Te Pūnaha Matatini
Data ▪ Knowledge ▪ Insight

Note: This paper has not yet undergone formal peer review

Potential reduction in transmission of COVID-19 by digital contact tracing systems

26 August 2020

Michael J. Plank[1,4], Alex James[1,4], Audrey Lustig[2,4], Nicholas Steyn[3,4], Rachelle N Binny[2,4], Shaun C. Hendy[3,4]

1. School of Mathematics and Statistics University of Canterbury, New Zealand.
2. Manaaki Whenua, Lincoln, New Zealand.
3. Department of Physics, University of Auckland, New Zealand.
4. Te Pūnaha Matatini: Centre of Research Excellence in Complex Systems, New Zealand.

Executive Summary

- To maintain elimination of COVID-19, digital contact tracing systems should be designed to complement manual contact tracing, for example by enhancing coverage or speed of tracing, rather than as a separate or fully automated system.
- To reduce the effective reproduction number to around 1 requires a combination of rapid testing and case isolation, a well-functioning manual contact tracing system, digital contact tracing with an uptake rate of at least 75% and recording 90% of close contacts, and highly effective quarantine of traced contacts.
- Ensuring that individuals with COVID-19 symptoms get tested quickly and are able to isolate effectively is just as important as investment in contact tracing.
- Digital systems based on QR codes with no proximity detection are likely to be less effective as a result of recording fewer contacts.
- Bluetooth apps and card-based proximity detection systems perform comparably at a given level of coverage, but other factors such as usability, reliability and longevity need to be considered.
- In the event of a large ongoing outbreak, scalability and false positive rates are more important, but significant population-wide control measures are also likely be required to prevent a major epidemic.
- Tracing and quarantining second-order contacts of a confirmed case provides a relatively small additional benefit. This could be useful in the very early stages of an outbreak, but for a larger outbreak ensuring fast and effective quarantine of first-order contacts should be a higher priority.

Abstract

Digital tools are being developed to support contact tracing as part of the global effort to control the spread of COVID-19. These include smartphone apps, Bluetooth-based proximity detection, location tracking, and automatic exposure notification features. Evidence on the effectiveness of alternative approaches to digital contact tracing is so far limited. We use an age-structured branching process model of the transmission of COVID-19 in different settings to estimate the potential of manual contact tracing and digital tracing systems to help control the epidemic. We investigate the effect of the uptake rate and proportion of contacts recorded by the digital system on key model outputs: the effective reproduction number, the mean outbreak size after 30 days, and the probability of elimination. We show that effective manual contact tracing can reduce the effective reproduction number from 2.4 to around 1.5. The addition of a digital tracing system with a high uptake rate over 75% could further reduce the effective reproduction number to around 1.1. Fully automated digital tracing without manual contact tracing is predicted to be much less effective. We conclude that, for digital tracing systems to make a significant contribution to the control of COVID-19, they need be designed in close conjunction with public health agencies to support and complement manual contact tracing by trained professionals.

Introduction

Contact tracing has become a key tool in the global effort to control the spread of COVID-19. Contact tracing has been crucial in controlling several disease outbreaks, notably SARS, MERS and Ebola (WHO & CDC, 2015; Kang et al., 2015). While contact tracing alone is unlikely to contain the spread of COVID-19 (Hellewell et al., 2020; Kucharski et al., 2020), in countries like New Zealand where cases have been reduced to very low numbers (Cousins, 2020; Binny et al., 2020), it may allow population-wide social distancing measures to be relaxed. In countries with more widespread epidemics, it can allow safe reopening.

Manual contact tracing typically involves interviewing confirmed cases about their recent contacts, getting in touch with those contacts and asking them to take measures to prevent onward transmission of the disease in case they are infected. Such measures may include limiting their interactions with others, formal quarantine, getting tested, or remaining vigilant for symptoms. Manual contact tracing is intensive work and requires highly trained public health professionals to be implemented effectively (Verrall, 2020). It is also difficult to scale manual contact tracing up to deal with very large outbreaks.

In response, many countries have attempted to develop digital contact tracing systems using smartphone apps. There are multiple different approaches to this problem, for example QR code-based systems, Bluetooth-proximity based apps, automated exposure notification features, and systems that do not rely on smartphones such as card-based proximity detection. These systems can offer, to varying extents, three main benefits to controlling COVID-19: (i) an increase in the proportion of contacts who are traced (e.g. contacts that would not be traced by case recall but are recorded by the digital system); (ii) a reduction in the time taken to identify and notify traced contacts (e.g. via an exposure notification feature); (iii) improved scalability over manual contact tracing.

The effectiveness of digital contact tracing is still unproven, with limited real-world data (Anglemyer et al., 2020). Here, we use a model of COVID-19 transmission and contact tracing to evaluate the potential of digital contact tracing systems to reduce the spread of COVID-19. We evaluate the benefits of digital contact tracing both alone and in combination with manual tracing, over a range of uptake rates, tracing probabilities, and the effectiveness of quarantine.

Methods

Transmission and contact tracing model.

We use an age-structured branching process model for COVID-19 transmission and contact tracing that is an extension of the age-structured model of James et al. (2020) to include different contact types (see below). We assume that the outbreak is sufficiently small that the effect of infection-induced population immunity can be ignored. The time from infection to symptom onset is gamma distributed with mean 5.5 days and standard deviation 2.3 days (Lauer et al., 2020). Infectiousness is a Weibull function, time-shifted such that 35% of transmission occurs prior to symptom onset (Ferretti et al., 2020; Ganyani et al., 2020). Infections have an age-dependent probability of being subclinical (Davies et al., 2020) that decreases linearly from 40% in the 0-10 year age group to 5% in the over 70 years age group. Subclinical infections are assumed to be 50% as infectious as clinical cases.

Contacts are categorised into one of four different types: home, work, school and casual. Each contact has a probability, called the secondary attack rate (SAR), of resulting in transmission. The SAR is assumed to be 20% for home contacts and 6% for work, school and casual contacts (Kucharski et al., 2020). Age-specific contact rates in each of these four settings are based on the contact rates estimated by Prem at al. (2017) for the New Zealand population. The average number of age-specific home contacts was taken directly from the results of Prem et al. (2017) and assumes that household contacts do not vary from day to day. The average number of work, school and casual contacts made during the infectious period was chosen to be a fixed multiple of the number of daily work, school and casual contacts in Prem et al. (2017), chosen to give a basic reproduction number (in the absence of any control measures, case isolation, or contact tracing) of $R_0 = 2.6$ (Jarvis et al., 2020, Kucharski et al., 2020). We model heterogeneity in number of contacts via gamma distributed individual multipliers for each of the four settings (Lloyd-Smith et al., 2005). Heterogeneity is assumed to be higher in work, school and casual contacts than in home contacts, reflecting a greater occurrence of superspreading events in these settings. We model homophily by assuming that the number of contacts of a secondary case is correlated to the number of contacts of the index case in the setting in which transmission occurred. We assume that each transmission via a non-home contact results in infection of a new household, which is assigned a household size according to the age-specific distribution of home contacts. Each household is assumed to consist of a fixed group of individuals, so that subsequent infections within the same household deplete the pool of susceptible home contacts (see Appendix for details).

In the absence of any contact tracing, clinical cases are assumed to be eventually tested with probability $p_{detect} = 1$ (i.e. all clinical cases are eventually detected). The delay from onset of symptoms to testing is assumed to be gamma distributed with mean 6.8 days and standard deviation 5.9 days (estimated from New Zealand data for the March-April outbreak). Cases are isolated at the same time as getting tested and this prevents any further transmission. There is a further delay between getting tested and the test result being returned that is a minimum of 0.5 days, plus an exponentially distributed random variable with a mean of 0.5 days. Subclinical cases do not get tested are do not isolate.

The contact tracing model is illustrated in Figure 1. When a new positive test result is returned and contact tracing for that case begins, we refer to the individual testing positive as the index case and to contacts of the index case as secondary cases. Contacts who do not end up getting infected are not modelled explicitly (but see Discussion about effects of quarantining false positives). Tracing of the index case's contacts begins when the index case returns a positive test result. Under manual tracing, each contact has a probability of being traced, and a time taken to trace, both of which may differ across the four settings. Traced contacts who are not currently symptomatic (i.e. either subclinical or pre-symptomatic) go into quarantine, which is assumed to reduce onward transmission to a level $c_{quar} < 1$ relative to no quarantine. Traced contacts who are symptomatic go into isolation immediately on symptom onset, which is assumed to completely prevent any further onward transmission. Untraced contacts do not go into quarantine, and experience a delay between symptom onset and isolation and testing (see above). Effective contact tracing therefore reduces transmission in two ways: (i) quarantining of contacts who are not currently symptomatic; and (ii) prompt isolation of contacts immediately on symptom onset.

Home contacts are always assumed to be traced instantly (i.e. immediately after the index case returns a positive test result) with probability 1, and this happens independently of any digital contact tracing system. Under manual contact tracing, work contacts are traced with probability 0.5, school contacts with probability 0.8, and casual contacts with probability 0.25. Work and casual contacts are assumed to have a tracing time that is gamma distributed with mean 3 days and standard deviation 1.7 days. School contacts are assumed to be traced more rapidly but not instantly (0.5 days after the index case returns a positive test result).

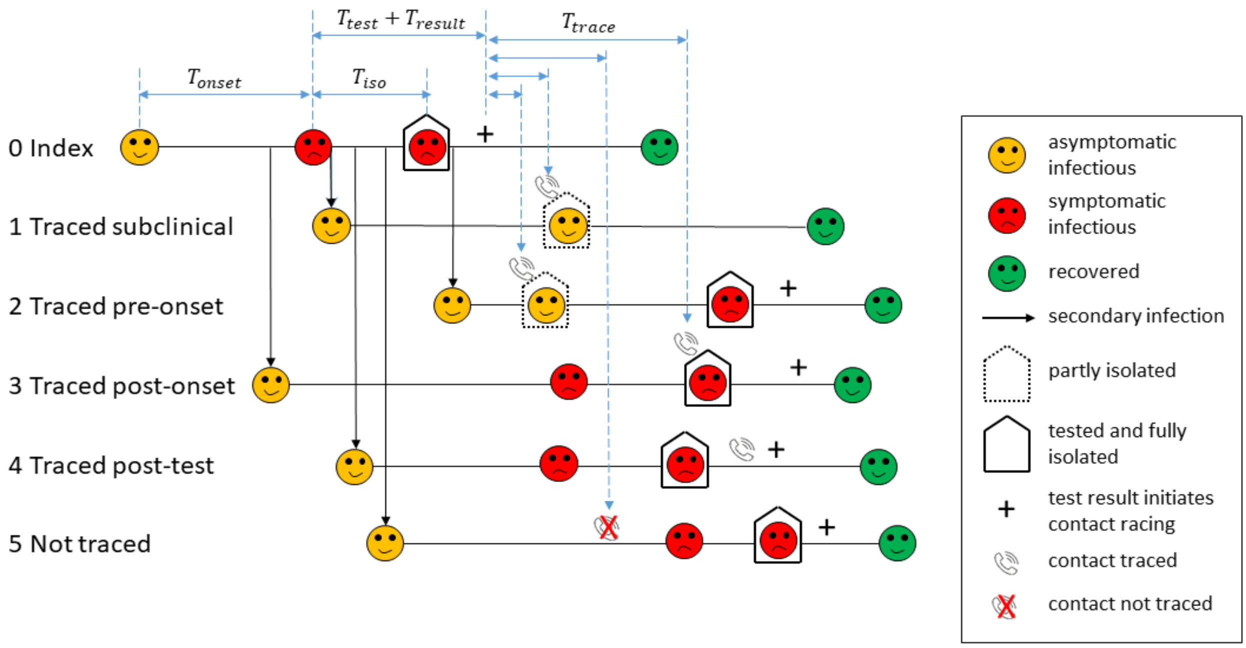


Figure 1. Schematic diagram of the contact tracing model. Infectious individuals are initially asymptomatic (yellow). For the index case who was not traced (0), there is a delay between onset of symptoms (red) and getting tested. Isolation occurs at the same time as getting tested. There is a subsequent delay to the test result being returned (+) and tracing of contacts. Traced contacts (1-4) are quarantined when contacted by public health officials (phone icons) and are isolated and tested immediately on symptom onset. Traced contacts (3) who are already symptomatic prior to being traced are isolated immediately when contacted. Traced contacts (4) that have already isolated prior to being traced are not affected. Contacts that cannot be traced (5) may still get tested and isolated, but this is likely to take longer. Subclinical individuals (1) do not get tested or isolated, but will be quarantined if they are a traced contact.

Digital contact tracing systems

We model alternative digital contact tracing systems by varying key parameters of the contact tracing model. Home contacts are assumed to be always traced rapidly by the manual system, so digital contact tracing applies to school, work and casual contacts. For each scenario, we assume there is an uptake rate u and that each individual in the population is a user of the digital tracing system (e.g. has installed the app) with probability u, independently of all other individuals. This ignores any correlation in the usage probabilities of close contacts.

Provided the index case and secondary case are both users of the system, the contact is digitally logged with probability P_d, which we assume is the same for school, work and casual contacts (see Table 1). We set a default value of $P_d = 90\%$, but also investigate values of P_d smaller than this. A tracing probability of 100% is likely to be unachievable, for example because there will be situations where one or both individuals failed to carry the smartphone or card with them. Contacts that are digitally logged are assumed to be quarantined immediately after the index case returns a positive test result, the same as for home contacts. If neither or one of the index case and the secondary case is a user of the system, or both are users but the contact was not logged by the digital system, the contact is not traced digitally, but may still be later traced manually.

In addition to the benefits of instant tracing described above, some digital tracing systems may also help improve coverage of manual tracing. For example, location-based tracking or a digital diary feature may enable contact tracers to identify and follow up contacts who would otherwise be missed. We do not explicitly investigate these scenarios, though they could be modelled via an increase in the manual tracing probabilities of relevant settings in Table 1.

Rather than making estimates of tracing probabilities for alternative digital systems, we investigate how the performance of the contact tracing system varies with the uptake rate u and the probability of digital tracing P_d (see Table 1). We assume that isolation of symptomatic traced contacts is 100% effective in preventing onward transmission, but quarantine of pre-symptomatic and subclinical traced contacts is only partially effective. We investigate two scenarios: (1) quarantine reduces transmission by 50% ($c_{quar} = 0.5$); (2) quarantine reduces transmission by 80% ($c_{quar} = 0.2$). We assume the effectiveness of quarantine/isolation is the same for contacts that are traced digitally and contacts that are traced manually. This is likely to require effective manual follow up of digitally traced contacts as opposed to relying solely on automatic exposure notifications.

We also investigate the additional benefit from including recursive tracing of second-order contacts (i.e. quarantining the contacts of contacts of a confirmed case) in the model. Recursive tracing and effective quarantine of second-order contacts is more difficult to achieve in practice because of the much larger number of second-order contacts and the lower risk of them being infected. In the case on an ongoing outbreak with a large number of cases, the number of uninfected individuals being quarantined under recursive tracing is likely to be prohibitively large (Firth et al., 2020). However, recursive tracing could potentially be useful in suppressing a small outbreak in its very early stages. We assume that tracing of second-order contacts begins 2 days after tracing of the first-order contact and can occur digitally or manually following the same rules described above (Figure 2). This means that second-order contacts who are traced digitally are quarantined 2 days after quarantining the first-order contact; second-order contacts traced manually are quarantined later. Any second-order contacts made subsequent to quarantine of the first-order contact cannot be traced recursively (but may still be traced on or after the positive test result of the first-order contact). Third-order contacts were not traced recursively.

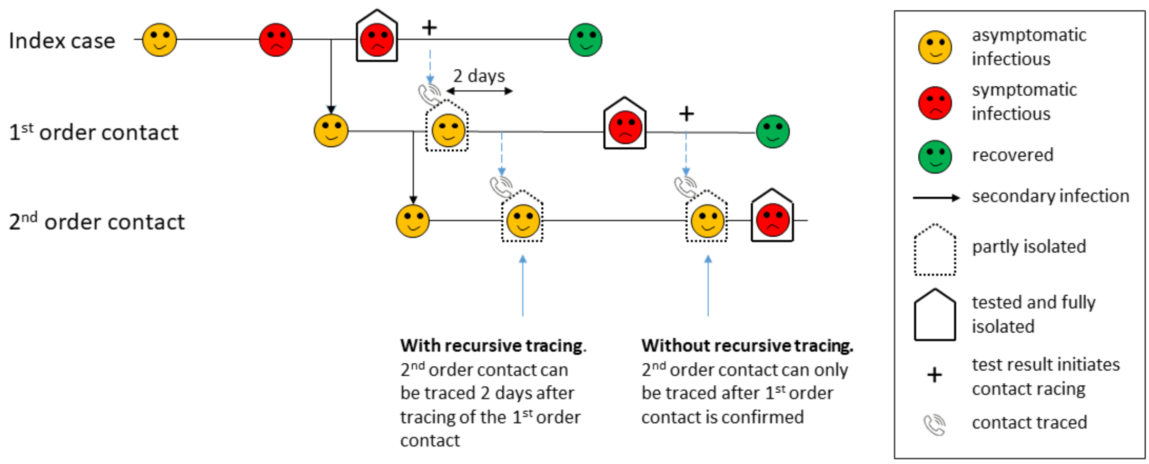


Figure 2. Recursive tracing of second-order contacts. With recursive tracing, second-order contacts of a confirmed case can be traced and quarantined two days after the first-order case is traced and quarantined (assuming both contacts are traced digitally).

	Probability traced manually	Manual tracing time (days)	Probability traced digitally (provided index case and contact are both users)	
Home	100%	0	-	
School	80%	0.5	P_d	default 90%
Work	50%	$\Gamma(\text{mean} = 3, \text{s.d.} = 1.7)$	P_d	default 90%
Casual	25%	$\Gamma(\text{mean} = 3, \text{s.d.} = 1.7)$	P_d	default 90%

Table 1: Contact tracing probabilities and times. Manual contact tracing has different tracing probabilities and times for home, school, work and casual contacts. Home contacts are always traced and this is assumed to happen with no delay. School contacts are traced with probability 80% and this takes half a day. Work and casual contacts have lower tracing probabilities and a delay of 3 days on average. We investigate manual contact tracing supported by a digital contact tracing system that has probability P_d (default value 90%) of instantly tracing school, work and casual contacts, provided both the index case and the contact are users of the digital system. Contacts that are not traced by the digital system may still later be traced manually.

Results

We measured the reduction in spread of COVID-19 by looking at three model outputs: (a) the effective reproduction number R_{eff} (average number of secondary infections per case); (b) the mean outbreak size (total number of cases per seed case) after 30 days; (c) the probability of extinction of an outbreak starting from a single seed case. Together these outputs measure the relative effectiveness of the contact tracing system in containing the virus. These results are robust to the initial number of seed cases: if there are multiple initial seed cases, the reproduction number is not affected; the mean outbreak size is simply multiplied by the initial number of cases, and the probability of elimination is raised to the power of the initial number of cases.

For each combination of contact tracing parameters, we ran multiple simulations each of which was initialised with a single infected seed case. Results are shown for uptake rates ranging from 0 to 90% and were calculated by averaging over a sufficient number of simulations to provide an aggregate total of at least 100,000 cases.

With case isolation in the absence of any contact tracing, the effective reproduction number was$R_{eff} = 2.4$, the mean outbreak size after 30 days was 78, and the probability of extinction was approximately 47%. Manual-only contact tracing (which corresponds to a digital uptake rate of $u = 0$ in Fig. 3) with moderately (50%) effective quarantine of pre-symptomatic or subclinical individuals reduced R_{eff} to 1.55, the mean outbreak size to approximately 34 and increased the probability of extinction to 67%.

When quarantine is moderately effective (reduces transmission by 50%, blue curves in Fig. 3), the addition of digital tracing with high uptake rate (>75%) and high probability of logging contacts ($P_d = 90\%$) reduced R_{eff} to around 1.22, mean outbreak size to 20, and increased the probability of extinction to 80%. If quarantine is more effective (reduces transmission by 80%, red curves in Fig. 3), digital tracing can reduce R_{eff} to approximately 1.12 and increase probability of elimination to 90%. Adding recursive tracing of second-order contacts (orange curves in Fig. 3) provides a relatively small reduction in R_{eff} to 1.05, although this does increase the probability of elimination from 90% to 95%.

Lower uptake rates (<75%) result in poorer performance although there is still some noticeable benefit of digital tracing at an uptake rate of around 40%, provided the probability of a close contact being logged is high ($P_d = 90\%$ in Fig. 2). If P_d is much lower than this, performance will deteriorate. However, the results are not as sensitive to P_d as they are to uptake rate u, because the requirement for both the index and the secondary case to be users of the system means there is a quadratic dependence on uptake rate. This means that a 20% reduction in P_d is approximately equivalent to a 10% reduction in uptake rate.

We also considered a scenario in which there is no manual contact tracing, except for home contacts which are still assumed to be traced instantly (Figure 4). This could represent a situation where a larger outbreak has exceeded the capacity of the manual contact tracing system, so non-household contacts can only be traced digitally. In this scenario, we assume that quarantine of pre-symptomatic individuals is only moderately (50%) effective and there is no recursive tracing, representing a digital-only contact tracing system without consistent follow-up from trained public health professionals. We measure the effectiveness of contact tracing by R_{eff}, as the other two measures (mean outbreaks size per seed case and probability of elimination) are less relevant in the case of a large outbreak. In this scenario, digital contact tracing makes a larger relative contribution to controlling the spread of COVID-19. However, it also means that digital tracing alone is unlikely to be able to contain an outbreak: even with a very high uptake rates (80%) of an effective digital tracing system ($P_d = 90\%$), R_{eff} is around 1.46. This implies that digital contact tracing would need to be combined with significant population-wide control measures in order to avoid a major epidemic.

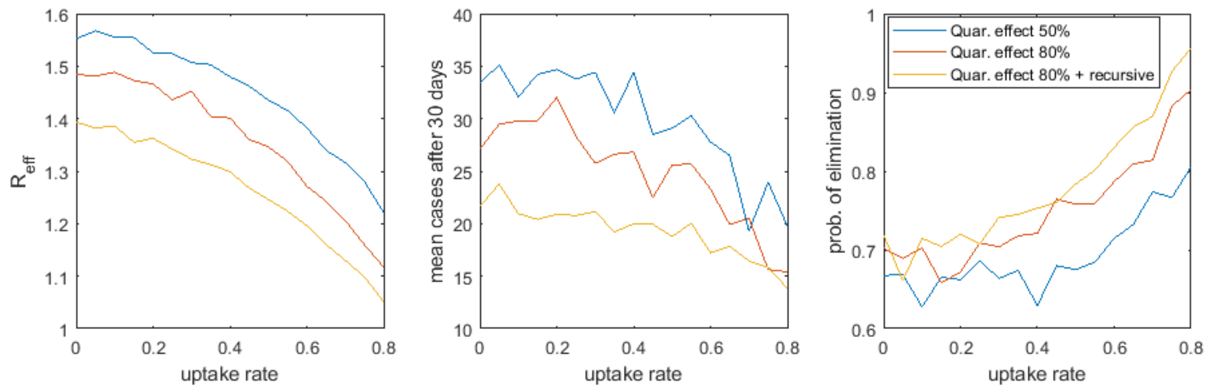


Figure 3: Effectiveness of manual contact tracing plus digital tracing at a range of uptake rates, under different effectiveness of quarantine, with and without recursive tracing. Effectiveness of manual contact tracing supported by a digital tracing system measured by: (a) effective reproduction number R_{eff}; (b) mean number of cases per seed case after 30 days; (c) probability of elimination of an outbreak starting from a single seed case. Proportion of contacts logged by digital tracing system when both individuals are users of the system is $P_d = 90\%$. Quarantine reduces onward transmission by 50% (blue), by 80% (red), by 80% and with recursive tracing (orange). Isolation of symptomatic cases completely prevents onwards transmission. Results are averaged over sufficient simulations of the branching process to give a total 100,000 cases.

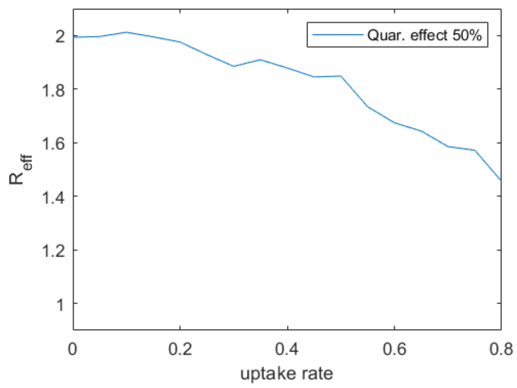


Figure 4: Effectiveness of digital only contact tracing at a range of uptake rates. Effectiveness of a digital tracing system measured by effective reproduction number R_{eff}. Home contacts are still traced manually, but other contacts can only be traced digitally. Proportion of contacts logged by digital tracing system when both individuals are users of the system is $P_d = 90\%$. Quarantine reduces onward transmission by 50%. Isolation of symptomatic cases completely prevents onwards transmission. Results are averaged over sufficient simulations of the branching process to give a total 100,000 cases.

Table 2 shows a comparison of the effectiveness of alternative technological approaches to digital contact tracing, modelled as having different probabilities of recording contacts. Table 2 shows the effective reproduction number for manual contact tracing plus digital tracing systems at different levels of uptake, with highly effective (80%) quarantine and without recursive tracing of second-order contacts. Systems based on QR codes alone without proximity detection are likely to perform less well because of the additional steps required for a contact to be recorded: the location needs to have a QR code displayed and both the index case and secondary case need to scan it. This does not diminish the benefit to manual tracing of users scanning QR codes to maintain a record of their movements, but we do not explicitly model this here. Fully decentralised Bluetooth apps are estimated to be less effective than centralised apps at the same level of uptake, because of a reduced likelihood of users reacting to automatic exposure notifications from a decentralised system without follow up from manual contact tracers. A card-based proximity system is estimated to perform similarly to a centralised Bluetooth app, though with a slightly reduced effectiveness because notifications cannot be sent natively and need to be made via a separate system which requires current contact details.

Digital tracing system type	**Assumed proportion of contacts lagged**	**Effective reproduction number R_{eff}**		
		40% uptake	60% uptake	80% uptake
QR code exposure notification	$P_d = 40\%$ of contacts logged	1.47	1.44	1.41
Decentralised app-based Bluetooth proximity detection	$P_d = 90\%$ of contacts logged, but only 50% of notifications result in quarantine	1.47	1.42	1.40
Centralised app-based Bluetooth proximity detection	$P_d = 90\%$ of contacts logged	1.40	1.27	1.12
Card-based proximity detection	$P_d = 90\%$ of contacts logged, but only 85% result in an instant notification	1.41	1.34	1.22

Table 2. Comparison of alternative technological approaches to digital contact tracing. Effective reproduction number R_{eff} of manual contact tracing plus different types of digital tracing system at 40% uptake, 60% uptake and 80% uptake, with highly effective (80%) quarantine and without recursive tracing. A system based on QR codes with no proximity-detection is modelled as logging a low proportion (40%) of contacts, because the location needs to have a QR code displayed, as well as both contacts taking the additional step of scanning the code. Bluetooth apps are modelled as logging a high proportion (90%) of contacts, but in decentralised systems only 50% of users are assumed to quarantine following a notification. A card-based proximity detection system is assumed to have similar detection probability as a Bluetooth app, but only 85% of contacts lead to an instant notification because the card is separate from user's phones.

Discussion

Successful control of COVID-19 is likely to require a range of intervention strategies, including some or all of: moderate population-wide social distancing; widespread use of face coverings; restrictions on large gatherings or other interventions targeting superspreading events; case isolation and household quarantine; manual and digital contact tracing (Ferreti et al. 2020; Hellewell et al., 2020; Kucharski et al., 2020). Establishing trusted relationships with cases and contacts is crucial both to increasing contact tracing coverage and to supporting effective quarantine and isolation. There is still limited evidence on the effectiveness of digital tracing systems (Anglemyer et al., 2020) and reliance on digital tracing alone is unlikely to be sufficient. However, there is potential for digital tracing to enhance the coverage and/or speed of contact tracing systems run by trained public health professionals (James et al., 2020). Digital systems should therefore be seen as an opportunity to improve coverage and/or speed by acting in a complementary role to manual contact tracing. Emphasis should be on how these systems can provide additional useful information to contact tracers in a timely way or fill potential gaps in the manual system. This implies that thorough consultation with public health agencies undertaking contact tracing should be a pre-requisite for an effective digital support system.

In this paper, we modelled the effect of manual contact tracing supported by a digital tracing system with varying levels of uptake and effectiveness. Our results show that manual contact tracing can significantly reduce the spread of COVID-19, but on its own is not sufficient to make the effective reproduction number R_{eff} less than 1. Manual contact tracing supported by a digital tracing system can further reduce spread, depending on the effectiveness of quarantine of unconfirmed (pre-symptomatic and subclinical) cases. If quarantine is moderately effective (reduces transmission by 50%), a manual plus digital tracing system with very high uptake (>75%) could reduce R_{eff} to around 1.2. If quarantine is highly effective (reduces transmission by 80%), R_{eff} can be reduced to approximately 1.1, which could be sufficient to contain the majority of small outbreaks. However, if the uptake rate is less than around 75%, the reduction in R_{eff} is not sufficient to contain an outbreak without additional measures.

Our results suggest that the additional benefit from recursive tracing of second-order contacts is relatively small. In the case of a small outbreak, this may be worthwhile as it could increase the probability of elimination from 90% to 95% provided uptake is high and quarantine is effective. However, given that it is likely more difficult to effectively quarantine second-order contacts, and the number of uninfected people being quarantined would be much higher, it may be more beneficial to focus on effective quarantine of first-order contacts than attempting to locate second-order contacts.

In our model, untraced clinical cases take 6.8 days on average to get tested after onset of symptoms. This assumption was based on the onset to reporting times in data from the March-April outbreak in New Zealand. The aim of contact tracing is to find close contacts of confirmed cases and therefore enable early quarantine or isolation. However, it is important to recognise that reducing the time between symptom onset and isolation can bring significant benefits, even in the absence of contact tracing. This implies that enhancing public awareness of COVID-19 symptoms and the need to get tested quickly, ensuring equitable access to healthcare, and maintaining rapid testing capacity are equally important as investment in contract tracing.

We have not considered the consequences of false positives from digital contact tracing systems, i.e. people who are recorded as potential contacts by the digital system but have not in fact been at risk of exposure to COVID-19. For countries such as New Zealand that have reduced cases to very low numbers and where the primary goal is to achieve or maintain elimination of community transmission, we assume that the number of false positives is not a primary consideration.

If case numbers exceed the capacity of the manual contact tracing system, its performance will rapidly deteriorate. Under this scenario, digital tracing can make a significant contribution to slowing the growth of the outbreak, but if it became the dominant form of tracing it is likely to be insufficient to reduce R_{eff} under 1. This implies that population-wide control measures would likely be needed to prevent a major epidemic. A low false positive rate is a more important consideration in this situation.

Our model allowed for age-specific mixing patterns in home, work, school, and casual contacts, and for heterogeneity and homophily in individual contact rates. However, the model ignores other sources of heterogeneity, for example in types of workplace. Some workplaces will be much more amenable to rapid contact tracing than others, for example where employees tend to work in a consistent physical location each day. This could be modelled via individual heterogeneity and homophily in contact tracing probabilities as well as contact rates, however the mean tracing probability is likely to remain the most important parameter.

We have taken a technology-agnostic approach to modelling digital contact tracing. Among the most important parameters for any digital tracing system are the uptake rate (number of people using the system) and the probability of close contacts being logged. Our results show that to contain an outbreak, a well-functioning manual contact tracing system needs to be combined with a digital tracing system that should ideally have at least 75% uptake and record and record 90% of close contacts. Different systems may have different uptake rates, for example due to usability and privacy issues, so a careful study of expected uptake rates is critical to choosing the best system. The proportion of contacts logged will be affected by the usability of the system and by individual behaviour. For example, contacts will be missed in situations where a user forgets or loses their phone or card, has the phone switched off or a flat battery, or has Bluetooth deactivated. Systems that rely solely on QR code scanning for exposure notification are likely to perform less well than systems with proximity-based detection at the same level of uptake. This is because successful tracing requires the location to have a compatible QR code displayed, both individuals to have the app installed and to take the additional step of scanning the code. There may be benefits to a QR code app in keeping a record of movements or digital diary, but we did not explicitly consider these here.

Acknowledgements

The authors acknowledge the support of StatsNZ, ESR, and the Ministry of Health in supplying data in support of this work. The authors are grateful to Andrew Chen, Matt Parry and Philippa Yasbek for discussions about digital contact tracing systems and feedback on an earlier version of this paper. This work was funded by the Ministry of Business, Innovation and Employment and Te Pūnaha Matatini, New Zealand's Centre of Research Excellence in complex systems.

References

Anglemyer A, Moore THM, Parker L, Chambers T, Grady A, Chiu K, Parry M, Wilczynska M, Flemyng E, Bero L (2020). Digital contact tracing technologies in epidemics: a rapid review. Cochrane Database of Systematic Reviews 2020, Issue 8. Art. No.: CD013699.DOI: 10.1002/14651858.CD013699.

Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, Liu X, Wei L, Truelove SA, Zhang T, Gao W, Cheng C, Tang X, Wu X, Wu Y, Sun B, Huang S, Sun Y, Zhang J, Ma T, Lessler J, Feng T (2020) Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet, DOI 10.1016/S1473-3099(20)30287-5.

Binny RN, Hency SC, James A, Lustig A, Plank MJ, Steyn N (2020). Probability of elimination for COVID-19 in Aotearoa New Zealand. Te Pūnaha Matatini, 5 June 2020, https://www.tepunahamatatini.ac.nz/2020/06/06/probability-of-elimination-for-covid-19-in-aotearoa-new-zealand

Cousins S (2020). New Zealand eliminates COVID-19. Lancet 395, 1474.

Davies NG, Klepac P, Liu Y, Prem K, Jit M, CMMID COVID-19 working group, Eggo RM (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine*, DOI 10.1038/s41591-020-0962-9.

Ferretti L, Wyman C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, Parker M, Bonsall D, Fraser C (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*. DOI: 10.1126/science.abb6936.

Firth JA, Hellewell J, Klepac P, Kissler S, CMMID COVID-19 Working Group, Jucharski AJ, Spurgin LG (2020). Using a real-world network to model localized COVID-19 control strategies. Nature Medicine, DOI: 10.1038/s41591-020-1036-8.

Ganyani T *et al.* (2020) Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data. Eurosurveillance **25**, 2000257.

Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, Munday JD, Kucharski AJ, Edmunds WJ, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Funk S, Eggo RM (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Global Health: 8: e488-496.

James A, Plank MJ, Binny RN, Lustig A, Steyn N, Hendy S, Nesdale A, Verrall A (2020). Successful contact tracing systems for COVID-19 rely on effective quarantine and isolation. medRxiv preprint, DOI: 10.1101/2020.06.10.20125013.

Jarvis CI, Zandvoort KV, Gimma A, Prem K, CMMID COVID-19 working group, Klepac P, Rubin GJ, Edmunds WJ. (2020). Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Medicine*. DOI: 10.1186/s12916-020-01597-8.

Kang M, Song T, Zhong H, Hou J, Wang J, Li J, Wu J, He J, Lin J, Zhang Y (2016). Contact Tracing for Imported Case of Middle East Respiratory Syndrome, China, 2015. Emerging Infectious Diseases 22: 1644-1646.

Kucharski A, Klepac P, Conlan A, Kissler S, Tang M, Fry H, Gog J, Edmunds J (2020). Effectiveness of isolation, testing, contact tracing and physical distancing on reducing transmission of SARS-CoV-2 in different settings. Lancet, DOI 10.1016/S1473-3099(20)30457-6.

Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine.* DOI: 10.7326/M20-0504.

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

Prem K, Cook AR, Jit M (2017). Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Computational Biology.* 13(9): e1005697.

Tindale L, Coombe M, Stockdale JE, Garlock E, Lau WYV, Saraswat M, Lee Y-HB, Zhang L, Chen D, Walinga J, Colijn C (2020). Transmission interval estimates suggest pre-symptomatic spread of COVID-19. *MedRxiv preprint*. DOI: https://doi.org/10.1101/2020.03.03.20029983.

Verrall, A (2020). Rapid Audit of Contact Tracing for Covid-19 in New Zealand. *Ministry of Health.* 310
WHO & CDC (2015). Implementation and management of contact tracing for Ebola virus disease: emergency guideline. Retrieved from: http://www.who.int/csr/resources/publications/ebola/contact-tracing/en/

Appendix – model specification

Index variables

i, j – individual cases (i for index case, j for secondary case)

k, l – age groups

$k(i)$ – age group of case i

s – setting (home/work/school/casual)

Contact structure

$M_{kl}^{(s)}$ – contact matrix representing the average number of contacts that an individual in age group k has in age group l and setting s (home/work/school/casual). These were obtained from the results of Prem et al. (2017) using the number of daily contacts for New Zealand combined into 10 year age bands. The total numbers of work, school and casual contacts were scaled up by a common factor from the number of daily work, school and casual contacts to model an infectious period longer than 1 day. The total number of household contacts was set equal to the number of daily household contacts, to reflect the fact that household contacts do not usually change from one day to the next. The common scaling factor for work, school and casual contacts was set to be 3.0 to give a basic reproduction number of $R_0 = 2.6$ in the absence of any control measures. This corresponds to a reproduction number of approximately 2.4 with case isolation and no contact tracing. Contact matrices are shown in Supplementary Figure S1.

$Y_i^{(s)}$ – Individual multiplier for number of contacts individual i has in setting s. If $Y_i^{(s)} > 1$, individual i has a greater than average number of contacts in setting s and is more likely to be a superspreader. This parameter may be correlated between index case and secondary cases transmitted in setting s (see below).

$A^{(s)}$ – attack rate for contacts in setting s, assumed to be 20% for household contacts and 6% for non-household contacts (Kucharski et al., 2020), independent of age (Bi et al., 2020).

$C_i = 0.5$ if case i is subclinical and $C_i = 1$ otherwise. A case in age group k is subclinical with probability $p_{sub}(k)$

Branching process dynamics

The total expected number of new cases infected by case i between time t and time $t + \delta t$ is:

$$\lambda_i(t) = C_i \sum_{l,s} Y_i^{(s)} M_{k(i),l}^{(s)} A^{(s)} \ F_i(t) \int_t^{t+\delta t} W\left(\tau - T_{onset,i}\right) d\tau$$

where T_{onset}, i is the symptom onset time for case i, $F_i(t)$ is equal to c_{quar} if case i is in quarantine at time t, c_{iso} if case i is in isolation at time t, or 1 otherwise. The function W is the probability density function of a Weibull distribution (see Supplementary Table S1).

When case i infects a new case, the new case occurs in age group l and setting s with probability

$$p_l^{(s)} = \frac{Y_i^{(s)} M_{k(i),l}^{(s)} A^{(s)}}{\sum_{l,s} Y_i^{(s)} M_{k(i),l}^{(s)} A^{(s)}}$$

Each new secondary case j needs to be assigned individual multipliers $Y_j^{(s)}$ for each of the setting types. The multiplier for the setting ($s = s^*$) in which case j was infected is correlated with the multiplier for the index case i. This models assortative mixing, meaning people with high contact rates in setting s tend to be linked to other people with high contact rates and vice versa). Multipliers for different settings ($s \neq s^*$) are assumed to be uncorrelated with the multipliers for the index case. For simplicity, we assume that for work, school and casual transmission, the secondary case's multiplier for that setting is equal to the index case's multiplier for that setting with probability q, and is independent from the index case's multiplier with probability $1 - q$:

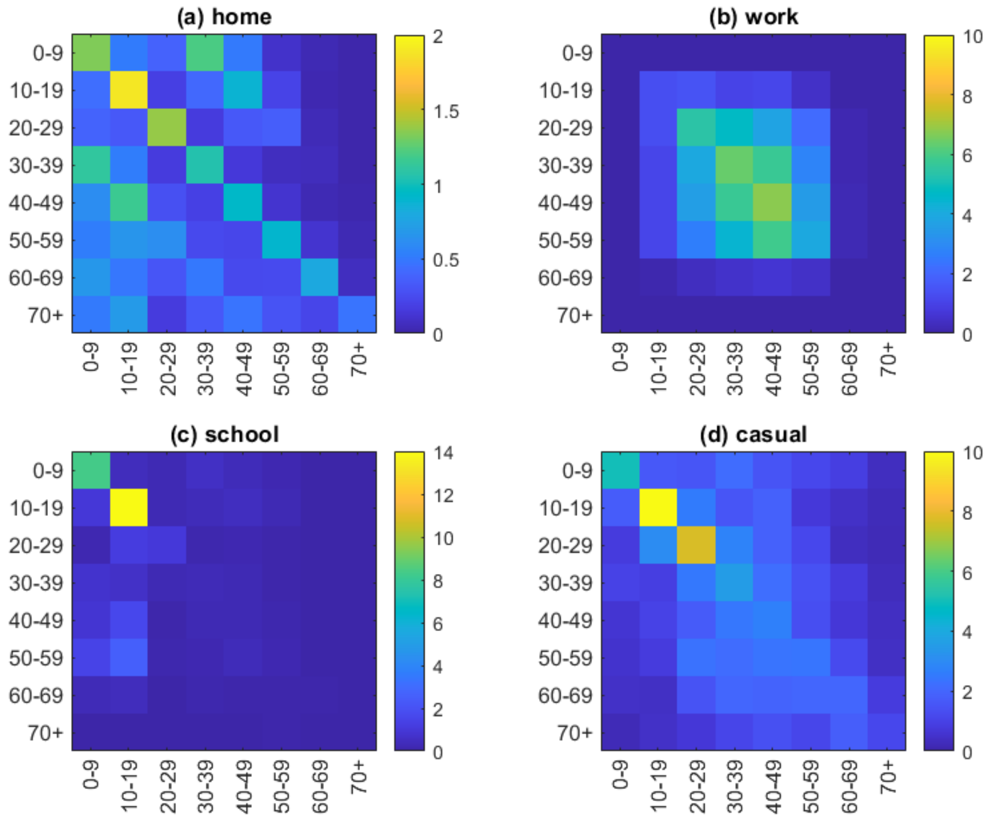
$$Y_j^{(s)} \sim \begin{cases} Y_i^s \text{ with probability } q & s = s^* \in \{\text{work,school,casual}\} \\ \Gamma\left(\mu = 1, \sigma^{(s)}\right) & s \neq s^* \end{cases}$$

Home transmissions are treated differently to model spread of the virus among a household containing a fixed number of members. Each case is assigned a household ID and an expected number of home contacts, given by $Y_i^{(\text{home})} \sum_l M_{k(i),l}^{(\text{home})}$ where $Y_i^{(\text{home})} \sim \Gamma\left(\mu = 1, \sigma^{(home)}\right)$.When a secondary case j occurs via household transmission it shares the same household ID as the index case i. The expected number of home contacts decreases by one with each new case in that household. This means that, eventually, transmission chains within a given household will always go extinct as the pool of susceptible household members is depleted. For simplicity, network effects and infection-induced immunity for work, school and casual contacts are ignored, i.e. all non-household contacts of the secondary case are assumed to be mutually exclusive of the non-household contacts of the primary case.

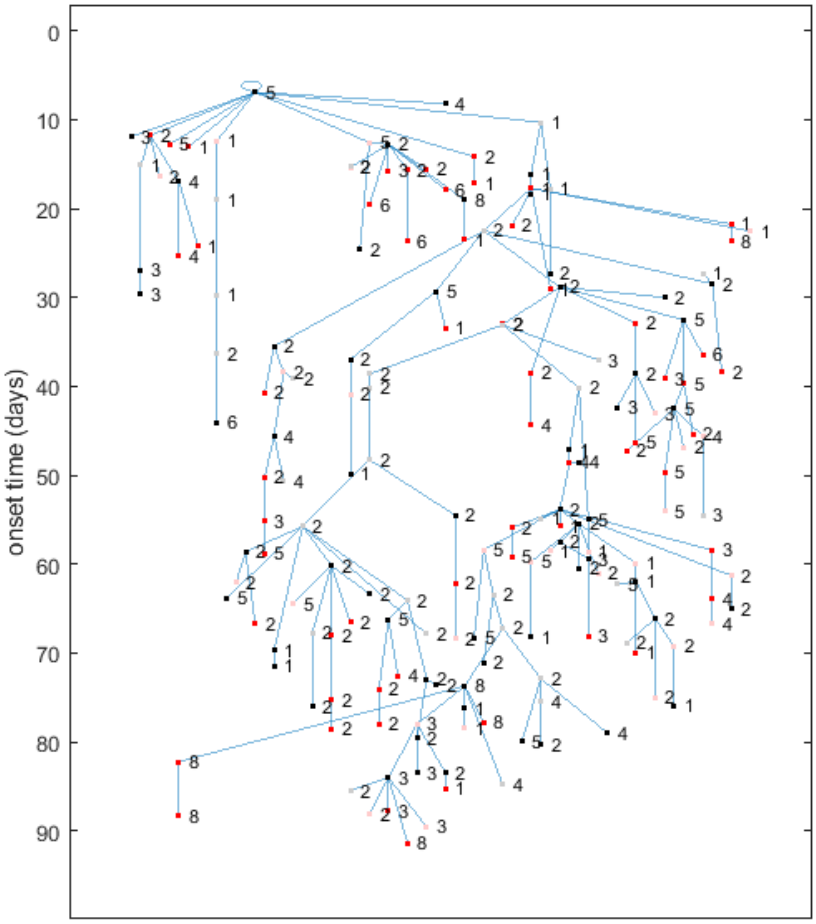
The parameter $\sigma^{(s)}$, the standard deviation of the distribution of Y_i^s, represents the degree of heterogeneity in individual contact rates in setting s. Higher values of $\sigma^{(s)}$ correspond to a greater degree of superspreading in the transmission dynamics (Lloyd-Smith et al., 2005). We set $\sigma^{(s)} = 1$ for home contacts and $\sigma^{(s)} = \sqrt{2}$ for work, school and casual contacts. This represents greater heterogeneity in contact rates outside the home and more potential for superspreading events in non-household settings, and corresponds to a negative binomial over-dispersion parameter of $k = 0.5$ for infections outside the home. Supplementary Figure S2 shows an example simulation of the branching process and contact tracing model.

Parameter	Value	Source
Secondary attack rate	$A^{(home)} = 20\%$ $A^{(work,school,casual)} = 6\%$	Kucharski et al. (2020)
Basic reproduction number (without case isolation or contact tracing)	$R_0 = 2.6$	Kucharski et al. (2020) Jarvis et al. (2020)
Variance in individual contact rates	$\sigma^{2\ (home)} = 1$ $\sigma^{2\ (work,school,casual)} = 2$	
Probability of cases in age group k (in 10-year age bands) being subclinical	$p_{sub}(k) = [40\%, 35\%, \ldots, 10\%, 5\%]$	Davies et al. (2020)
Relative infectiousness of subclinical infections	$c_{sub} = 50\%$	Davies et al. (2020)
Distribution of infection to onset (days)	$T_{onset} \sim \Gamma(\text{shape} = 5.8, \text{scale} = 0.95)$	Lauer et al.
Distribution of generation times (days)	$T_s - T_{onset} + t_p \sim Weibull(\text{shape} = 2.83, \text{scale} = 5.67)$	Ferretti et al.
Distribution of onset to isolation and testing (untraced contacts) (days)	$T_{iso} \sim \Gamma(\text{shape} = 1.3, \text{scale} = 5.2)$	NZ case data
Distribution of testing to test result (days)	$T_{result} - 0.5 \sim Exp(0.5)$	NZ case data
Distribution of test result to contact tracing for manually traced non-household contacts (days)	$T_{trace} \sim \Gamma(\text{shape} = 3, \text{scale} = 1)$	NZ case data
Proportion of secondary infections occurring before symptom onset (in the absence of case-targeted control)	$p_{pre} = 35\%$	Ganyani et al. (2020), Tindale et al. (2020)
Relative infectiousness after quarantine	$c_{quar} = 50\%, 20\%$	See scenario parameters
Relative infectiousness after isolation	$c_{iso} = 0$	See scenario parameters
Homophily parameter	$q = 0.5$	

Supplementary Table S1. Model parameter values and distributions.



Supplementary Figure S1. Contact matrices in 10-year age bands derived from results of Prem et al. (2017). Each matrix shows the average number of home, work, school or casual contacts that an individual in age group k has with an individual in age group l over the course of their infectious period. The number of home contacts are taken directly from the results of Prem et al. (2017) for New Zealand. The number of work, school and casual matrices are scaled up by a factor of 3 from the number of daily work, school, and other contacts in Prem et al. (2017) to allow for an infectious period longer than 1 day. The factor of 3 was chosen to make $R_0 = 2.6$ in the absence of any control measures.



Supplementary Figure S2. Example branching process simulation starting from a single infected seed case. Cases are represented as nodes with the vertical coordinate corresponding to time of symptom onset. Transmission routes are blue lines. Node labels indicate age group (1 = 0-10 years, 2 = 10-20 years, etc.), red indicates traced contacts, black indicates untraced, grey/pink indicates subclinical infections.