

1 Early intervention is the key to success in COVID-19 control

2 Rachelle N. Binny^{1,4}, Michael G. Baker^{5,6}, Shaun C. Hendy^{3,4}, Alex James^{2,4}, Audrey
3 Lustig^{1,4}, Michael J. Plank^{2,4}, Kannan M. Ridings^{3,4}, Nicholas Steyn^{2,3,4}

- 4 1. Manaaki Whenua, Lincoln, New Zealand.
- 5 2. School of Mathematics and Statistics, University of Canterbury, New Zealand.
- 6 3. Department of Physics, University of Auckland, New Zealand.
- 7 4. Te Pūnaha Matatini: the Centre for Complex Systems and Networks, New Zealand.
- 8 5. Department of Public Health, University of Otago, Wellington, New Zealand
- 9 6. Maurice Wilkins Centre for Molecular Biodiscovery, New Zealand.

11 Abstract

12 New Zealand responded to the COVID-19 pandemic with a combination of border restrictions
13 and an Alert Level system that included strict stay-at-home orders. These interventions were
14 successful in containing the outbreak and ultimately eliminating community transmission of
15 COVID-19. The timing of interventions is crucial to their success. Delaying interventions may
16 both reduce their effectiveness and mean that they need to be maintained for a longer period.
17 Here, we use a stochastic branching process model of COVID-19 transmission and control to
18 simulate the epidemic trajectory in New Zealand and the effect of its interventions during its
19 COVID-19 outbreak in March-April 2020. We use the model to calculate key measures,
20 including the peak load on the contact tracing system, the total number of reported COVID-19
21 cases and deaths, and the probability of elimination within a specified time frame. We
22 investigate the sensitivity of these measures to variations in the timing of interventions and
23 show that changing the timing of Alert Level 4 (the strictest level of restrictions) has a far
24 greater impact than the timing of border measures. Delaying Alert Level 4 restrictions results

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

25 in considerably worse outcomes and implementing border measures alone, without Alert Level
26 4 restrictions, is insufficient to control the outbreak. We conclude that the rapid response in
27 introducing stay-at-home orders was crucial in reducing the number of cases and deaths and
28 increasing the probability of elimination.

29

30 **Keywords:** infectious disease outbreak; non-pharmaceutical interventions; stochastic model;
31 COVID-19; coronavirus; border restrictions

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51 **Introduction**

52

53 An outbreak of COVID-19, a novel zoonotic disease caused by the SARS-CoV-2 virus, was
54 first detected in Wuhan, China in November 2019. The virus spread rapidly to other countries
55 resulting in a pandemic being declared by the World Health Organisation in March 2020.

56 Governmental policy responses to COVID-19 outbreaks have varied widely among countries,
57 in terms of the nature and stringency of policy interventions, how quickly these interventions
58 were implemented (Table S2) (Desvars-Larrive et al., 2020) and their effectiveness at

59 reducing spread of the virus (Flaxman et al., 2020; Hsiang et al., 2020; Binny et al., 2020a).

60 While it is tempting to judge the success of interventions by comparison across jurisdictions,
61 this assessment may be confounded by local context that may influence success, as well as by

62 the fact that policy choices can be driven by the severity of initial outbreaks. Models of

63 disease spread played an important role in the design and timing of interventions, but they

64 can also be used post hoc, to evaluate the effectiveness of those interventions. For example,

65 Flaxman et al. (2020) and Brauner et al. (2020) fitted models of disease dynamics to case

66 count and death data in different countries to estimate the effect of specific non-

67 pharmaceutical interventions on the transmission rate of COVID-19. Modelling studies have

68 indicated the importance of timely interventions for achieving local elimination and averting

69 resurgent waves of COVID-19 in locations like New Zealand and the Australian state of

70 Victoria (Parag et al., 2021).

71

72 In response to the escalating COVID-19 pandemic and the outbreak that was establishing in

73 New Zealand in March 2020, a number of policy interventions were implemented to mitigate

74 risk at the border and reduce community transmission. From 15 March 2020 (11.59pm), all

75 international arrivals were required to ‘self-isolate’ (home quarantine) for 14 days. On 19

76 March 2020, the border was closed to everyone except returning citizens and residents. A
77 system of four alert levels was introduced on 21 March with the Alert Level initially set at
78 Level 2. On 23 March, it was announced that the Alert Level was increasing to Level 3, and
79 that the country would move to Alert Level 4 as of 11.59pm on 25 March, signalling that New
80 Zealand was taking a decisive COVID-19 response that would eventually become an
81 elimination strategy (Baker et al., 2020b). At the time Alert Level 4 came into effect, there had
82 been 315 reported (confirmed and probable) cases. Alert Level 4 stayed in place until 27 April
83 when restrictions were eased to Alert Level 3. On 13 May, after 16 days at Alert Level 3, daily
84 new cases had dropped to three and there was a phased easing into Alert Level 2 (Table 1). The
85 seven weeks spent under stringent Alert Level 3 or 4 restrictions, which included stay-at-home
86 orders (see Appendix Table S4 for full list of measures) alongside systems for widespread
87 testing, contact tracing and case isolation, were effective at reducing transmission (effective
88 reproduction number, $R_{eff} = 1.8$ [95% CI 1.44, 1.94] prior to Alert Level 4; $R_{eff} = 0.35$ [0.28,
89 0.44] during Alert Level 4; Binny et al., 2020b, James et al., 2020c). Daily numbers of new
90 cases declined to zero or one by mid May and the last case of COVID-19 associated with the
91 March outbreak was reported on 22 May. A descriptive epidemiological study by Jefferies et
92 al. (2020) provides a comprehensive account of SARS-CoV-2 transmission patterns, testing
93 patterns, demographic features and disease outcomes in New Zealand during this period. On 8
94 June, after 17 consecutive days with no new reported cases, New Zealand officially declared
95 elimination of community transmission following WHO guidelines (Baker et al., 2020c). This
96 decision is supported by modelling work estimating the probability of elimination at this time
97 (Binny et al., 2020c; Parag et al., 2021; Wilson et al., 2020). Between 22 May and 11 August,
98 the only new cases detected were in international arrivals who were required to spend 14 days
99 in government-managed isolation or quarantine facilities (Baker et al., 2020c). On 9 August
100 2020, New Zealand reached a milestone of 100 days with no community transmission.

101
102
103

Table 1: Dates of implementation for policy interventions during New Zealand’s COVID-19 response. All interventions were implemented at 11.59pm.

Date implemented	Policy Intervention
15 March	14-day self-isolation (i.e. home quarantine) for all international arrivals
19 March	Border closed except to returning residents and citizens
21 March	Alert Level 2
23 March	Alert Level 3
25 March	Alert Level 4
9 April	Mandatory 14-day government-managed quarantine for all international arrivals
27 April	Alert Level 3
13 May	Alert Level 2 (schools and bars remain closed, gathering size limit 10)
18 May	Schools reopen
21 May	Bars reopen
25 May	Gathering size limit increased to 100
8 June	Alert Level 1: all restrictions lifted except border measures

104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119

A comparison of the outcomes of New Zealand’s COVID-19 response with predicted outcomes from hypothetical alternative actions is important for evaluating the effectiveness of the interventions made and to help refine future response strategies. In this work, we first model the factual scenario using New Zealand’s actual intervention timings, then compare this to counterfactual (alternative ‘what if’) scenarios where policy interventions were implemented earlier or later than occurred in reality, to assess what impact this could have had for COVID-19 spread in New Zealand. For each scenario, we simulate a model of COVID-19 spread and compare key measures, including the peak load on the contact tracing system, the cumulative numbers of cases and deaths, and the probability of elimination predicted by the model. Here, we define elimination as there being no active cases (we assume a case remains ‘active’ for 30 days after date of exposure) that could contribute to future community transmission. This definition excludes cases in international arrivals after 9 April 2020 who were required to spend 14 days in government-managed isolation and quarantine facilities. Similar definitions employing 28 days after date of exposure (which is twice the maximum incubation period)

120 have been proposed (Baker et al., 2020a). The date of exposure for an infected individual is
121 often unknown and is instead inferred from date of symptom onset; our 30-day definition
122 provides some buffer for potential error arising during this estimation.

123

124 In particular, we assess how important New Zealand’s decision to move ‘hard and early’ was
125 for the successful elimination of community transmission following the March-April outbreak.
126 To this end, we compare scenarios with different timings until the start of Alert Level 4 to see
127 how these choices could have affected the size of the outbreak. While Alert Level 4 was
128 successful in achieving elimination, the benefits of elimination had to be weighed against the
129 negative impacts of stringent stay-at-home measures, for example job losses, financial
130 insecurity, and disruption to education and economic activity. If careful border management
131 could have avoided the need for a lockdown or reduced its intensity, this approach may have
132 been preferable. For instance, Taiwan’s early border closure, travel restrictions and 14-day
133 quarantine for those entering the country have meant that, to date, Taiwan has avoided a mass
134 lockdown (Summers et al., 2020). We explore whether introducing border restrictions
135 (requiring 14-day self-isolation for all international arrivals) earlier in New Zealand might have
136 been sufficient to eliminate or reduce transmission from international arrivals to the extent
137 where stringent Alert Level 4 restrictions could have been avoided or less restrictive measures
138 been sufficient. Compared to other countries, New Zealand was very quick to close its border
139 to all except returning citizens and residents (Table S2). We explore a scenario where border
140 closure is delayed by 5 days to assess how much larger the outbreak might have been had New
141 Zealand been slower to act. We also consider a scenario where timings of two interventions are
142 altered, to check whether this combination amplifies or counteracts the outcome trends found
143 in scenarios where timings are varied for interventions individually. However we do not
144 perform exhaustive comparisons of all possible combinations of different timings for all

145 interventions. Finally, we consider a scenario with no Alert Level 4/3 restrictions to compare
146 the size of outbreak that New Zealand could have experienced if border restrictions and closure
147 had been the only control measures.

148

149 In this study, we focus on the timing of interventions; we do not explicitly consider the duration
150 of interventions, although this will be investigated in future work. Indeed, the likelihood of
151 elimination was one of the factors taken into account in New Zealand government decision-
152 making concerning the duration of Alert Levels (DPMC, 2020).

153

154 **Methods**

155

156 We simulated a stochastic model of COVID-19 spread in New Zealand (James et al., 2020a;
157 Hendy et al., 2021) under a factual scenario using actual timings for border restrictions,
158 border closure and Alert Level 4, and for counterfactual scenarios in which implementation
159 of these interventions were either delayed or started earlier. Case data were obtained from
160 Ministry of Health, containing arrival dates, symptom onset dates, isolation dates and
161 reporting dates for all international cases arriving in New Zealand between February and June
162 2020. The model is a continuous time branching process that is seeded with internationally
163 imported cases and simulates the numbers of new clinical and subclinical cases that are
164 acquired through local transmission each day (see Appendix for full model specification and
165 Table S1 for list of model parameters). It accounts for delays from infection to symptom
166 onset, and from symptom onset to date of reporting. We assume that the time between an
167 individual becoming infected and infecting another individual (the generation time) follows a
168 Weibull distribution, with a mean and median of 5 days and standard deviation of 1.9 days
169 (Ferretti et al., 2020). The model incorporates individual heterogeneity in transmission rate

170 (e.g. super-spreaders) and assumes that a proportion of clinical cases are undetected by
171 testing and do not get reported, as described in James et al. (2020a). It assumes that
172 subclinical cases are not reported and do not self-isolate. Except for heterogeneity in an
173 individual's reproduction number, individuals are otherwise assumed to be homogenous and
174 the population well mixed. Each clinical case is assumed to have a fixed probability of
175 resulting in fatality, calculated using an infection fatality rate (IFR) of 0.88%; this value was
176 obtained by fitting age-specific COVID-19 IFR estimates from international studies (Verity
177 et al, 2020) to the age distribution of the New Zealand population from 2018 Census data
178 (Statistics New Zealand, 2020).

179

180 We account for three classes of interventions for reducing onward transmission: (i) self-
181 isolation (i.e. home quarantine), simulated as a reduction in an individual's transmission rate
182 relative to their transmission when not isolated; (ii) government-managed isolation and
183 quarantine (MIQ), which we assume is 100% effective at preventing onward transmission;
184 (iii) population-wide control, modelled as a reduction in transmission rate by a factor $C(t)$ at
185 time t , relative to no population-wide control, due to restrictions under each of the four Alert
186 Levels. The relative transmission rate is $C(t)=1$ in the absence of population-wide control and
187 higher Alert Levels correspond to smaller values of $C(t)$ and an associated reduction in
188 effective reproduction number R_{eff} . Reproduction number has a large impact on the
189 probability of extinction in a branching process and is related, via the generation time
190 distribution, to the rate of exponential growth during the early phase of an outbreak (Wallinga
191 & Lipsitch, 2007). Rather than reducing reproduction number in the branching process, the
192 effect of border closure is to reduce the number of seed cases, i.e. the number of new
193 independent branches.

194

195 In each scenario, we kept the duration of each Alert Level (AL) the same as actually occurred,
196 i.e. 33 days at AL4 followed by 16 days at AL3. We explored the following scenarios (see
197 Table 1):

- 198 **0. Border restrictions, border closure and AL4 implemented on actual dates.**
- 199 **1. Early AL4.** Border restrictions and closure implemented on actual dates, and start of
200 AL4 implemented 5 days early.
- 201 **2. Delayed AL4.** Border restrictions and closure implemented on actual dates, and start
202 of AL4:
 - 203 a. delayed by 5 days,
 - 204 b. delayed by 10 days,
 - 205 c. delayed by 20 days.
- 206 **3. Early border restrictions.** Border restrictions 5 days earlier; border closure and AL4
207 on actual dates.
- 208 **4. Delayed border closure.** Border closure delayed by 5 days; border restrictions and
209 AL4 on actual dates.
- 210 **5. Change in timing of AL4, border restrictions and closure:**
 - 211 a. Border restrictions 5 days early and AL4 5 days early; border closure on actual
212 date.
 - 213 b. Border closure and AL4 delayed by 5 days; border restrictions on actual date.
- 214 **6. No AL4.** No AL3 or AL4 implemented; border restrictions and closure on actual dates.

215 Border restrictions, border closure and start of AL4 were all implemented at 11.59pm so we
216 start simulating their effects on the day after their implementation date. For scenarios 0, 1, 2, 3
217 and 6, the model was seeded with the same number of international cases as were reported in
218 case data, using their actual dates of arrival, symptom onset and reporting. For each of these
219 case cases, the date of exposure was estimated backwards from the date of symptom onset. In

220 scenarios where border restrictions were implemented on the actual start date (15 March;
221 Scenarios 0, 1, 2, 4 and 6), the self-isolation dates of international cases were set to the same
222 isolation dates as were actually reported. In all scenarios, prior to 9 April the modelled effect
223 of self-isolation is to reduce an individual's transmission rate to 65% of their transmission when
224 not isolated (following the assumed value used in a model of COVID-19 transmission
225 dynamics by Davies et al., 2020). This reflects some risk of onward transmission for cases self-
226 isolating at home. After 9 April, the model assumes that all international cases are placed in
227 MIQ facilities and do not contribute to local transmission. We also simulated a Poisson-
228 distributed random number of international subclinical cases in proportion to the number of
229 international clinical cases (assuming 1/3 of all cases are subclinical), with arrival and symptom
230 onset dates that were randomly sampled with replacement from the international case data. We
231 assume that these international subclinical cases are not detected and therefore do not self-
232 isolate, but those arriving after 9 April are placed in MIQ and therefore do not contribute to
233 community transmission.

234
235 To simulate border restrictions starting 5 days early (Scenarios 3 and 5a), international cases
236 arriving between the earlier start date and the actual start date (11 - 15 March, inclusive) were
237 assumed to be self-isolated on their date of arrival. To simulate a 5-day delay to border closure
238 (Scenarios 4 and 5b), we delayed the arrival dates (and associated symptom onset, reporting
239 and isolation dates) of international seed cases arriving after 19 March by 5 days. We then
240 allowed for new international cases arriving over these 5 days (e.g. additional non-residents
241 that may have chosen to travel had the border remained open for longer) by seeding an
242 additional Poisson-distributed random number of international cases from 20 March to 24
243 March, with an average daily number of seeded cases equal to the actual average daily number
244 of international cases arriving during the week prior to 19 March (33 international cases per

245 day). These additional seeded cases were assumed to self-isolate on arrival and their delays
246 from arrival-to-symptom onset and arrival-to-reporting were randomly sampled with
247 replacement from the corresponding delays in the actual international case data (obtained from
248 Ministry of Health). We did not attempt to simulate scenarios with delayed border restrictions
249 or earlier border closure because these would have required additional, more substantial
250 modelling assumptions about isolation dates of international arrivals and about the reduction
251 in the volume of international arrivals resulting from border closure. Model predictions would
252 have been highly sensitive to these major assumptions and, without data available to validate
253 them, this would introduce considerable model uncertainty.

254

255 For each scenario, we assessed the following key measures describing the dynamics of a
256 COVID-19 outbreak:

- 257 1. The maximum contact tracing (or health system) load, by calculating the maximum
258 number of daily new reported cases and the date on which this occurred.
- 259 2. The number of daily new reported cases at the end of AL4.
- 260 3. Cumulative number of reported cases and the cumulative number of deaths at the end
261 of the seven week period of Alert Level 3-4 restrictions (i.e. end of AL3).
- 262 4. Probability of elimination, $P(\text{elim})$, 5 weeks after the end of AL3.

263 The first measure is useful for assessing whether the contact tracing or health system capacity
264 would have been exceeded. The second measure indicates the daily incidence of cases after
265 four weeks of the most stringent restrictions under AL4, and at the time when restrictions are
266 eased to AL3 (meaning schools, years 1 to 10, and Early Childhood Education centres can re-
267 open with limited capacity, and non-essential businesses can re-open premises but cannot
268 physically interact with customers; see Appendix Table S4). The third measure quantifies the
269 overall health cost of the outbreak. Given New Zealand's elimination strategy (Baker et al.,

270 2020b), we included the fourth measure to assess the likelihood of achieving elimination of
271 community transmission under the different intervention timings. This is an important
272 consideration because in scenarios resulting in low probabilities of elimination after AL3
273 restrictions are eased, there is a higher risk of cases persisting undetected in the community
274 and sparking a new outbreak under the weaker AL2 restrictions. For example, gatherings of up
275 to 100 people are permitted under AL2, increasing the risk of super-spreading events. If a new
276 outbreak did occur then another lockdown may be required and the overall health cost would
277 be even higher than that of our third key measure.

278

279 In Scenarios 1-5, the model was simulated from 1 February 2020 until a date 5 weeks after the
280 end of AL3 (end dates therefore vary between scenarios depending on Alert Level timings). In
281 Scenario 6, simulations were run up to 1 February 2021 to allow sufficient time for the outbreak
282 to run its full course. We performed 5000 realisations of the model and report the average value
283 of each key measure as well as the interval range within which 90% of simulation results were
284 contained (in square brackets throughout). In the model, $P(\text{elim})$ was calculated as the
285 proportion of all model realisations that resulted in elimination, defined as no remaining cases
286 within 30 days of infection on the date 5 weeks after the end of AL3 (or in Scenario 6 with no
287 AL4/3 restrictions, no remaining cases within 30 days of infection on 18 June 2020, i.e. 5 weeks
288 after the end of actual AL3).

289

290 Simulations were run using previously published estimates of the reproduction number R_{eff} that
291 provided the best fit to actual data before and during AL4 (Binny et al., 2020b). For the period
292 prior to AL4, $R_{\text{eff}} = 1.8$ which reflects the average level of transmission in the period prior to
293 21 March (when there was a heightened public awareness of the risk of community
294 transmission, as the first domestic cases were detected in New Zealand) and during the 2-day

295 periods in AL2 and AL3. During AL4, $R_{eff} = 0.35$. Under the Alert Levels 3, 2 and 1, which
296 followed the lockdown, the daily numbers of new cases were too low to obtain reliable
297 estimates of the effective reproduction number R_{eff} from these data using standard techniques
298 (e.g. see Obadia et al. (2012) for a review of methods). Instead we simulated the model for
299 assumed values of $R_{eff} = 0.95, 1.7$ and 2.4 , respectively. The estimate of $R_{eff} = 2.4$ at AL1
300 (minimal restrictions, schools and businesses fully open and no limits on gathering size) is in
301 line with estimates reported in Plank et al (2020), for the pre-lockdown period of New
302 Zealand's August-September outbreak, when community transmission was considered to be
303 eliminated and the country was at AL1. Plank et al (2020) obtained these estimates using two
304 independent methods: (i) performing Monte Carlo reconstructions of epidemiological
305 transmission trees using contract tracing data (James et al., 2020c), and (ii) fitting the Hendy
306 et al. (2021) stochastic branching process model to daily reported case count data. The value
307 of R_{eff} at AL2 would likely be lower than at AL1 but greater than one due to relatively high
308 activity levels and contact rates as stay-at-home orders are lifted and public venues, businesses
309 and schools re-open. The estimate of $R_{eff} = 1.7$ at AL2 is in the range of estimated values for
310 the pre-lockdown period of the March-April outbreak given in Plank et al (2020). The R_{eff} for
311 AL3 was chosen to be less than 1, however we tested the sensitivity of our results to using a
312 value greater than one. For the scenario with no stringent Alert Level restrictions (Scenario 6),
313 we simulated the model using $R_{eff} = 1.8$ for the entire period (i.e. the same value as was used
314 in all Scenarios for the period prior to AL4). This value is lower than that of AL1, reflecting a
315 likelihood of substantial behaviour change and heightened public awareness of the risks of
316 community transmission, even in the absence of any alert level restrictions.

317

318 *Sensitivity analyses*

319

320 We further investigated how varying the length of the delay (in days) until the start of AL4
321 (cf. the delays chosen in Scenario 2) affected key measures. Similarly, we assessed the effect
322 on the key measures of introducing border restrictions 10 days early (cf. 5 days early in
323 Scenario 3). We also tested the sensitivity of our results to different choices of relative
324 transmission rate $C(t)$ (corresponding to changes in R_{eff}) under AL3, in particular for a less
325 effective AL3 ($C(t)=0.46$; $R_{eff}=1.1$) and a more effective AL3 ($C(t)=0.30$; $R_{eff}=0.7$). Previous
326 studies have considered the sensitivity of the branching process model predictions to
327 variations in model parameters whose values are uncertain or context-specific (James et al.,
328 2020a; Binny et al, 2020b; Binny et al, 2020c). Multiple dependencies exist between certain
329 model parameters and the overall population reproduction number R_{eff} , as described in James
330 et al (2020a). There is further dependency between R_{eff} and the probability of elimination
331 $P(elim)$ such that a reduction in R_{eff} corresponds to an increase in $P(elim)$. James et al (2020a)
332 compared model outputs for different values of the proportion and relative infectiousness of
333 subclinical infections, the individual heterogeneity in transmission rate, and the mean
334 generation time. Changes in these parameters that resulted in a change in the overall
335 population reproduction number, R_{eff} , caused a corresponding change in the outbreak
336 trajectory. However, if R_{eff} was set to a fixed value the model was robust to changes in these
337 parameters. Increasing the individual heterogeneity in transmission rate increases the
338 variation between independent realisations of the model and increases $P(elim)$ (Lloyd-Smith
339 et al, 2005). The best-fit R_{eff} estimates for AL4 were also relatively insensitive to changes in
340 model parameters (Binny et al., 2020b).

341

342 **Results**

343

344 *Scenario 0*

345 To check that the model could accurately replicate the outbreak, we first simulated our model
346 under a factual scenario with border restrictions, border closure and AL4 implemented on the
347 dates they actually occurred. The predicted dynamics of daily new reported cases were a very
348 good visual match to observed daily case data (Fig. 1) and predicted key measures showed
349 good agreement with the values that were actually observed (Table 2, bold text). After moving
350 into AL4, the model prediction and the actual number of daily new reported cases both levelled
351 off at 70-80 for around one week before case numbers started to decline (Fig. 1). In actual case
352 data, the maximum of 84 new cases per day was observed at the start of this flat-topped peak,
353 while our model predicted a similar maximum (80 [67, 99] new cases per day) occurring 6 days
354 later. By the end of AL3, the model predicted similar cumulative totals to the 1503 cases and
355 22 deaths actually reported. Five weeks after AL3 restrictions were relaxed, elimination of
356 community transmission of COVID-19 was achieved in 66% of model simulations, giving
357 $P(\text{elim})=0.66$ (Table 2). As we discuss later, higher estimates of probability of elimination
358 reported when New Zealand declared elimination on 9 June were obtained by making use of
359 additional information from case data on the number of consecutive days with zero new
360 reported cases (Binny et al., 2020c). In the following counterfactual scenarios with alternative
361 timings of interventions, we use Scenario 0 as a baseline for comparing key measures.

362

363 *Scenario 1: Early AL4*

364 Under a scenario where AL4 was implemented 5 days earlier (one day after border closure),
365 the model predicts slightly lower values for most key measures than were actually observed:
366 daily new cases peaked at a lower level of 69 [61, 79] cases around 26 March and at the end of
367 AL4 had dropped to a similar level of 4 new cases per day as was actually observed (Fig. 1).
368 By the end of the 7 weeks of AL4/3 it predicts approximately 500 fewer cases in total and 10
369 fewer deaths (Table 2; Scenario 1 cf. Scenario 0). However, this estimate should be taken with

370 caution because of the small numbers of daily cases and fine-scale variations involved: for
371 instance, whether an outbreak occurred in an aged care facility or not. Five weeks after AL3,
372 the probability of elimination was 63%, slightly lower than in Scenario 0. This counter-intuitive
373 result is due to the presence of an international case in the data that had an arrival date prior to
374 the start of AL4 (25 March) but a much later symptom onset date near the end of AL4. In
375 Scenario 0, when international cases are seeded in the model using the actual dates of arrival,
376 symptom onset, reporting and isolation from the case data (see Appendix for details), this
377 individual's peak transmission rate occurs during AL4. However, in Scenario 1, the earlier start
378 to AL4 means that the individual is instead most infectious during AL3, where R_{eff} is higher,
379 so this individual infects more people, on average, in this scenario than in Scenario 0. Similarly,
380 any simulated subclinical infections with the same arrival and symptom onset dates (which are
381 sampled with replacement from the international case data) will also be most infectious during
382 AL3. These subclinical infections do not appear in the numbers of reported cases but will
383 reduce the probability of elimination. If this international case outlier is excluded from the data,
384 the model predicts a very similar probability of elimination in both scenarios.

385

386
387
388
389

Table 2: Key measures from alternative scenarios of early or delayed implementation of policy interventions: the maximum number of daily new reported cases, date on which the peak occurs, the number of daily new cases at the end of the simulated Alert Level 4 period, the cumulative number of cases and the total number of deaths at the end of the simulated 7 week period of Alert Level 4/3 restrictions (dates given in 'AL3 ends' and footnotes). For each measure, except $P(\text{elim})$, the mean value from 5000 simulations is reported alongside the interval range, in parentheses, in which 90% of simulations results are contained.

Scenario	Border self-isolation	Border closed	AL4 starts	AL3 ends	Max. new daily cases	Date of peak	New daily cases at end of AL4	Cumulative reported cases	Total deaths	$P(\text{elim})$, 5 weeks after end of AL3
Actual	15-Mar	19-Mar	25-Mar	13-May	84	25-Mar	3	1503	22	-
0	15-Mar	19-Mar	25-Mar	13-May	80 [67, 99]	31-Mar [26-Mar, 02-Apr]	4 [1, 8]	1448 [1208, 1796]	23 [14, 33]	0.66
1	15-Mar	19-Mar	20-Mar	8-May	69 [61, 79]	26-Mar [25-Mar, 26-Mar]	4 [1, 7]	953 [839, 1132]	14 [8, 21]	0.63
2a	15-Mar	19-Mar	30-Mar	18-May	108 [84, 139]	06-Apr [01-Apr, 09-Apr]	7 [3, 12]	2373 [1918, 2999]	39 [26, 55]	0.57
2b	15-Mar	19-Mar	4-Apr	23-May	179 [137, 233]	11-Apr [09-Apr, 14-Apr]	12 [6, 19]	3988 [3161, 5115]	67 [48, 91]	0.38
2c	15-Mar	19-Mar	14-Apr	28-May	503 [382, 661]	21-Apr [20-Apr, 23-Apr]	34 [22, 49]	11534 [8854, 15048]	200 [147, 266]	0.07
3	10-Mar	19-Mar	25-Mar	13-May	79 [67, 97]	31-Mar [26-Mar, 02-Apr]	4 [1, 8]	1422 [1194, 1765]	22 [14, 32]	0.66
4	15-Mar	24-Mar	25-Mar	13-May	91 [77, 110]	01-Apr [31-Mar, 02-Apr]	5 [1, 9]	1594 [1359, 1934]	25 [16, 35]	0.55
5a	10-Mar	19-Mar	20-Mar	8-May	68 [61, 79]	26-Mar [25-Mar, 26-Mar]	4 [1, 7]	941 [826, 1119]	14 [8, 21]	0.63
5b	15-Mar	24-Mar	30-Mar	18-May	120 [97, 152]	06-Apr [01-Apr, 08-Apr]	7 [3, 12]	2501 [2069, 3121]	41 [28, 56]	0.53
6	15-Mar	19-Mar	-	-	47592 [47240, 47962]	14-Jun [11-Jun, 17-Jun]	1127 [841, 1492] ¹	60443 [45761, 79201] ² ; 1,812,900 [1,809,600, 1,816,300] ³	1187 [891, 1565] ² ; 31905 [31606, 32204] ³	0.00 ⁴

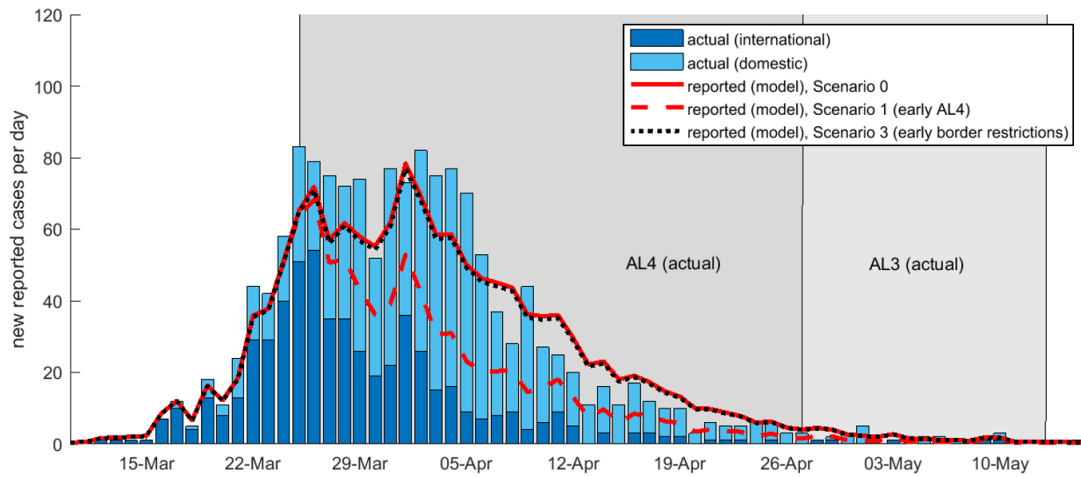
390
391
392
393
394

¹Evaluated on 27th April 2020 (end of actual AL4); continues to increase after this date.

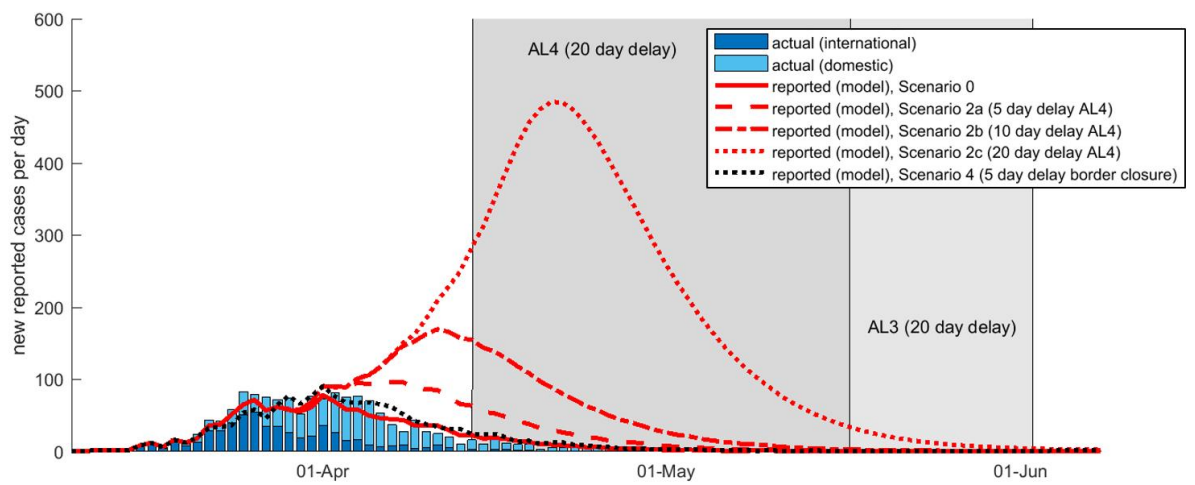
²Evaluated on 13th May 2020 (end of actual AL3); continues to increase after this date.

³Evaluated at end of outbreak. Across all realisations, the outbreak had run its full course by approx. October 2020, on average, and the last case reported by 20 December 2020 at the latest.

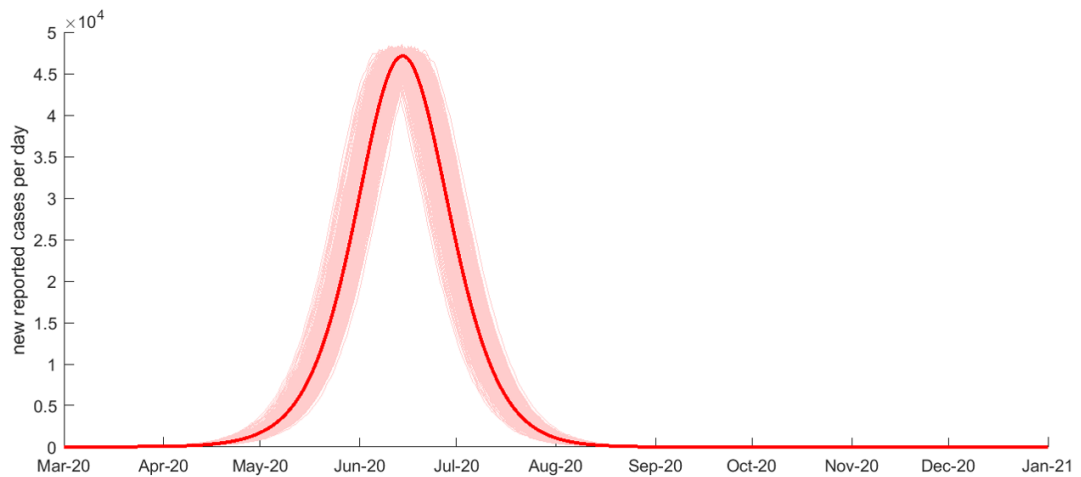
⁴Evaluated on 18th June 2020 (5 weeks after end of actual AL3).



395



396



397

398

399

400

401

402

403

404

Figure 1: Effect of alternative timings of interventions on the trajectory of the outbreak. Number of new reported cases per day predicted by the model alongside observed reported domestic (light blue bars) and international cases (dark blue bars) (data source: MoH). Model simulated for interventions implemented on their actual start dates and for alternative scenarios with different timings of AL4, border restrictions or border closure **Top:** Scenarios with AL4 started 5 days early (border restrictions and closure on actual start dates) (red dashed; Scenario 1) compared with a scenario where border restrictions were implemented five days early (border closure on actual start date) (red dotted; Scenario

405 3). **Middle:** Delayed start to Alert Level 4 (delays of 5, 10 and 20 days; Scenarios 2a-c; red broken
406 lines) (with border restrictions and closure on actual start dates). Five day delay to border closure
407 (with border restrictions and AL4 on actual start dates) (black dotted line; Scenario 4). **Bottom:** No
408 AL4/3 restrictions (border restrictions and closure on actual start dates; Scenario 6) results in an
409 uncontrolled outbreak; faint red lines show the outbreak in individual realisations of the model, bold
410 red line is the average over all 5000 simulations. Note, y-axis scale differs between figures.

411

412 *Scenario 2: Delayed AL4*

413 Delaying the move into Alert Level 4 would have led to a higher peak in daily new cases, and
414 greater cumulative totals of cases and deaths. For a delay of 20 days (Scenario 2c), the outbreak
415 would have reached a considerably higher maximum of close to 500 daily new cases (cf. 80
416 cases in Scenario 0; Table 2, Fig. 1 and Fig. 2). This number would certainly have overwhelmed
417 the contact tracing system, which was already pushed close to capacity in places by the 70-80
418 daily new cases in late March (Verrall, 2020). After a week in AL4, case numbers would start
419 to decline and by the end of the 4 weeks in AL4 daily new cases would still have been as high
420 as 34 [22, 49] (close to the actual number of domestic daily reported cases when New Zealand
421 went into AL4 on 25 March). By the end of the 7 week period of stringent restrictions (i.e. end
422 of AL3) the incidence would have dropped to approximately 4 new cases per day (Fig. 1), but
423 the cumulative total could have climbed to 11,534 [8854, 15048] reported cases and 200 [147,
424 266] deaths, substantially more than Scenario 0 and the 1,503 cases and 22 deaths actually
425 reported on 13 May. Additionally, the probability of elimination 5 weeks after the end of AL3
426 was only 7%, much lower than in Scenario 0.

427

428 *Scenario 3: Early border restrictions*

429 We next investigated a scenario where border restrictions were put in place 5 days earlier, but
430 border closure and AL4 were started on their actual dates. Border restrictions would therefore
431 have been in place for 9 days (cf. actual 4 days) before the border was closed. Our model

432 predicted this would have had very little impact for the initial trajectory (Fig. 1) or eventual
433 outbreak size, with values for all key measures similar to those in Scenario 0 (Table 2). This
434 suggests that key measures are more sensitive to varying the timing of AL4 than to the timing
435 of border restrictions. In reality, out of the 563 international cases who arrived prior to the start
436 of MIQ and could have contributed to local transmission, only 78 (14%) arrived before border
437 restrictions were implemented on 15 March and were not required to self-isolate. Furthermore,
438 out of these 78 cases, 52 arrived between 10 and 15 March and 19 of these were reported to
439 have voluntarily self-isolated immediately on arrival (the model simulates these 19 cases as
440 being self-isolated on arrival in all scenarios). Therefore, under this scenario, early self-
441 isolation requirements reduce transmissions by a factor $c_{iso}=0.65$ (Appendix Table S1) for only
442 an additional 33 international cases. This reduction is not sufficient to prevent an outbreak, nor
443 does it reduce transmission to an extent where AL4/3 restrictions would not have been
444 necessary to control the outbreak. In contrast, starting AL4 early (Scenario 1) has a stronger
445 effect because it substantially reduces transmissions by a factor $C(t)=0.15$ (Appendix Table
446 S1) for all locally acquired cases (clinical and subclinical) and international cases.

447

448 *Scenario 4: Delayed border closure*

449 Under a scenario where closure of the border (to all except returning residents and citizens)
450 was delayed by 5 days (24 March; 9 days after border restrictions and 1 day before AL4), our
451 model predicted slightly worse outcomes, on average, for key measures compared to Scenario
452 0. However, due to the stochasticity of individual simulations, the range of key measures
453 always had overlap with the Scenario 0 values and actual values, suggesting a 5 day delay to
454 border closure alone would not have made a significant difference. A delayed border closure
455 did, however, have a greater impact for the probability of elimination 5 weeks after AL3
456 restrictions were relaxed, which was only 55%, compared to 66% in Scenario 0. This reduced

457 probability of elimination is partly due to the additional international clinical cases (captured
458 in the key measures of reported cases) and international subclinicals (not captured in reported
459 cases) arriving prior to the delayed border closure. It is also likely affected by the international
460 case outlier with the pre-MIQ arrival date and late onset date, discussed above.

461

462 *Scenario 5: Change in timing of AL4, border restrictions and closure*

463 After varying the timing of interventions individually, we considered the effects of varying
464 timings for two interventions. If border restrictions were implemented 5 days early and AL4
465 came into effect 5 days early (Scenario 5a), this would have led to outcomes very similar to
466 those predicted in Scenario 1 (where only AL4 started early) (Table 2). This again suggests
467 that results are more sensitive to changes in timing for the start of AL4 than to an earlier start
468 to border restrictions.

469

470 The outcomes from delaying both border closure and the start of AL4 by 5 days (Scenario 5b)
471 were similar to a 5-day delay to AL4 (Scenario 2a), though slightly worse due to the combined
472 effect of delays to both interventions. Therefore, the timing of AL4 also appears to have a
473 greater impact than timing of border closure. Compared to the factual Scenario 0, daily new
474 cases would have reached a larger maximum of 120 [97, 152] cases on 6 April and by the end
475 of the 7 week period in AL4/3 there would have been close to 1,050 more cases in total and
476 nearly 20 more deaths (Table 2). The probability of elimination 5 weeks after AL3 would have
477 also been reduced to 53%.

478

479 *Scenario 6: No AL4*

480 Finally, we explored the impact of only having border restrictions and border closure in place,
481 but without implementing AL4/3. Under this scenario, the international cases who arrived prior

482 to 9 April and were either in self-isolation or were not isolated have a chance of seeding an
483 outbreak which, without AL4/3 measures to reduce R_{eff} below one, leads to community
484 transmission and a large uncontrolled outbreak. New Zealand would have seen close to 1127
485 [841, 1492] new cases per day by 27 April, the date on which New Zealand moved from AL4
486 to AL3 in reality. By 13 May (the date on which New Zealand moved from AL3 to AL2), there
487 could have been over 60,000 cumulative reported cases and over 1100 deaths. New cases would
488 have continued to increase, reaching a peak of 47,592 [47,240, 47,962] daily new cases on 14
489 June (Table 2). By the end of the outbreak, around October 2020 on average, there could have
490 been over 1.81 million reported cases in total and 31,905 [31,606, 32,204] deaths. No
491 simulations resulted in elimination by 18 June (5 weeks after end of actual AL3), indicating a
492 0% chance of COVID-19 having been eliminated by this time, compared to the 66% chance on
493 this date in Scenario 0. This is an important result demonstrating that border measures alone
494 would have been insufficient to prevent a serious outbreak from occurring and that stringent
495 AL4/3 restrictions were necessary to have a chance of eliminating community transmission.

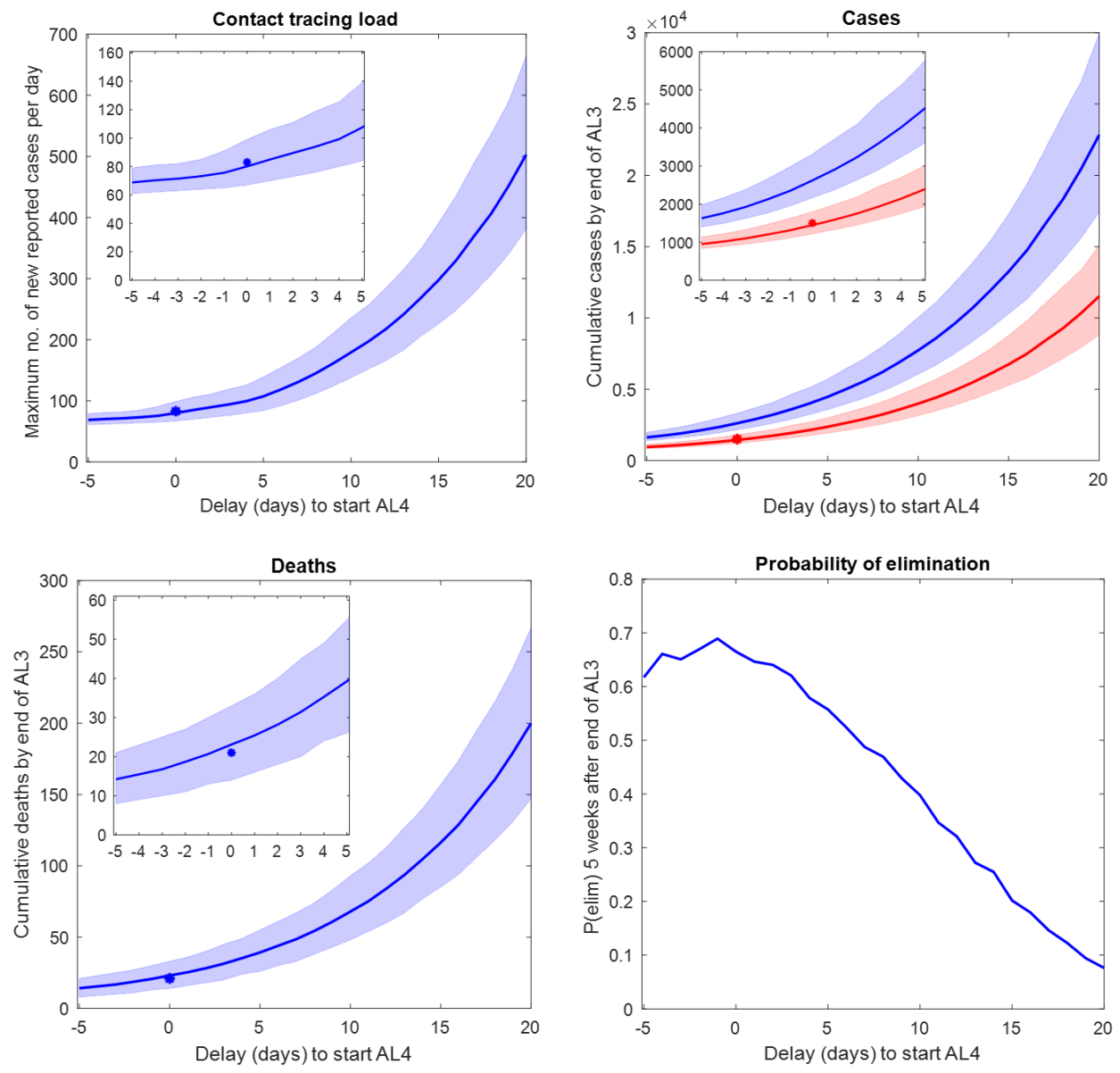
496

497 *Sensitivity analysis*

498 We assessed the effect that different lengths of delay (in days) until the start of AL4 (Fig. 2)
499 had on key measures: maximum load on contact tracing system; cumulative total reported
500 cases; total infected cases (including both clinical and subclinical); total deaths at end of AL3;
501 and probability of elimination 5 weeks after the end of AL3. Measures of numbers of cases and
502 deaths increased exponentially with increasing delay to AL4, emphasising the importance of
503 acting quickly to reduce the risk of large outbreaks arising. Probability of elimination decreased
504 linearly with increasing delays to AL4. Counter-intuitively, earlier starts to AL4 slightly
505 reduced the probability of elimination; again, this is caused by the international case outlier

506 discussed previously. If the outlier is excluded from the international case data, the predicted
 507 probability of elimination is insensitive to AL4 starting 1 to 5 days early.

508
 509



510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520

Figure 2: Sensitivity of predicted cases and deaths to varying the delay until start of Alert Level 4, up to a maximum delay of 20 days. A negative delay of -5 days represents starting Alert Level 4 5 days early (20th March 2020). Border restrictions and closure were implemented on the same dates as actually occurred. **Top, left:** Maximum load on contact tracing system (measured as maximum number of new reported cases per day) predicted by the model (blue line) and actual maximum number of daily reported cases (asterisk). **Top, right,:** Cumulative number of infected individuals (both clinical and sub-clinical) (blue line) and reported cases (red line) predicted by the model and actual number of reported cases (red asterisk). **Bottom, left:** Cumulative number of deaths predicted by the model (blue line) and actual number (blue asterisk). **Bottom, right:** Probability of elimination, $P(\text{elim})$, 5 weeks after the end

521 of AL3. Shaded regions indicate the interval range in which 90% of simulation results are contained.
522 Note, y-axis scale differs between figures. Insets show close-ups of results for delays from -5 to 5 days.

523

524

525 Introducing border restrictions ten days earlier still results in an outbreak and gives very similar
526 results to Scenario 3 (5 days early), with a maximum of 77 [65, 94] new daily cases, 1385
527 [1166, 1706] cumulative reported cases at the end of AL3, $P(\text{elim}) = 0.68$, and other measures
528 the same as in Scenario 3. With border restrictions ten days earlier, the only difference
529 compared to Scenario 3 is that an additional 16 cases who arrived between 5 and 10 March
530 have their transmission rates reduced (a further 2 cases arriving in this 5-day period were
531 voluntarily self-isolated on arrival in reality, so are simulated with self-isolation on arrival in
532 all scenarios). This has little impact on the overall contribution to local transmission by all
533 international cases who arrive prior to the start of MIQ, providing further support that border
534 restrictions alone are less effective at reducing community transmission compared to stringent
535 Alert Level restrictions.

536

537 We also tested the sensitivity of all key measures to using different values of R_{eff} under AL3
538 (Table S3; $R_{\text{eff}} = 1.1, 0.95$ and 0.7). Different choices of AL3 R_{eff} had very little effect on
539 predicted cumulative totals of cases at the end of AL3 and no effect on total deaths at end of
540 AL3. However, the predicted probability of elimination was sensitive to varying AL3 R_{eff} ; for
541 all scenarios, assuming a lower $R_{\text{eff}} = 0.7$ (more effective AL3) gave a $P(\text{elim})$ that was
542 approximately 0.14 higher than with $R_{\text{eff}} = 0.95$, while a higher $R_{\text{eff}} = 1.1$ (less effective AL3)
543 reduced $P(\text{elim})$ by approximately 0.07. This did not affect our general conclusion that the
544 timing of AL4 has a greater impact on key measures than timing of border measures.

545

546

547 **Discussion**

548

549 New Zealand's decision to act quickly and to implement stringent restrictions to reduce SARS-
550 CoV-2 transmission meant that, to date, New Zealand has experienced amongst the lowest
551 mortality rates reported worldwide (Kontis et al., 2020). On 8 June 2020, nearly 11 weeks after
552 AL4 was initiated, New Zealand declared elimination of COVID-19. Over the course of the
553 March-April outbreak, a total of 1504 cases and 22 deaths were reported before elimination
554 was achieved. Our results suggest that the timing of Alert Level 4 is a much stronger driver of
555 reductions in daily new cases than timings of border restrictions and closure. This finding
556 makes sense because the effect of AL4 in the model is to greatly reduce R_{eff} for all cases,
557 domestic and international arrivals, to 0.35, while border restrictions reduce the delay until case
558 isolation of international cases only (i.e. international cases have their transmission rates
559 reduced earlier) and border closure reduces the daily numbers of international cases only. Out
560 of the scenarios we considered, an earlier start to AL4 by 5 days resulted in the greatest
561 reduction in numbers of cases and deaths, with approximately 500 fewer cases in total and 10
562 fewer deaths. However, in reality, the rapid escalation of the COVID-19 situation in mid-March
563 may have made an earlier start to AL4 impractical and would have allowed less time to prepare
564 for ongoing provision of essential services under AL4.

565

566 Introducing border restrictions requiring 14-day self-isolation for international arrivals earlier
567 than 15 March would have been unlikely to have much impact on the trajectory of New
568 Zealand's March-April outbreak, unless such measures were started prior to the first case on
569 26 February and used methods that were particularly effective (notably full MIQ). The 563
570 international cases arriving between 15 March and 9 April were already required to self-isolate;
571 had border restrictions been in place prior to the arrival of New Zealand's first case, this would

572 have required self-isolation for, at most, an additional 56 international cases (22 cases who
573 arrived prior to 15 March self-isolated voluntarily immediately on their arrival). In mid-March,
574 there was a lower global prevalence of COVID-19 and between 2 and 12 cases arrived at the
575 border each day in the week prior to 15 March. With a higher global prevalence and
576 correspondingly higher numbers of international cases arriving per day, earlier implementation
577 of border restrictions may have had a greater impact than our model predicted for this outbreak.
578 Self-isolation is less stringent than MIQ and relies heavily on public compliance. Without
579 additional safety nets, such as official monitoring and support for people who are self-isolating,
580 there is a greater risk of the virus spreading into the community than in MIQ facilities. For
581 example, risk of non-compliance may be higher for individuals who are concerned about loss
582 of income (Bodas & Peleg, 2020). Without Alert Level restrictions in place to require strong
583 community-wide social distancing, any infected individuals who do not self-isolate effectively
584 are more likely to spark an outbreak. Self-isolation restrictions for international arrivals can
585 therefore reduce the frequency of cases leaking into the community (James et al, 2020c) but
586 are unlikely to be sufficient to prevent an outbreak entirely, unless additional measures are also
587 put in place.

588

589 Delaying border closure by 5 days could have led to a slightly larger outbreak, but not as large
590 as if AL4 had been delayed by 5 days. The full effect on local transmission potential of the
591 additional international cases expected under a delayed border closure was partially dampened
592 because international cases arriving after 9 April were still placed in MIQ and assumed not to
593 contribute to community transmission. If the timing of this MIQ policy was also delayed, a
594 larger outbreak may have occurred, but we did not model such a scenario here. If the start of
595 AL4 had been delayed by 20 days, our results suggest New Zealand could have experienced
596 over 11,500 reported cases and 200 deaths, reducing the chance of elimination to only 7%. As

597 with other severe viral disease, the infection fatality risk for COVID-19 is greater for Māori
598 and Pacific peoples (close to 50% higher for Māori than for non-Māori) (Steyn et al., 2020;
599 Wilson et al., 2012; Verrall et al., 2010). Therefore, in scenarios resulting in significantly
600 higher numbers of COVID-19-related deaths (e.g. Scenario 2c), Māori and Pacific
601 communities would likely have been disproportionately affected, however a population-
602 structured model would be required to assess this consequence in detail. Delaying AL4 would
603 have also increased the chance of a longer lockdown period being required to reduce daily new
604 case numbers to low levels. With a 20 day delay to AL4, New Zealand could still have been
605 experiencing close to 35 new reported cases per day at the end of AL4. While in reality, a 33-
606 day period in AL4 was sufficient to reduce daily new cases to below 10, and the Government
607 announced an easing to AL3, these higher case numbers predicted for a delayed start to AL4
608 may have motivated an extension to the lockdown to allow more time for cases to drop below
609 a safe threshold.

610

611 In terms of the key measures we considered, the counterfactual scenario with no AL4/3
612 restrictions (Scenario 6) had disastrous outcomes, including close to 2 million reported cases
613 and tens of thousands of deaths. This demonstrates that: 1) under the conditions (e.g. level of
614 pandemic preparedness and global COVID-19 prevalence) particular to New Zealand's March
615 – April outbreak, border restrictions and border closure alone would not have been sufficient
616 to control the outbreak; and 2) New Zealand's national lockdown restrictions in combination
617 with its border management, rapid testing and contact tracing, were effective measures that
618 prevented a considerably larger and more prolonged outbreak (i.e. Scenario 6) from occurring
619 during that period.

620

621 Our model uses a value of $R_{eff}=0.35$ during AL4, which was estimated by Binny et al (2020b)
622 by fitting the model to case data over the entire period under AL4, and is consistent with a later
623 estimate of R_{eff} from reconstructions of the epidemiological tree (James et al., 2020c). This is
624 a relatively low value of R_{eff} compared to other countries who implemented interventions
625 roughly equivalent to AL4 (Flaxman et al, 2020; Binny et al, 2020a). A combination of a highly
626 effective social distancing in AL4, fast contact tracing, effective case isolation, and the fact
627 that the outbreak occurred at the end of the Southern hemisphere summer, likely contributed to
628 this low R_{eff} (James et al, 2020b). Disentangling the effects of different interventions introduced
629 in quick succession is challenging (see e.g. Flaxman et al., 2020). For this reason, we apply a
630 single R_{eff} value for the period prior to AL4, rather than different values for the period prior to
631 21 March and the 2-day periods under AL2 and AL3. Due to the delays from exposure to
632 symptom onset and to testing/reporting, AL2-3 likely contributed to the initial flattening in the
633 curve of New Zealand's daily reported cases during the early phase of AL4. In the late phase
634 of AL4, these lagged effects of AL2-3 would no longer be reflected in daily reported cases.
635 Because our model explicitly accounts for the infection-to-onset and reporting delays, and our
636 best-fit values were obtained by fitting to case data from the entire AL4 period, our estimate
637 $R_{eff}=0.35$ for AL4 effectiveness is independent of the lagged effects of earlier interventions.

638

639 For scenarios where the load on contact tracing exceeded system capacity (e.g. Scenario 2c
640 with a maximum 500 daily new cases), this effect would have likely resulted in longer delays
641 to isolation of cases and a higher R_{eff} . We did not attempt to model this potential feedback effect
642 and so our results for scenarios where contact tracing system capacity is exceeded may
643 underestimate the outbreak size. Our model assumed a relatively high and constant proportion
644 (75%) of clinical cases are detected and reported. In reality, this proportion can vary over time
645 as testing and contact tracing policies are revised, or as contact tracing and health systems

646 become overloaded in a large uncontrolled outbreak (e.g. Scenario 6). This makes it difficult
647 to benchmark predicted case numbers against empirical data from outbreaks in other countries,
648 where testing and contact tracing regimes may differ from New Zealand. Infection fatality rates
649 also vary between different countries and over time, for example fatality rates can decrease as
650 new knowledge, treatments and technologies become available. For simplicity, our model
651 assumed a constant IFR of 0.88%, which is within the range of IFR estimates reported for other
652 countries and very close to the median IFR reported for countries with COVID-19 mortality
653 rates that are less than the global average (Ioannidis, 2020). We did not attempt to directly
654 model the burden of COVID-19 on the healthcare system (e.g. numbers of cases requiring
655 hospitalisation or intensive care), or the effects of an overwhelmed healthcare system. Once
656 numbers of daily new cases requiring hospitalisation or ICU admission exceed New Zealand's
657 healthcare system capacity, this could result in increased fatality rates and considerably more
658 deaths (Hendy et al., 2021). These effects would have been most pronounced under the scenario
659 with no AL4/3 restrictions.

660

661 While we report average values for outbreak dynamics, each individual realisation of the
662 stochastic model can deviate (sometimes widely) from the average behaviour. When case
663 numbers are small, as they were in New Zealand, the predicted dynamics are particularly
664 sensitive to fine-scale variations. While $R_{eff} < 1$ means that an outbreak will eventually die out,
665 on average, it is still possible for a small number of cases to spark an outbreak in a particular
666 stochastic realisation if interventions are relaxed too soon. Conversely, when case numbers are
667 small, an outbreak can still die out by chance even when $R_{eff} > 1$. It is therefore important to
668 account for this stochasticity when weighing the effectiveness and risks of different
669 intervention strategies, for example by considering the probability of elimination. On 18 June,
670 five weeks after AL3 restrictions were relaxed, the probability that community transmission of

671 COVID-19 had been eliminated in model simulations was estimated to be 66% in Scenario 0.
672 This estimate is calculated by finding the proportion of stochastic realisations that resulted in
673 elimination. In reality, as the outbreak died out and more days with zero new cases were
674 observed, this provided additional information about which trajectory New Zealand was most
675 likely experiencing. Each additional consecutive day with no new reported cases reduced the
676 likelihood of being on an upward trajectory. Making use of this additional information meant
677 that New Zealand's estimated probability of elimination on 18 June, given there had been 27
678 consecutive days with zero new cases reported, was approximately 95%, higher than in
679 Scenario 0 (Binny et al., 2020c). However, this estimate required up-to-date information about
680 recent case numbers. The results reported in this paper compare average outcomes under
681 different scenarios, which is appropriate for evaluating the effect of alternative actions and
682 guiding future decision making. In general, for the other scenarios we explored, bringing in
683 earlier interventions had very little impact on probability of elimination, while delaying border
684 closure or AL4 reduced the chance of elimination. Definitions of the end of an outbreak and
685 approaches to inferring probability of elimination from incidence data vary between studies.
686 Estimates can be highly sensitive to R_{eff} (Djaafara et al., 2020) and to surveillance factors,
687 including the degree of time-varying under-reporting, reporting delays, and the interaction of
688 local and imported cases (Parag et al., 2020). Our estimates of $P(elim)$ are sensitive to changes
689 in model parameters that result in changes to R_{eff} , however this does not affect our overall
690 conclusions that timing of AL4 is more important for controlling the outbreak than timing of
691 border measures, and that border measures alone are insufficient.

692

693 Our results are important for reflecting on the effectiveness of intervention timing in New
694 Zealand's COVID-19 response, relative to alternative scenarios, to help guide future response
695 strategies. Early intervention was critical to the successful control of New Zealand's March-

696 April outbreak. For modelling future disease outbreaks, epidemiological parameters should be
697 updated to reflect changes in national pandemic preparedness (e.g. improved policy and
698 response plans) and behavioural changes influencing the dynamics of future outbreaks. For
699 instance, the degree of compliance with alert level restrictions in future may differ dramatically
700 from the March-April outbreak, resulting in different values of R_{eff} . Further work is needed to
701 explore the social dynamics affecting transmission and the effectiveness of interventions, for
702 instance whether wearing masks in public spaces becomes more common, or whether more
703 people will choose to work from home or avoid travel if a suspected new outbreak is reported
704 or if government action is perceived to be inadequate.

705

706 The key measures of outbreak dynamics assessed here should be considered alongside other
707 measures of economic, social and health impacts (e.g. job losses, consumer spending, impacts
708 for mental health, rates of domestic violence or disrupted education). Particular attention needs
709 to be given to identifying vulnerable groups who may experience inequitable impacts so that
710 future policies can be tailored to support these groups. At the end of AL3, health benefits (e.g.
711 number of cases and deaths avoided) differed between scenarios. For cost-benefit analyses,
712 age-dependent morbidity and mortality rates of COVID-19 (Kang & Jung, 2020) allow
713 numbers of cases and deaths to be quantified in terms of disability-adjusted life years (DALYs)
714 avoided (or quality-adjusted life years gained), which can (with obvious issues) be converted
715 to monetary units to facilitate comparison with economic costs. Because the duration spent
716 under AL4/AL3 was fixed at 7 weeks for Scenarios 0-5, the short-term economic costs of the
717 different scenarios would have been similar, so we did not convert health benefits into DALYs
718 here. After the end of AL3, benefits and costs would differ between scenarios depending on
719 the value of R_{eff} for AL1-2 and on whether or not elimination was achieved. Increased levels
720 of activity and contact rates under AL1-2 mean that R_{eff} is very likely to have been greater than

721 one. In scenarios with lower probabilities of elimination, it is more likely that New Zealand
722 would continue to experience new cases while under AL1-2 which, with $R_{eff} > 1$, would likely
723 lead to another outbreak and require a second lockdown (with its associated costs). Conversely,
724 scenarios with higher probabilities of elimination mean there is a greater chance of the outbreak
725 dying out entirely and less risk of a second lockdown being required. Future work could
726 consider the costs and benefits of alternative scenarios where the duration of time spent in AL4
727 and AL3 is dictated by the need to achieve a certain threshold probability of elimination.

728

729 **Acknowledgements**

730 The authors acknowledge the support of StatsNZ, ESR, and the Ministry of Health in supplying
731 data in support of this work. We are grateful to Samik Datta, Nigel French, Markus Luczak-
732 Roesch, Melissa McLeod, Anja Mizdrak, Fraser Morgan and Matt Parry for comments on an
733 earlier version of this manuscript. This work was funded by the Ministry of Business,
734 Innovation and Employment and Te Pūnaha Matatini, a New Zealand Centre of Research
735 Excellence.

736

737

738 **References**

739

740 Baker MG, Kvalsvig A, Verrall AJ (2020a). New Zealand's COVID-19 elimination
741 strategy. *The Medical Journal of Australia*, 213(5):198-200.e1,
742 doi:10.5694/mja2.50735

743

744 Baker MG, Kvalsvig A, Verrall AJ, Telfar-Barnard L, Wilson N (2020b). New Zealand's
745 elimination strategy for the COVID-19 pandemic and what is required to make it work.
746 *The New Zealand Medical Journal (Online)*, 133(1512):10-4.

- 747
- 748 Baker MG, Wilson N, Anglemyer A (2020c). Successful elimination of covid-19 transmission
749 in New Zealand. *New England Journal of Medicine*, 383:e56, doi:
750 [10.1056/NEJMc2025203](https://doi.org/10.1056/NEJMc2025203)
- 751
- 752 Binny RN, Hendy SC, James A, Lustig A, Plank MJ, Steyn N (2020a). Effect of Alert Level 4
753 on effective reproduction number: review of international COVID-19 cases. medRxiv
754 preprint, doi: [10.1101/2020.04.30.20086934](https://doi.org/10.1101/2020.04.30.20086934)
- 755
- 756 Binny RN, Lustig A, Brower A, Hendy SC, James A, Parry M, Plank MJ, Steyn N (2020b).
757 Effective reproduction number for COVID-19 in Aotearoa New Zealand. medRxiv
758 preprint, doi: [10.1101/2020.08.10.20172320](https://doi.org/10.1101/2020.08.10.20172320)
- 759
- 760 Binny RN, Hendy SC, James A, Lustig A, Plank MJ, Steyn N (2020c). Probability of
761 elimination for COVID-19 in Aotearoa New Zealand. medRxiv preprint, doi:
762 [10.1101/2020.08.10.20172361](https://doi.org/10.1101/2020.08.10.20172361)
- 763
- 764 Bodas M & Peleg K (2020). Self-Isolation compliance in the COVID-19 era influenced by
765 compensation: findings from a recent survey in Israel. *Health Affairs*, 39(6):936-941,
766 doi:
767 [10.1377/hlthaff.2020.00382](https://doi.org/10.1377/hlthaff.2020.00382)
- 768
- 769 Brauner JM, Mindermann S, Sharma M, *et al.* (2020). The effectiveness of eight
770 nonpharmaceutical interventions against COVID-19 in 41 countries. medRxiv preprint,
771 doi: [10.1101/2020.05.28.20116129](https://doi.org/10.1101/2020.05.28.20116129)
- 772
- 773 [DPMC] Department of Prime Minister and Cabinet (2020). CAB-20-SUB-0270 Review of
774 COVID-19 Alert Level 2, 8 June 2020. Proactive release. url:
775 [https://covid19.govt.nz/assets/resources/proactive-release-2020-july/AL2-Minute-](https://covid19.govt.nz/assets/resources/proactive-release-2020-july/AL2-Minute-and-Paper-CAB-20-MIN-0270-Review-of-COVID-19-Alert-Level-2-8-June-2020.PDF)
776 [and-Paper-CAB-20-MIN-0270-Review-of-COVID-19-Alert-Level-2-8-June-](https://covid19.govt.nz/assets/resources/proactive-release-2020-july/AL2-Minute-and-Paper-CAB-20-MIN-0270-Review-of-COVID-19-Alert-Level-2-8-June-2020.PDF)
777 [2020.PDF](https://covid19.govt.nz/assets/resources/proactive-release-2020-july/AL2-Minute-and-Paper-CAB-20-MIN-0270-Review-of-COVID-19-Alert-Level-2-8-June-2020.PDF) (accessed 9 September 2020)
- 778

- 779 Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ, CMMID COVID-19 Working
780 Group, (2020). Effects of non-pharmaceutical interventions on COVID-19 cases,
781 deaths and demand for hospital services in the UK: a modelling study. *The Lancet.
782 Public Health*, 5(7):e375-e385, doi: 10.1016/s2468-2667(20)30133-x
783
- 784 Desvars-Larrive A, Ahne V, Álvarez S, *et al.* (2020). CCCSL: Complexity Science Hub Covid-
785 19 Control Strategies List. Version 2.0. Available from: [https://github.com/amel-
786 github/covid19-interventionmeasures](https://github.com/amel-github/covid19-interventionmeasures) (accessed 20 August 2020)
787
- 788 Djaafara BA, Imai N, Hamblion E, Impouma B, Donnelly CA, Cori A (2020). A Quantitative
789 Framework for Defining the End of an Infectious Disease Outbreak: Application to
790 Ebola Virus Disease. *American Journal of Epidemiology*, kwaa212, doi:
791 10.1093/aje/kwaa212.
792
- 793 Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, Parker M, Bonsall D,
794 Fraser C (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control
795 with digital contact tracing. *Science*, 368(6491):eabb6936, doi:
796 <https://doi.org/10.1126/science.abb6936>
797
- 798 Flaxman S, Mishra S, Gandy A, *et al.* (2020). Estimating the effects of non-pharmaceutical
799 interventions on COVID-19 in Europe. *Nature*, 584:257–261, doi: [10.1038/s41586-
800 020-2405-7](https://doi.org/10.1038/s41586-020-2405-7)
801
- 802 Hendy SC, Steyn N, James A, Plank MJ, Binny RN, Lustig A (2021). Mathematical modelling
803 to inform New Zealand’s COVID-19 response. *Journal of the Royal Society of New
804 Zealand*, doi: 10.1080/03036758.2021.1876111
- 805
- 806 Hsiang S, Allen D, Annan-Phan S, *et al.* (2020). The Effect of Large-Scale Anti-Contagion
807 Policies on the Coronavirus (COVID-19) Pandemic. medRxiv preprint, doi:
808 [10.1101/2020.03.22.20040642](https://doi.org/10.1101/2020.03.22.20040642)
809

810 Ioannidis JPA (2020). Infection fatality rate of COVID-19 inferred from seroprevalence data.
811 *Bulletin of the World Health Organization*. Early access online, url:
812 https://www.who.int/bulletin/online_first/BLT.20.265892.pdf (accessed 4 November
813 2020)

814
815 James A, Plank MJ, Binny RN, Hannah K, Hendy SC, Lustig A, Steyn N (2020a). A structured
816 model for COVID-19 spread: modelling age and healthcare inequities.
817 medRxiv preprint, doi: [10.1101/2020.05.17.20104976](https://doi.org/10.1101/2020.05.17.20104976)

818
819 James A, Plank MJ, Binny RN, Lustig A, Steyn N, Hendy SC, Nesdale A, Verrall A (2020b).
820 Successful contact tracing systems for COVID-19 rely on effective quarantine and
821 isolation. medRxiv preprint, doi: [10.1101/2020.06.10.20125013](https://doi.org/10.1101/2020.06.10.20125013)

822
823 James A, Plank MJ, Hendy SC, Binny RN, Lustig A, Steyn N (2020c). Model-free estimation
824 of COVID-19 transmission dynamics from a complete outbreak. medRxiv preprint, doi:
825 [10.1101/2020.07.21.20159335](https://doi.org/10.1101/2020.07.21.20159335)

826
827 Jefferies S, French N, Gilkison C, Graham G, Hope V, Marshall J, McElnay C, McNeill A,
828 Muellner P, Paine S, Prasad N, Scott J, Sherwood J, Yang L, Priest P (2020). COVID-
829 19 in New Zealand and the impact of the national response: a descriptive
830 epidemiological study. *The Lancet Public Health*, 5(11):e612 - e623.

831
832 Kang SJ & Jung SI (2020). Age-Related Morbidity and Mortality among Patients with COVID-
833 19. *Infection & chemotherapy*, 52(2):154–164, doi: [10.3947/ic.2020.52.2.154](https://doi.org/10.3947/ic.2020.52.2.154)

834
835 Kontis V, Bennett JE, Rashid T, Parks RM, Pearson-Stuttard J, Guillot M, Asaria P, Zhou B,
836 Battaglini M, Corsetti G, McKee M (2020). Magnitude, demographics and dynamics
837 of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21
838 industrialized countries. *Nature Medicine*, 14:1-0.

839
840 Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005). Superspreading and the effect of
841 individual variation on disease emergence. *Nature*, 438(7066):355-359.

842

843 Obadia T, Haneef R, Boëlle PY (2012). The R0 package: a toolbox to estimate reproduction
844 numbers for epidemic outbreaks. *BMC medical informatics and decision making*, 12(1),
845 147.

846
847 Parag KV, Cowling BJ, Donnelly CA (2021). Deciphering early-warning signals of the
848 elimination and resurgence potential of SARS-CoV-2 from limited data at multiple
849 scales. medRxiv preprint, doi: [10.1101/2020.11.23.20236968](https://doi.org/10.1101/2020.11.23.20236968)

850
851 Parag KV, Donnelly CA, Jha R, Thompson RN (2020). An exact method for quantifying the
852 reliability of end-of-epidemic declarations in real time. *PLoS Computational Biology*
853 16(11):e1008478, doi: [10.1371/journal.pcbi.1008478](https://doi.org/10.1371/journal.pcbi.1008478)

854
855
856 Plank MJ, Binny RN, Hendy SC, Lustig A, James A, Steyn N (2020). Effective reproduction
857 number and likelihood of cases outside Auckland. url: [https://cpb-ap-](https://cpb-ap-se2.wpmucdn.com/blogs.auckland.ac.nz/dist/d/75/files/2020/09/reff-update-and-travel-model-results-for-release-FINAL.pdf)
858 [se2.wpmucdn.com/blogs.auckland.ac.nz/dist/d/75/files/2020/09/reff-update-and-](https://cpb-ap-se2.wpmucdn.com/blogs.auckland.ac.nz/dist/d/75/files/2020/09/reff-update-and-travel-model-results-for-release-FINAL.pdf)
859 [travel-model-results-for-release-FINAL.pdf](https://cpb-ap-se2.wpmucdn.com/blogs.auckland.ac.nz/dist/d/75/files/2020/09/reff-update-and-travel-model-results-for-release-FINAL.pdf) (accessed 9 October 2020)

860
861 Statistics New Zealand (2018). 2018 Census of population and dwellings counts. url:
862 [https://www.stats.govt.nz/information-releases/2018-census-population-and-dwelling-](https://www.stats.govt.nz/information-releases/2018-census-population-and-dwelling-counts)
863 [counts](https://www.stats.govt.nz/information-releases/2018-census-population-and-dwelling-counts) (accessed 4 November 2020)

864
865 Steyn N, Binny RN, Hannah K, Hendy SC, James A, Kukutai T, Lustig A, McLeod M, Plank
866 MJ, Ridings K, Sporle A (2020). Estimated inequities in COVID-19 infection fatality
867 rates by ethnicity for Aotearoa New Zealand. medRxiv preprint, doi:
868 [10.1101/2020.04.20.20073437](https://doi.org/10.1101/2020.04.20.20073437)

869
870 Summers JL, Lin H-H, Cheng H-Y, Telfar Barnard L, Kvalsvig A, Wilson N, Baker MG
871 (2020). Potential lessons from the Taiwan and New Zealand health responses to the
872 COVID-19 pandemic. *The Lancet Regional Health - Western Pacific*. (Accepted, in
873 press).

874
875 Verity R et al (2020). Estimates of the severity of coronavirus disease 2019: a model-based

876 analysis. *The Lancet*, 20(6), 669-677, doi: 10.1016/S1473-3099(20)30243-7
877
878 Verrall A (2020). Rapid Audit of Contact Tracing for Covid-19 in New Zealand. Ministry of
879 Health, New Zealand.
880
881 Verrall A, Norton K, Rooker S, *et al.* (2010). Hospitalizations for pandemic (H1N1) 2009
882 among Māori and Pacific Islanders, New Zealand. *Emerg Infect Dis.*, 16:100-2.
883
884 Wallinga J and Lipsitch M (2007). How generation intervals shape the relationship between
885 growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological*
886 *Sciences*, 274:599-604, doi: 10.1098/rspb.2006.3754
887
888 Wilson N, Parry M, Verrall AJ, Baker MG, Schwehm M, Eichner M (2020). When Can
889 Elimination of SARS-CoV-2 Infection be Assumed? Simulation Modelling in a Case
890 Study Island Nation. medRxiv preprint, doi: [10.1101/2020.05.16.20104240](https://doi.org/10.1101/2020.05.16.20104240)
891
892 Wilson N, Telfar-Barnard L, Summers J, *et al.* (2012). Differential mortality by ethnicity in 3
893 influenza pandemics over a century, New Zealand. *Emerg. Infect. Dis.*, 18:71-77.
894
895
896
897
898
899
900
901