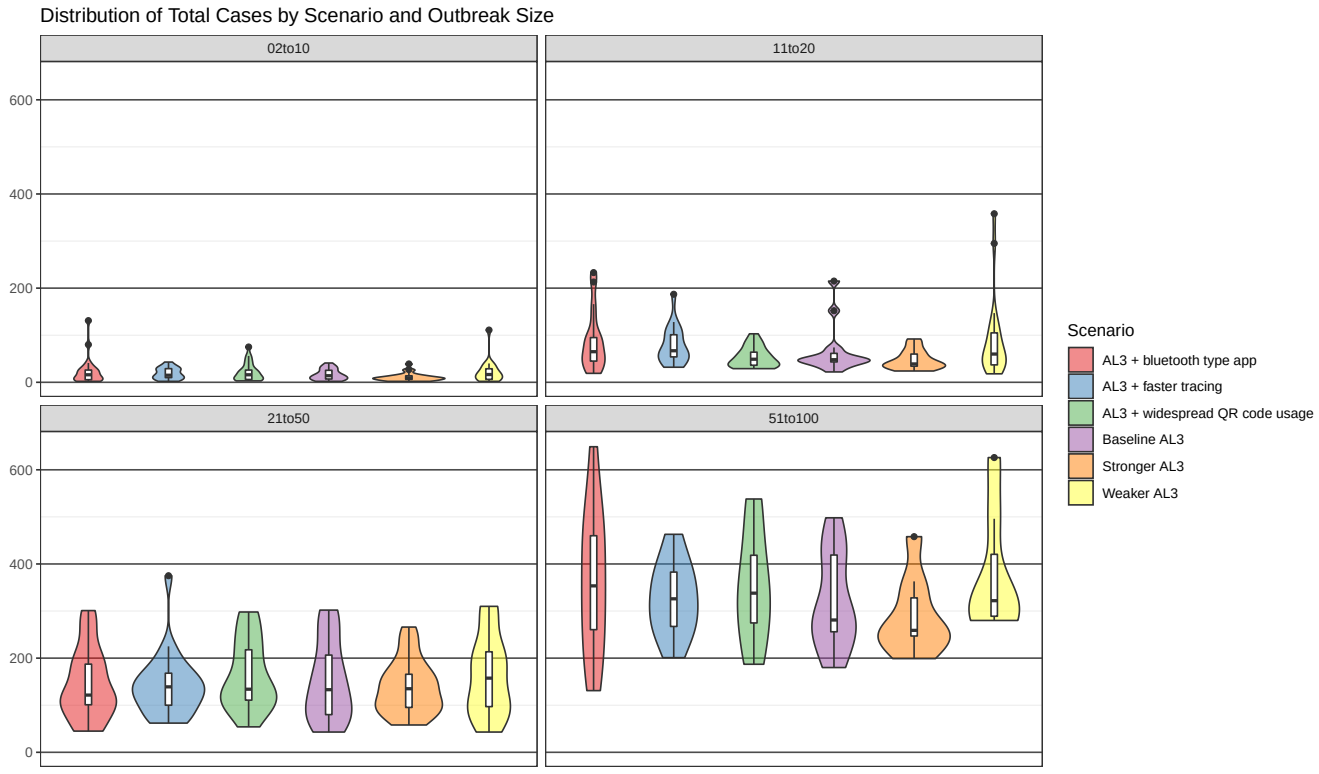

Figure 4. Distribution of Total Cumulative Cases by Scenario and Outbreak Size at Initial Detection. This plot shows the distribution of the total cumulative number of cases (i.e. both confirmed cases and unknown cases). Results are split by outbreak size at initial detection (four sub-figures) and across the scenarios of Baseline AL3, Stronger AL3, Weaker AL3, AL3 with faster tracing, AL3 with a Bluetooth tracing technology, and AL3 with wide use of COVID Tracer app. Density plots indicate the lower quartile, median, upper quartile, and distribution of initial outbreak sizes across the set of simulations for each outbreak size and intervention scenario.

This lack of clear effect is likely due to **these scenarios being tested relative to a baseline AL3**. Thus **we expect the additional impact of various interventions to be limited**. For example, the effects we are testing here for Bluetooth tracing technology and wide use of COVID Tracer app are the increase in known close contacts and casual contacts, respectively.

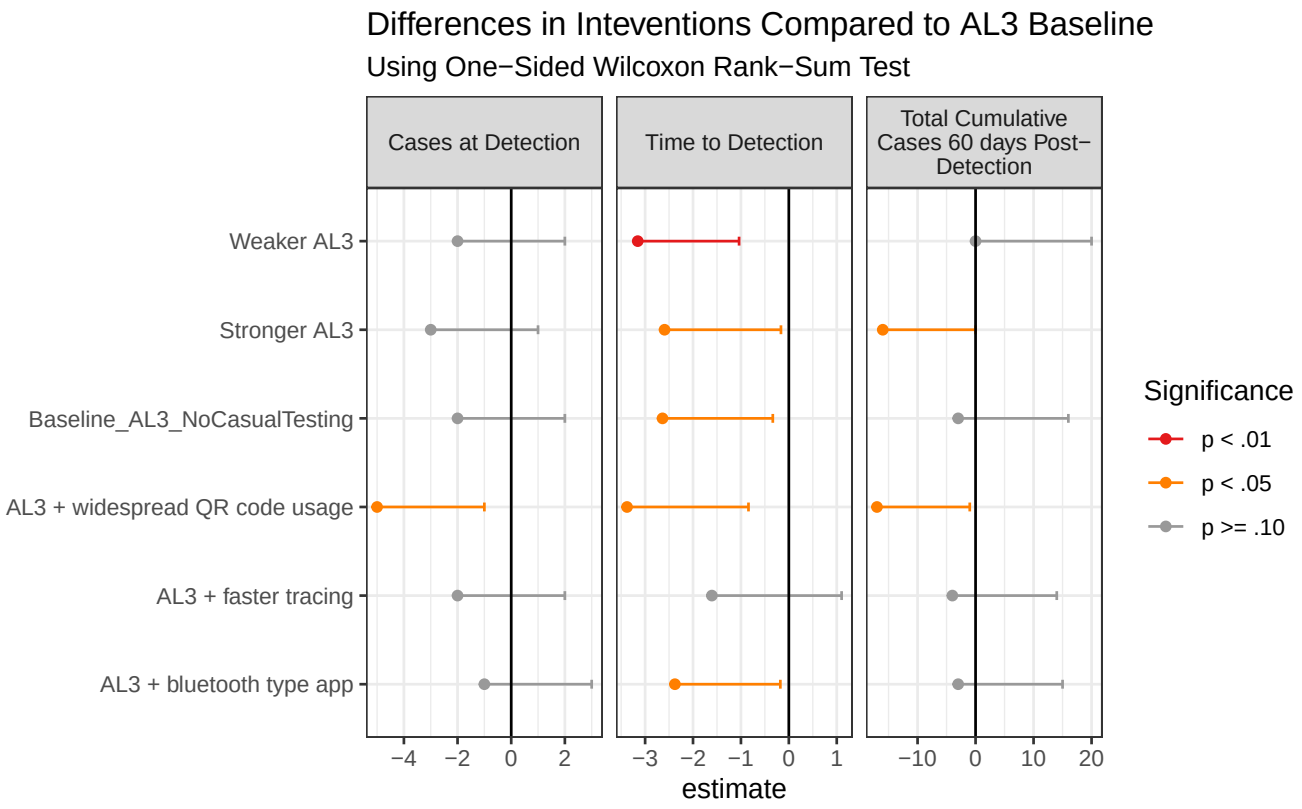

Figure 5. Differences in interventions relative to AL3 baseline. This figure shows the interval estimates of shift relative to baseline of each 'Stamp it out' intervention on time to complete isolation, total cumulative, and total cumulative confirmed cases 60 days post-detection. We see no obvious effects .

These contacts are already reduced 90% at Alert Level 3, so the impact of knowing more of the remaining 10% is small. In contrast, **we would expect the improved contact tracing scenarios to have a greater impact in AL1 or AL2**, where personal interactions are at a higher level, and gains from contact tracing are more substantial.

Comparison with results from the TPM branching process model

In the branching process based Te Pūnaha Matatini report[1], the parameterisation used for both AL2.5 and AL3 is $R_0 = 0.75$. We find that these are in fact both representative of a reasonably strict AL3, with current levels of contact tracing. We do not see an $R_{\text{eff}} < 1$ for AL2.5 transmission reductions. Hence, the total time at AL3 and AL2.5 in that report should be compared with the duration at AL3 in this report.

Furthermore, we find that the impact of Bluetooth tracing and QR code technologies is much lower at AL3 than would be expected, since the main impact of these interventions is on community close and casual contacts, both of which are already severely restricted at AL3. It will be important for future work to look the impact of the different technologies at different alert levels, in order to more accurately parameterise the branching process model before it should be used for estimating these impacts.

Acknowledgements

We would like to acknowledge the help of a various institutions and officials, including Statistics NZ, NeSI, ESR, Treasury, and MoH — particularly Mary van Andel, Trish Priest, and the NCTS team for providing data used in this modelling. This work was funded and supported by Te Pūnaha Matatini and MBIE.

References

1. James, A. *et al.* Modelling support for the continued elimination strategy. Tech. Rep., Te Pūnaha Matatini (2020).
2. James, A., Hendy, S. C., Plank, M. J. & Steyn, N. Suppression and Mitigation Strategies for Control of COVID-19 in New Zealand. Tech. Rep., Te Pūnaha Matatini (2020). DOI: 10.1101/2020.03.26.20044677.
3. Plank, M. J. *et al.* A stochastic model for COVID-19 spread and the effects of Alert Level 4 in Aotearoa New Zealand. Tech. Rep., Te Pūnaha Matatini (2020).
4. QriousNZ. Qrious inter-regional ground travel data. https://www.transport.govt.nz/mot-resources/research-papers/inter-regional-ground-travel-data-from-qrious/ (2019). Accessed: 2020-08-10.
5. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434 (1976).
6. Kendall, D. G. An artificial realization of a simple "Birth-and-Death" process. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **12**, 116–119 (1950).
7. Kiss, I. Z., Miller, J. C. & Simon, P. L. *Mathematics of Epidemics on Networks: From Exact to Approximate Models* (Springer International Publishing, 2017).
8. Bratsun, D., Volfson, D., Tsimring, L. S. & Hasty, J. Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci.* **102**, 14593–14598 (2005).
9. Barrio, M., Burrage, K., Leier, A. & Tian, T. Oscillatory regulation of hes1: discrete stochastic delay modelling and simulation. *PLoS Comput. Biol* **2**, e117 (2006).
10. 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).
11. OpenABM-Covid19 model. https://github.com/BDI-pathogens/OpenABM-Covid19. Accessed: 2020-08-03.
12. Harvey, E., Maclaren, O. J., O'Neale, D. & Wu, D. Asymptomatic cases of covid-19 in nz: a study of data from the auckland august re-emergence case data. Tech. Rep., Te Pūnaha Matatini (2020).
13. Byambasuren, O. *et al.* Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. *medRxiv* DOI: 10.1101/2020.05.10.20097543 (2020). https://www.medrxiv.org/content/early/2020/09/13/2020.05.10.20097543.full.pdf.
14. Harvey, E. & Turnbull, S. 2020-11-20 fluTracking Incidence Calculations. Tech. Rep., Te Pūnaha Matatini (2020).
15. Kucirka, L., Lauer, S., Laeyendecker, O., Boon, D. & Lessler, J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based sars-cov-2 tests by time since exposure. *Annals Intern. Medicine* **173**, 262–267, DOI: https://doi.org/10.7326/M20-1495 (2020).
16. Ma, J. Estimating epidemic exponential growth rate and basic reproduction number. *Infect Dis Model.* **5**, 129–141 (2020).
17. Wallinga, J. & Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. Biol. Sci.* **274**, 599–604 (2007).

Appendix

Initial conditions

'Stamp it out' simulations are simulated by initially running the Baseline 'Prepare for it' scenario and then continuing the simulation after applying the specified intervention, one day after the first case is detected. Since the test seeking in Prepare for it period is a stochastic process, (with a relatively low probability), it is possible for some simulations to proceed without a case ever being detected. These simulations will typically die out with only a small number of cumulative infections. For example, if we begin a batch of 100 simulations for a single set of parameter values, it is possible that only 85 of these simulations will have one or more case that is detected. Since the first *detected* case is a stochastic property of the simulation, the time of first detection and the total size of the outbreak at first detection are not directly controlled for in those simulations that continue, post-detection, as Stamp it out scenarios. Although we begin with the same number of simulations for each intervention scenario, the final number of realisations to which the interventions are applied, and the size of the initial outbreak when the interventions are applied, will differ slightly between the different intervention scenarios. It is therefore necessary to consider whether the initial conditions, *at first detection*, are comparable across the different scenarios. Table 5 and Figure 6 below show the distribution of initial outbreak sizes within each of the scenarios and sets of initial conditions for the Stamp it out simulations.

Scenario	Distribution of different initial outbreak sizes			
	2-10 (n)	11-20 (n)	21-50 (n)	51-100 (n)
Baseline AL3	6 [4,7.5] (23)	14 [11.25,15.5] (14)	34 [25.5,40] (27)	63 [59.5,81.5] (15)
Stronger AL3	4 [3,7] (26)	14 [12.75,17] (20)	35 [26.5,41.5] (26)	69 [63,89] (9)
Weaker AL3	6.5 [3,8] (18)	15 [11.25,17] (22)	36 [26,40] (25)	58 [55,63] (9)
AL3 + faster tracing	5 [3,7] (25)	16 [13,18] (17)	37 [24,44] (25)	74.5 [69,82.25] (10)
AL3 + wide use of COVID Tracer app	6 [3,8] (29)	14.5 [13,16] (20)	36 [28,45.5] (19)	67 [55.5,75] (7)
AL3 + Bluetooth tracing technology	5 [3,8.5] (27)	15 [12,19] (15)	33 [27.35,41] (22)	60.5 [60,70.5] (11)

Table 5. Distribution of initial outbreak sizes within each size band, with the number of simulations in each size band indicated in parentheses (n). Results shown are *median [lower quartile, upper quartile]* unless otherwise stated. The proportions match the likelihood of each size based on Baseline AL1 'Prepare for it' simulations; there are more simulations for the more common outbreak sizes.

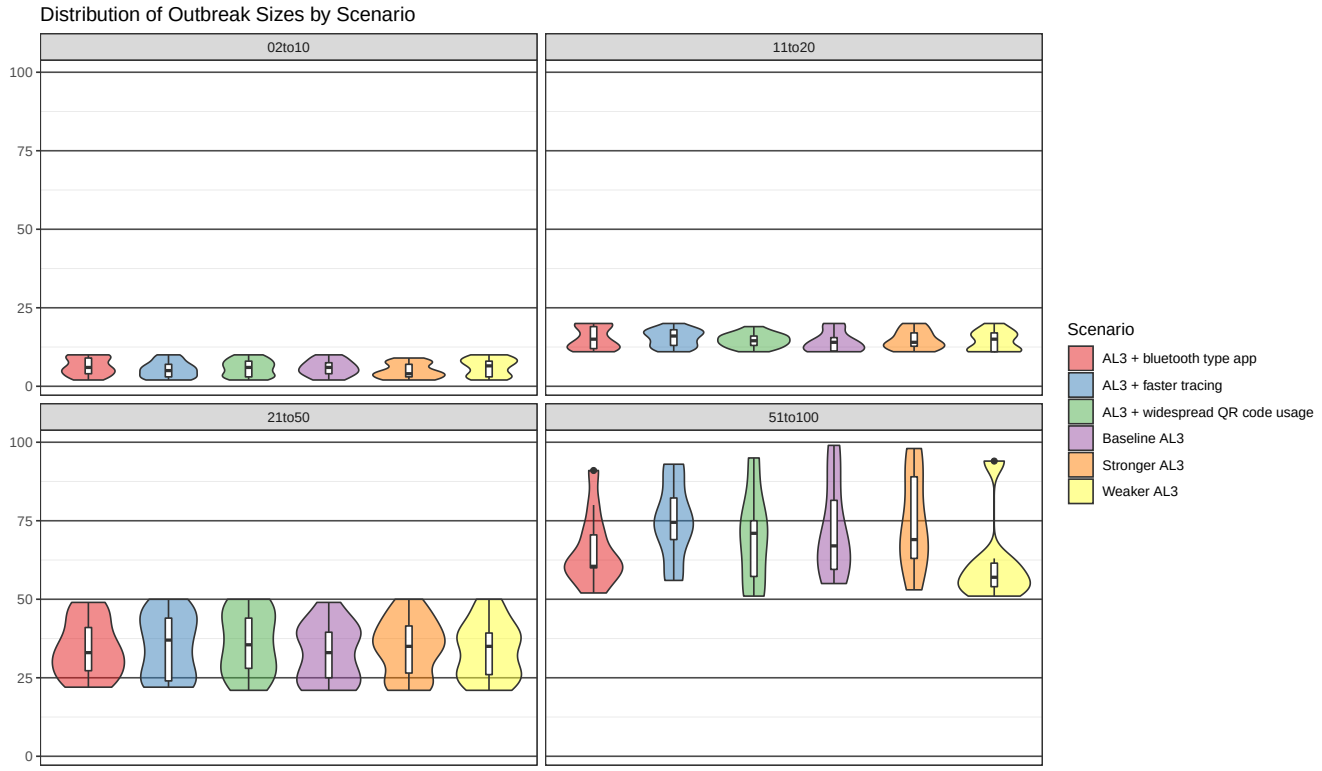

Figure 6. Distribution of Initial Outbreak Sizes by Scenario. This plot shows the distribution of initial outbreak sizes within each outbreak size bin, across the scenarios of Baseline AL3, Stronger AL3, Weaker AL3, AL3 with faster tracing, AL3 with a Bluetooth tracing technology, and AL3 with wide use of COVID Tracer app. The distributions of the initial outbreak sizes used for the Stamp it out simulations are comparable across intervention scenarios and within groups of initial outbreak size though the largest initial outbreaks (size 51–100 cases present at initial detection) has significantly higher variance than the smaller outbreaks and this is not distributed evenly across all the scenarios.