Digital Contact Tracing for COVID-19: A Primer for Policymakers

Rhema Vaithianathan*[1,2], Matthew Ryan[3], Nina Anchugina[1], Linda Selvey[4], Tim Dare[5], and Anna Brown[6]

[1]Centre for Social Data Analytics, Auckland University of Technology
[2]Institute for Social Science Research, The University of Queensland
[3]School of Economics, Auckland University of Technology
[4]School of Public Health, The University of Queensland
[5]Department of Philosophy, The University of Auckland
[6]Toi Āria: Design for Public Good, Massey University

May 28, 2020

Abstract

Recent modelling studies have shown that timely isolation of contacts can be effective in reducing the COVID-19 reproduction rate [1]. The epidemiological features of COVID-19 put it in the region where contact tracing is a viable public health strategy but it needs to occur at exceptional speed to control transmission. This need for speed has led policymakers to consider adopting contact tracing apps that enable rapid digital contact tracing. The rapidity and perfect scalability of digital contact tracing make it superior to manual contact tracing for an infection such as COVID-19. However, digital contact tracing has to enable notification of contacts the instant a user becomes positive. Ideally, with a rapid follow-up call by public health officials. A number of contact tracing apps currently deployed – such as the Australian COVIDSafe app – do not allow instantaneous notification and are therefore very limited in their ability to achieve control. Digital tracing apps also require sufficient uptake to realise the benefits. This requires building and maintaining strong social licence for use by demonstrating that the benefits to the user are high and that privacy and security risks are low. Like many medical devices and therapies, contact tracing apps are "credence goods", and claims of efficacy and safety must be largely taken on trust. Therefore Governments ought to subject these apps to robust and transparent evaluation so that users can judge for themselves whether downloading and use are worth the effort. This paper is intended for policymakers and provides an introduction to how digital contact tracing works, including a simplified model of its epidemiological foundations, and explores how to obtain enhanced uptake by achieving social licence for the use of the technology within the context of a public health emergency.

*Corresponding author. Email rvaithia@aut.ac.nz

Contents

1 Introduction

Tracing and isolating contacts of cases of COVID-19 is an important public health strategy for gaining control of spread [2]. However, manual contract tracing is not readily scalable, and as case numbers rise, jurisdictions outstrip their contact tracing capacity. Tracing becomes slow – and the resulting delays in tracing and isolating contacts means poor control of the epidemic. This paper is intended for policymakers and provides an introduction to how digital contact tracing works, including a simplified model of its epidemiological foundations, and explores how to obtain enhanced uptake by achieving "social licence" for the use of the technology.

In many countries, the public are concerned about the over-reliance on technocratic advise with regard to Government's response to COVID-19. As a recent opinion piece in The Guardian put it "*"Following the science" is a disingenuous policy because mathematical reckoning and human judgments are very different things*"[3]. This means that while formal infectious disease models [2, 1] are undoubtedly comprehensive, they are also difficult to understand for non-specialist policymakers who have to take much of it on trust. When it comes to justifying wholesale adoption of contact tracing apps by a country, there is a role for simpler heuristic treatment of complex models, that allow elected representatives and their advisors to better understand and communicate the science. A specific aim of this paper is to provide an easy to understand graphical treatment of the epidemiological justification for automated contact tracing apps.

Technology has played a major role in the control of COVID-19 for many countries. For example, South Korea and China have used mobile-phone and GPS data to find contacts who were proximate to the index-case [4]. This type of contact tracing is often non-consented, and is unacceptable in countries which place a high value on individual liberties and privacy. Even in China, it may have been counter-productive, by reducing the general levels of trust in Government [5].

Countries with strong traditions of individual liberty have been exploring digital contact tracing that uses low-energy Bluetooth technology rather than geo-location data; and requires user consent through the downloading of an app onto their smart phone. Avoiding geo-location data protects privacy; requiring consent respects individual liberty.

When a user downloads these contact tracing apps, they consent to exchanging information with other app users whose phones are within range of each other. Over time, each phone collects a set of "proximity logs" corresponding to other app users with phones within a predefined range. The exchange of information is not a phone number, but rather an encrypted ID. These "proximity logs" can be stored locally or on a centralised server. They can be deleted after a certain time period (such as 20 days).

1.1 Enhanced manual tracing

Australia and Singapore have implemented Bluetooth-based apps that have limited functionality. In particular, they do not allow automated notification of contacts, but are intended to make the job of manual tracing easier. The way they work is that when a user tests positive for COVD-19, the public health worker who is conducting manual contact tracing, has access to the proximity logs from the user's phone. These logs can be used – in concert with the information provided by the case themselves – to build a more comprehensive picture of contacts. The hope is that the proximity logs provide phone numbers of contacts that might have been unknown to the case.

These types of enhanced manual tracing apps – such as the Australian COVIDSafe app – are extremely limited in their ability to speed up contact tracing. The reason is that the public health worker cannot share the phone numbers in the proximity logs with the case because of privacy concerns. Typically, when people download these apps they are assured that their details will only be accessed by tracing staff and not shared. This means that the public health worker will have to conduct their tracing exactly as in the case of fully manual tracing. Once the task of manually gathering all the contact details has been completed, public health units will be able to cross check with the proximity log to find numbers that might not have been identified by the contact.

These apps offer more comprehensive contact tracing than manualised systems alone, but are not expected to greatly enhance speed. Since speed is the most crucial element for controlling spread of COVID-19, it is unlikely that these apps can reduce reproduction rates.

1.2 Digital contact tracing

With digital contact tracing, when a user tests positive for COVID-19 (or has symptoms), the app generates an *automated instantaneous notification* that asks the contact to isolate and/or get tested (and that a public health official will be in touch with them). Ideally, this should be followed up as soon as possible by a call from public health staff providing information about isolation, testing and monitoring of symptoms. However, the essential functionality is the ability to provide instantaneous in-app or SMS notifications.

These apps could be enhanced with machine learning tools which adjust notifications based on what is learned over time about the patterns of infections. For example, while initially they might notify all users who have been within 1.5 meters for more than 15 minutes of someone who tested positive, they could also adjust these parameters if data suggest that greater distance or shorter contact duration is sufficiently predictive of positive cases. The rest of this paper is concerned with these types of digital contact tracing apps – and not the Australian type of enhanced manual tracing apps.

2 Transmission of infections from first order contacts

The main preventive function of contact-tracing is to find and isolate first order-contacts, in order to prevent them from infecting others [1]. The impact of contact tracing on reducing transmission depends on the number of second-order infections generated by each infected contact [2].

The ability of contact tracing to manage the epidemic depends on two key epidemiological parameters. The first is the basic reproduction number, R_0, which is the average number of people to whom an infected person will transmit the infection. The second is the proportion of infections that are transmitted pre-symptomatically or from an asymptomatic source; this parameter is denoted by θ in Fraser *et al.* [6]. The higher these parameters, the harder it is to control an outbreak. For infections such as SARS, with low R_0 and θ values, standard isolation and contact tracing strategies are very effective. If R_0 or θ are very high then even the most efficient isolation and tracing strategies would be unable to contain the virus without additional social distancing measures. COVID-19 is an intermediate case – it has a low R_0 value but a moderately high θ, estimated at 0.62 by Ferretti *et al.* [1]. In the formulation of Fraser *et al.* [6], it sits in a boundary region (see their figure 2) where contact tracing is still a viable public health strategy but needs to occur at exceptional speed to control transmission effectively.

To get a better understanding of this, we start by outlining a simple model which quantifies the impact of delays in contact tracing in terms of number of infections caused by first-order contacts prior to isolation. We measure this delay from the time that the index case becomes symptomatic.

Clearly, transmission by first-order contacts is not the only mode of transmission that contributes to viral spread: the case could have transmitted pre-symptomatically, and transmission might have occurred from second and subsequent order contacts[2]. Our purpose is to provide a heuristic device for analysing the immediate infections that are prevented by tracing and isolation of first-order contacts, rather than to provide a fully elaborated model [1].

3 Analysis

3.1 Expected infections from first-order contacts

Appendix 1 provides the formula used to calculate the number of infections from first-order contacts. We assume an incubation time (time from infection to onset of symptom) and generation time distributions estimated by Ferretti *et al.* [1]. The incubation time is assumed to be lognormal with implied mean of 5.5 days and standard deviation of 1.43 days. The density of infections transmitted at any instant t as $R_0 w(t)$ where w is a Weibull pdf with shape

[1] Note that we are not considering the more limited type of contact tracing, which is to find the source of infection

[2] Additionally, to the extent that first-order contacts become symptomatic and are isolated as cases (rather than contacts), our formulation over-estimates the number of infections.

parameter 2.826 and scale parameter 5.665 and R_0 is the reproduction rate. We assume that the index case is tested and isolated as soon as they become symptomatic (clearly not always the case).

Figure 1 shows the expected number of infections that the first-order contacts will transmit for different R_0 values as a function of the delay until those contacts are isolated. The horizontal line measures the number of days that have lapsed since the case became symptomatic, and the vertical line measures the cumulative number of infections due to first-order contacts. The shape of the graph shows that the marginal impact of contact tracing delays is initially increasing and subsequently decreasing, with the curves becoming quite flat after about eight days.

The vertical intercepts in Figure 1 represents the number of infections that are passed on by first-order contacts before the case themselves are symptomatic – and aptly demonstrates the challenge in managing COVD-19 through contact tracing. For example, consider the line representing the reproduction rate $R_0 = 2.5$. This line intersects the vertical axis at around 0.5 – which means that the first-order contacts have already created half an infection, even before the index-case is symptomatic. This illustrates the problematically high θ value for COVID-19: significant first-order transmission may occur well before the index-case is isolated, allowing second-order transmission even prior to the start of tracing.[3]

Figure 1 confirms the large returns to timely contact tracing – which is at the heart of the case for automated tracing. It also shows the diminishing returns to speed. For example, a reduction in contact tracing delay from 4 to 2 days (with an R_0 of 2.5) avoids around 1 additional infection per initial index case. Reducing tracing time from 8 days to 6 days avoids less than 0.5. Indeed, contact tracing after 8 days provides little additional protection against first-order transmission (although potentially transmission from second and subsequent order contacts might still be prevented).

[3]This front-loading of infectiousness is particularly pronounced for the Weibull generation function which Ferretti *et al.* [1] found to have the best fit to their data. Figure 1 of their paper shows that the Weibull, amongst competing functional forms, gives a particularly high level of transmission within the first 48 hours.

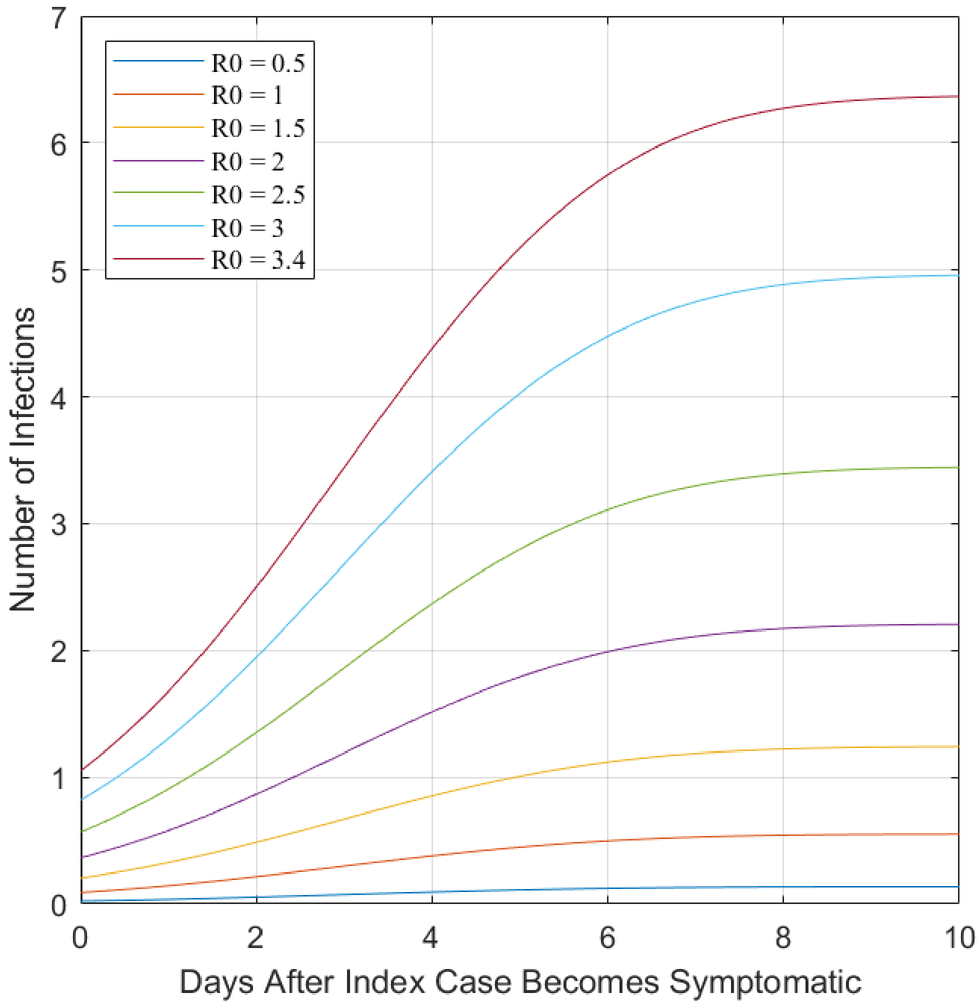

Figure 1: Expected Infections from First Order Contacts.

3.2 Speed of tracing and take-up

A crucial issue that confronts policymakers contemplating adopting digital contact tracing is how much take-up is needed to make it worthwhile. We now turn to a simple diagrammatic illustration of take-up. The target take-up rate depends on two things: (i) the maximum acceptable delay required to control the epidemic ("target delay") ; and (ii) the speed of manual tracing.

The target delay depends crucially on the efficacy of isolation and quarantining, the testing regime and other intervention package being proposed. If only a fraction of people who are identified as contacts are able to be successfully isolated, positive test results are slow to be returned, and not widely available for people as soon as they are symptomatic then the target delay would of necessity have to be shorter. If on the other hand, the app is part of a package of interventions such as physical distancing and targeted school closures that reduce transmission then target delay could be lengthened. Ferretti *et. al.* [1] show that with perfect testing and isolation, but no additional interventions, the target delay required to induce an effective reproduction rate less than 1 is shorter than 72 hours following the onset of symptoms.

Figure 2 illustrates the infection curve for $R_0 = 2.0$. We assume that positive test results are received within 24 hours, at which point notifications are sent via the app asking all contacts to isolate and manual contact tracing commences. If there is 50% take-up of the digital tracing app (as a percentage of the population), then approximately 25% (50% squared) of contacts will be notified automatically, as both parties must have the app installed. For the remaining 75% of contacts we rely on manual tracing.[4] For purposes of illustration, we assume manual tracing takes five days after test results.[5] Drawing a straight line between the two extreme scenarios (100% automated tracing and 100% manual tracing), we can see that the average number of infections generated by first-order contacts falls from 2 (with fully manual tracing) to around 1.6 with partially automated tracing (assuming 50% uptake of the app).

Using this approach, in Figure 3 we draw the combination of manual tracing days and take-up that can achieved a target delay of 2 and 3 days.[6] Suppose manual tracing can be completed within 4 days, then combined with a a 60% uptake, Figure 3 suggests that we could achieve a target of 72 hour (equivalent) delay – that is, achieve the same degree of infection control as if all tracing was done within 72 hours.

The question for policymakers is whether it is easier to achieve a 60% uptake and a 4 day manual tracing target or achieve manual tracing alone within 72 hours. Clearly, the challenge to maintaining a manual tracing target of 72 hours is that contact numbers and case numbers fluctuate depending on where in the epidemic curve we are and what other interventions are in place. Unfortunately, just as cases grow or social-distancing rules are eased, and speedy tracing is ever more important, manual tracing will be put under pressure and tracing times will rise. On the other hand, the app is perfectly scalable and notification is instantaneous regardless of the number of contacts or cases.

[4] In fact, this is probably unfair to automated tracing. It assumes that app take-up is independent across individuals. However, it is plausible that app take-up will be positively correlated within social networks linking the case to many of his or her contacts.

[5] We have no information about the manual contact tracing processes in various jurisdictions. The five days assumption is based on conversations we have had with Australian contact tracing staff - and is also the median between the "short" and "long" assumptions in [2]. This duration will vary by case and the number and nature of contacts, as well as the capacity of contact tracing teams. Social distancing measures increase the speed and ease of manual contact tracing, as it aims to reduce the number of contacts, particularly casual contacts. As social distancing measures are relaxed the number of contacts will increase as will the difficulty of manual contact tracing.

[6] We provide the equations for this calculation in Appendix 2.

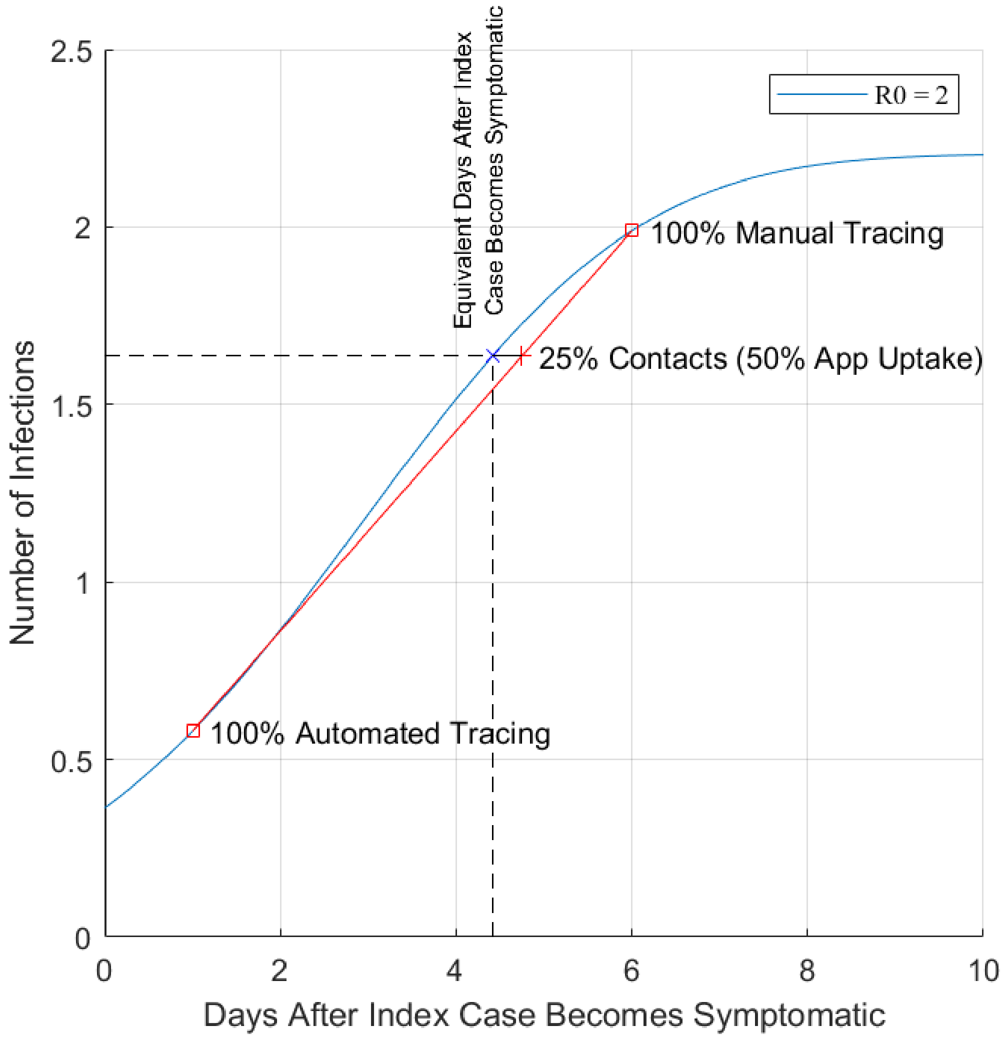

Figure 2: Illustration of the impact of automated tracing.

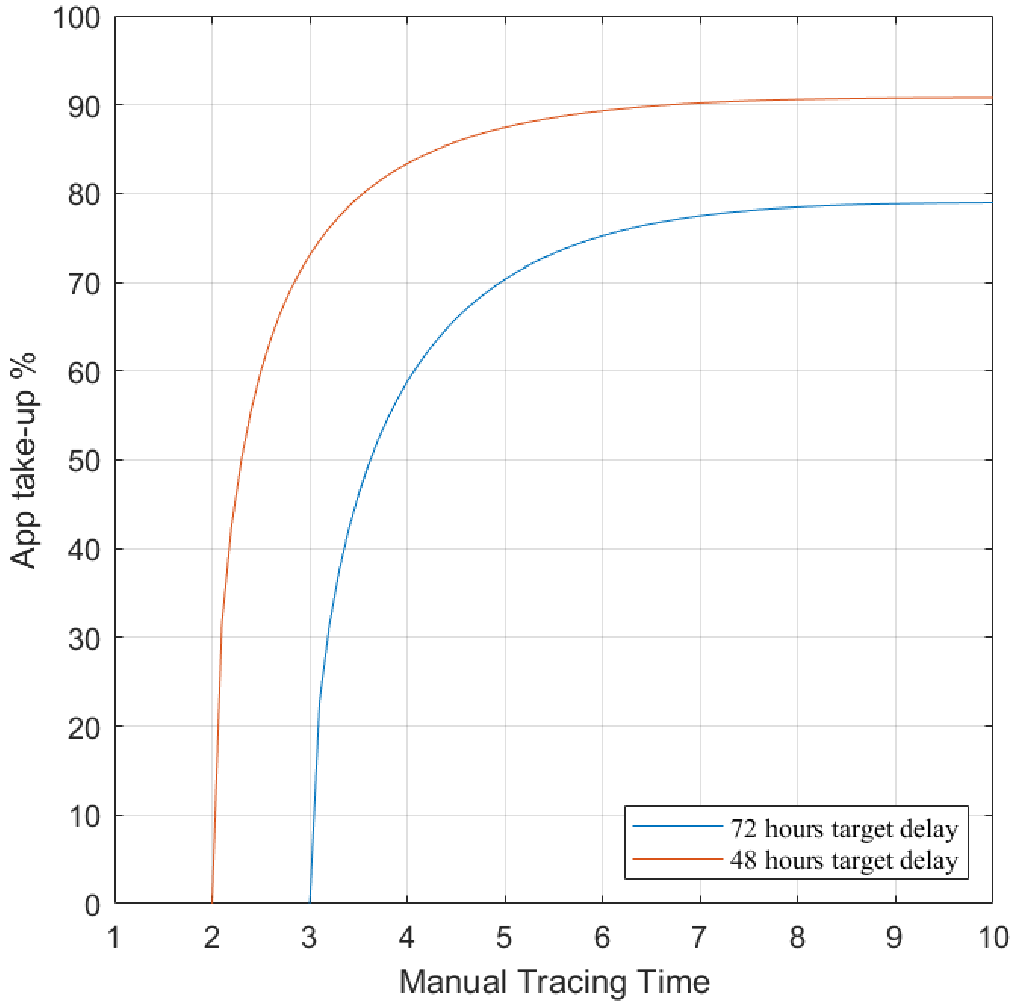

Figure 3: The combination of manual tracing and take-up that achieves a target delay.

4 Maximising take-up through ethics and social licence

The use of digital tools to address important social needs is often initially greeted with concern. Given the importance of high uptake in Figure 3, it is crucial that people are satisfied that such concerns about the app and how it will be used have been recognised and addressed.

The ethical justification of the app rests upon its capacity to deliver significant benefits to communities and individuals in ways which respect legitimate concerns about, for instance, consent, the security and use of information, the preservation of a role for human judgement, and the possibility that the app will exacerbate existing social inequalities. Polling data from Singapore and Australia both suggest that privacy is a key concern [7].

Barriers to take-up and trust in apps was discussed at a recent international roundtable of contributors from eight countries including the UK, Ireland, Germany, Switzerland, Norway, Denmark, Israel and New Zealand [8]. Governments should commit to clear guidelines addressing these predictable concerns [9, 10]. Where people are satisfied that they have been recognised and addressed – where they trust that their data will be used as they have agreed, and accept that enough value will be created – they are likely to be more comfortable with the use of that data. This acceptance is referred to as "social licence" [11].

Community engagement is recognised as being a key vehicle for achieving social licence. Engagement provides crucial information on issues of concern, can help with clear messaging and (ideally) guides the design and build of the app. However, the time-bound nature of responding to the current crisis, coupled with dynamic political environments and limits of legislation have seen little attempt at concentrated community engagement to build public trust. Robust and interactive public engagement (however big or small) should be prioritised by Governments.

Research in New Zealand suggests that people are comfortable with sharing their data as long as they see value in it for themselves, their family and community; and if they trust that their data will be kept secure. A key consideration for Governments is therefore how they will communicate the value of the tool (both initially and on an ongoing basis) and show that legitimate ethical concerns, including those about privacy and security, have been addressed.

4.1 Communicating the individual incentive for uptake

While there are clear diffuse and public health benefits, there are also significant (but perhaps less obvious) personal benefits from downloading the app. These benefits are often ignored in the commentary about this technology.

In particular, the fact that the majority of infections are transmitted to household members or work colleagues means that app users obtain a clear advantage in keeping family and co-workers safe. Of course, if the index case is a household member, then notification would likely be almost immediate whether or not the contact has the app. The app really comes into its own if the index case is a casual contact or work colleague. In this case, it could take manual tracers days to find the contact. By having the app, the delay would be avoided, reducing transmission from the contact to their family members and co-workers.

Additionally, the users' own value from the app is increasing in the rate of take-up amongst immediate social contacts, rather than take-up at the national level. For example, if all work colleagues have the app, then the rates of transmission within that social circle would fall significantly even if take-up is poor at the national level. This feature, sometimes referred to as a "local public good" property [12], means that users acting in their own narrow interest can generate the high levels of national take-up necessary for significant national benefit.

4.2 The challenge of credence goods

Goods or services where users are unable to judge efficacy or value – even after consuming the good or using the service – are called *credence goods* [13]. They are most commonly purchased from highly trusted (and regulated) professionals such as medical providers, where claims of efficacy are required to be backed up by evidence. For example, opting for treatment A over treatment B (where B might be doing nothing) relies on an assurance that A is more likely to produce the better outcome, but this fact cannot be directly "experienced" by the patient, even *after the treatment*.

The challenging nature of contact tracing apps is that once downloaded, it sits passively on user's phone. While on occasion, a user might be notified to isolate, a user cannot be sure that the app is working let alone providing the protection that they expect or is not creating some unintended adverse harms. Indeed, it was revealed that more than a week after the Australian app was launched, the data being collected was not actually actionable, as there had been no procedures set it place to access the data [14]. Therefore, claims that the app is having a beneficial effects – or any effect at all – must be largely taken on trust.

It is possible that a lack of evidence and low uptake will diminish trust, and have negative spillover effects to other recommendations that Government might be making in managing the pandemic (such as social distancing or wearing of masks). This suggests that Governments promoting contact tracing apps ought to ensure that their statements about efficacy are restricted to those that can be justified by the evidence.

Additionally, committing to ongoing real-time evaluation, that allow citizens to quickly judge whether the app has had the expected impact on COVID-19 transmission will increase trust as well as take-up. A positive finding showing that the app decreased delays in contact tracing and reduced infections by first-order contacts amongst those who had the app – will encourage uptake. Without this ongoing feedback loop from impact to uptake it is unlikely that the requisite diffusion will occur. Skeptical citizens will want to know that the app is actually doing what was intended.

4.3 An evaluation framework

While randomised control trials remain the most "credible" way of establishing evidence of efficacy and safety in health interventions, the urgency of the pandemic make it challenging to implement. It is likely that a quasi-experimental approach would have to be taken [15].

- **Communities who have high app uptake should have lower infection reproduction rates (reporting symptoms or testing positive) than people without the app.**

  Using data on take-up rates by postal-codes, one can readily calculate the correlation between take-up and local transmission rate. However, there are multiple problems with inferring that this association is causal. In particular, uptake is based on each user's assessment of expected transmission rates and take up of the app in their community. This means that uptake rate is not truly “exogenous” (in the parlance of evaluation science, uptake is correlated to unobservable factors that effect the outcomes of interest). In such situations, quasi-experimental methods rely on have suitable identification strategies – including a rich set of variables that can be used as instruments.

An added complication is that users who are part of a high-uptake group will receive more notifications and therefore are more likely to be tested or report positive symptoms. Moreover, those in higher uptake groups have a greater incentive to get tested knowing that this will generate valuable notifications to the user's social circle. A simplistic observational study that doesn't carefully account for such reverse causality, might find a contradictory result that high uptake leads to *higher* rates of confirmed infections.

This means that a credible evaluation strategy to establish whether high take-up is related to better control will depend on the availability of enough data to overcome the various threats to validity and a robust identification strategy. Access to rich and detailed research data – including the full and dynamic anonymised network data – is crucial to provide robust evidence of effectivness .

- **Proximity in the network to people with COVID-19, should be correlated to a user's propensity to report a positive test or symptoms.**

  Theoretically, if the app has complete coverage; all users report positive tests and symptoms; and environmental transmission is negligible, then the *only* known contacts should test positive. If a large proportion of users who test positive were never notified as a contact, this suggests that the notification system is ineffective.

  The implication of this logic is that a useful measure to track efficacy would be to monitor the number of users who test positive and were previously notified as contacts as a proportion of all app-users who test positive.

  However, such a measure will not be an unbiased statistic, because once people receive notification that they are contacts, they may be more inclined to get tested. For this reason, quasi-experimental method would have to be used to get at an unbiased measure of efficacy.

  For example, each individual with the app, maybe be assigned a risk score which excludes the status of first-order contacts, but is based on the number of second and higher order users who are positive. This measure – which is an exposure measure based on higher order contacts – is not observable by the user and therefore doesn't have the same behavioural response as the first-order contacts. It can therefore be used for validating the app. Firstly, by directly establishing that a high risk score is correlated with propensity to report a positive test. Secondly, by using it as a running variable for a fuzzy Regression Discontinuity design or an instrument in an instrumental variable design[16, 15]. The need to use second and higher order contacts as identification instruments points to the importance of ensuring sufficient data of usage and other metrics are gathered and retained.

4.4 Evaluation and centralised vs. decentralised models

The evaluation strategies itemised above are by no means exhaustive – but does point to some of the real challenges involved in establishing the efficacy of the app in the absence of an RCT. There is presently an active debate amongst researchers as to whether data collection should be decentralised and left on users' phones or pushed regularly to a centralised server (`https://github.com/DP-3T/documents/blob/master`).

The decentralised approach to contact tracing means that data on the full anonymous network is not retained. Without such information, obtaining unbiased estimates of impact is made more challenging [17].

Donated data doesn't solve this problem, because network data are, by their nature, data about relationships between individuals. Just because person A donates their data, it doesn't provide information on their tie to person B – unless the latter also agrees to donate their data. This then makes it less useful because of the need for double-coincidence for information about the network to be gleaned.

4.5 Comfort for citizens regarding their privacy and security

Most Governments have been embroiled in data breaches, including public health systems such as the NHS. Despite, these breaches people to continue to use the health system – because they trust that these medical services have value for themselves and their families.

While there are many attractive privacy design features that can be embedded in the app, most users are not privacy design experts – and communicating these features can be very challenging.

Rather than technical solutions to privacy – which is hard to communicate – emphasising the processes for achieving community comfort, especially amongst low trust groups, is an alternative approach. For example, when the Data Futures Working Group in New Zealand designed national guidelines for trusted data use, rather than speaking to privacy experts, they commissioned and led a "national conversation" across the country that asked ordinary New Zealand people what made them comfortable about sharing their data – and what sort of questions would have to be answered, or assurances given, to make them more comfortable sharing their data [11]. This type of process allows policymakers to speak to the concerns of ordinary citizens – and eschew some of the more technical and ephemeral privacy debates.

5 Conclusion

The increasing returns-to-scale and perfect scalability of digital contact tracing make it superior to manual contact tracing for a pandemic such as COVID-19, where speed of tracing and isolation are key to reducing reproduction. Manual tracing is likely to suffer from diminishing marginal product in the short run as more manual tracers are brought into service who have less experience and expertise [18].

However, while digital contact tracing has distinct advantages, it requires the right functionality – and enhanced manual tracing apps such as Australian COVDISafe, which do not enable instantaneous notification, will not achieve the enhanced speed required for success.

Apps also require sufficient uptake by populations to realise these benefits. This means that building and maintaining social licence for use has to be demonstrated – namely that the value to the user is high and that privacy and security risks are low.

There are a number of benefits that are worth communicating to users. Firstly, the app has large personal benefits because the majority of infections are to family and co-workers. This means that using the app delivers the greatest benefit to the users' immediate social groups.

Additionally, the app has features of a local public good (such as a local park), wherein high uptake in one's immediate social network is sufficient to realise value for the user. Individual work groups, extended families or circles of friends can maximise their own group protection by coordinating amongst themselves to download the app.

A challenge for Governments is that the app is a credence good, which means that users have limited ability to judge for themselves whether the app is benefiting them. It requires stronger levels of trust in the Government's advice, and also a scientific system for evidence-based feedback. For this reason, an independent impact evaluation should be considered.

While it is tempting to respond to privacy concerns with more and more complex privacy and cryptography features, in general citizens lack the ability to distinguish amongst alternative designs or to be reassured by them. An alternative approach is to talk directly to citizens (especially low-trust communities) through a (ideally national) conversation – about what questions they need answered and which concerns must be allayed before they feel comfortable downloading the app.

The danger is that without the data collected by a centralised approach, a robust evaluation will not be feasible. This is not to say that evaluation is not possible within a DP-3T framework, rather by eschewing data on the full anonymised network, those Governments choosing to adopt this framework would be advised to develop a robust evaluation strategy first.

Assurances of privacy are of little use if similar assurances of efficacy and safety cannot be provided.

Acknowledgement

This work was funded and developed by researchers and staff at the Centre for Social Data Analytics (based at Auckland University of Technology and the Institute for Social Science Research, The University of Queensland) with contributions from Auckland University of Technology, The University of Queensland, The University of Auckland and Massey University.

A Appendix 1

This section describes the calculation of expected infections from first-order contacts that would be prevented through contact tracing. Note that only first-order contacts are quarantined, i.e., there is no recursion.

- Let δ be the number of days after index case becomes symptomatic that contacts are isolated.
- We assume that no infected (first-order) contact was infected by someone other than the index case of interest.
- The incubation and generation time distributions from [1] are used for the calculations.
- Incubation time is time from infection to onset of symptoms. The Oxford group [1] use the estimate of [19], which is a lognormal with implied mean of 5.5 days and standard deviation of 1.43 days. Let $f(\tau)$ denote the pdf of the incubation time variable.
- The Oxford group [1] estimate the function $\beta(t)$ (= the "number" [actually, density] of infections transmitted at instant t) as $R_0 w(t)$ where w is a Weibull pdf with shape parameter 2.826 and scale parameter 5.665.
- The expected number of infections via first-order contacts prior to isolation is therefore equal to:

$$\begin{aligned} & R_0^2 \int_0^\infty f(\tau) \int_0^\tau w(t) \left[\int_t^{\tau+\delta} w(s-t)\ ds \right]\ dt\ d\tau \\ = \ & R_0^2 \int_0^\infty f(\tau) \int_0^\tau w(t)\, W(\tau+\delta-t)\ dt\ d\tau \qquad (*) \end{aligned}$$

  where $W(x) = 1 - \exp\left(-(x/\lambda)^k\right)$ is the cumulative distribution function for the Weibull ($\lambda = 5.665$ and $k = 2.826$).

  - For each possible incubation time (τ), we consider infections transmitted over the interval $[0, \tau]$ and for each of the contacts infected during this period, compute the expected number of infections transmitted by that contact during the interval $[t, \tau + \delta]$, where $t \in [0, \tau]$ is the time at which the contact was infected.

Our analysis is based on the quantity $(*)$, considered as a function of the tracing delay, δ.

B Appendix 2

The calculations behind Figure 3 are as follows. Let $H(\tau)$ denote the function $(*)$. Thus, $H(\tau)$ is the expected number of infections transmitted by first-order contacts when the tracing delay is δ days from the time that the index-case becomes symptomatic. Let x denote the number of days required for manual tracing (the quantity measured on the horizontal axis of Figure 3) and $y \in [0, 1]$ the fraction of the population that has the app (the quantity measured on the vertical axis, though expressed there as a percentage). We assume a 24 hour tracing time when using the app. The relationship between x and y indicated in Figure 3 is defined by the following equation:

$$y^2 H(1) + (1 - y^2)H(x) \;=\; H(\delta)$$

where δ is the target equivalent delay (2 days for the red curve; 3 days for the blue curve). It shows the combinations of x and y that achieve the target delay; that is, the combinations of x and y that give an expected number of transmissions from first-order contacts equivalent to a scenario in which all tracing occurs in δ days).

References

[1] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser, "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing," *Science*, 2020.

[2] J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, F. Sun, *et al.*, "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," *The Lancet Global Health*, 2020.

[3] K. Malik, "For all its sophistication, AI isn't fit to make life-or-death decisions." `https://www.theguardian.com/commentisfree/2020/may/16/for-all-its-sophistication-ai-isnt-fit-to-make-life-or-death-decisions-for-us`, May 2020.

[4] S. Park, G. J. Choi, and H. Ko, "Information technology–based tracing strategy in response to COVID-19 in South Korea—privacy controversies," *JAMA*, 2020.

[5] H. Cho, D. Ippolito, and Y. W. Yu, "Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs," *arXiv preprint arXiv:2003.11511*, 2020.

[6] C. Fraser, S. Riley, R. M. Anderson, and N. M. Ferguson, "Factors that make an infectious disease outbreak controllable," *Proceedings of the National Academy of Sciences*, vol. 101, no. 16, pp. 6146–6151, 2004.

[7] M. Farr, "Guardian Essential poll: suspicions about tracing app offset by approval of COVID-19 response." `https://www.theguardian.com/australia-news/2020/apr/28/guardian-essential-poll-suspicions-about-tracing-app-offset-by-approval-of-covid-19-response`, Apr 2020.

[8] "Centre for Data Ethics and Innovation blog "The ethics of contact tracing apps: International perspectives"." `https://cdei.blog.gov.uk/2020/05/12/the-ethics-of-contact-tracing-apps-international-perspectives/`.

[9] J. Morley, J. Cowls, M. Taddeo, and L. Floridi, "Ethical guidelines for SARS-CoV-2 digital tracking and tracing systems," *Available at SSRN 3582550*, 2020.

[10] T. Dare, "The ethics of contact tracing apps." `https://www.newsroom.co.nz/2020/05/06/1158576/the-ethics-of-contact-tracing-apps`, May 2020.

[11] "Our data, our way: what New Zealand people expect from guidelines for data use and sharing. findings from public engagement." `https://trusteddata.co.nz/massey_our_data_our_way.pdf`, Feb 2017.

[12] C. M. Tiebout, "A pure theory of local expenditures," *Journal of Political Economy*, vol. 64, no. 5, pp. 416–424, 1956.

[13] M. R. Darby and E. Karni, "Free competition and the optimal amount of fraud," *The Journal of Law and Economics*, vol. 16, no. 1, pp. 67–88, 1973.

[14] "Coronavirus app tracing capability not yet operational, despite 4 million downloads." `https://www.abc.net.au/news/2020-05-02/coronavirus-app-currently-not-fully-operational/12208924`, May 2020.

[15] T. Bärnighausen, P. Tugwell, J.-A. Røttingen, I. Shemilt, P. Rockers, P. Geldsetzer, J. Lavis, J. Grimshaw, K. Daniels, A. Brown, *et al.*, "Quasi-experimental study designs series—paper 4: uses and value," *Journal of Clinical Epidemiology*, vol. 89, pp. 21–29, 2017.

[16] J. Bor, E. Moscoe, P. Mutevedzi, M.-L. Newell, and T. Bärnighausen, "Regression discontinuity designs in epidemiology: causal inference without randomized trials," *Epidemiology (Cambridge, Mass.)*, vol. 25, no. 5, p. 729, 2014.

[17] A. Chandrasekhar and R. Lewis, "Econometrics of sampled networks," *Unpublished manuscript, MIT.[422]*, 2011.

[18] B. Armbruster and M. L. Brandeau, "Contact tracing to control infectious disease: when enough is enough," *Health Care Management Science*, vol. 10, no. 4, pp. 341–355, 2007.

[19] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, "The incubation period of 2019-nCoV from publicly reported confirmed cases: estimation and application," *medRxiv*, 2020.