Digital contact tracing: comparing the capabilities of centralised and decentralised data architectures to effectively suppress the COVID-19 epidemic whilst maximising freedom of movement and maintaining privacy.

Christophe Fraser[1], Lucie Abeler-Dörner[1], Luca Ferretti[1], Michael Parker[2], Michelle Kendall[1], David Bonsall[1]

[1]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK
[2]Wellcome Centre for Ethics and the Humanities and Ethox Centre, University of Oxford, UK

7 May 2020

Introduction

The aim of contact tracing is to provide an early warning to people that they have come into contact with someone who may be infected with COVID-19. Receiving this warning means that people can take action to prevent spreading the virus onwards, especially to vulnerable people. This action may involve quarantine for a number of days, or could be just an increased level of awareness, hygiene and physical distancing.

Digital contact tracing has the potential to assist people in receiving this warning as quickly as possible, before they become infectious. This speed advantage of a digital process is particularly important for COVID-19, where the virus is in many cases spread before people develop characteristic symptoms. Digital contact tracing may also help notifications reach people who otherwise wouldn't be recalled, such as people who have shared a space on public transport.

The aim of this report is to provide an overview of different considerations in the design of systems for digital contact tracing, and to address recent controversies that have arisen in the choice of possible architectures for digital contact tracing. The choice of architecture should be informed by considerations of what the system is trying to achieve. The aim of the intervention is to 1/ contribute to maintaining ongoing control of the epidemic as part of an integrated Test, Track and Trace strategy 2/ Minimise the number of individuals whose lives are disrupted by requests to isolate, distance, or quarantine whilst maintaining epidemic control 3/ Minimise the invasion of privacy needed to achieve aims 1 and 2.

Epidemiological and public health requirements for a COVID-19 contact tracing app

Digital contact tracing apps are a key component of many national strategies for suppressing COVID-19. Designing an effective app requires expertise from diverse fields including information security [Vaudenay, 2020], ethics [Morley et al. 2020; Parker et al. 2020], and behavioural science [Abeler et al. 2020; Altmann et al. 2020]. For these apps to achieve their core purpose of stopping COVID-19 transmission, epidemiological considerations must be at

the heart of their design. We present five key epidemiological and public health requirements which COVID-19 contact tracing apps should satisfy:

1. Sensitively and specifically identify infectious individuals

The purpose of contact tracing apps is to send notifications only to potentially infectious individuals. Failure to send notifications to infectious individuals by missing meaningful contact events, and falsely sending notifications for non-infectious individuals, should both be minimised. This means minimising both the number of false positives and false negatives in Table 1.

Table 1. Four outcomes of the contact tracing process:

All contacts	**Infected**	**Not infected**
Notified	True positives	False positives
Not notified	False negatives	True negatives

To achieve this, the algorithm must be **adjustable**. In this rapidly developing epidemic, our knowledge of the disease is continuously improving. It may vary across populations and social networks, and over time through the impact of interventions such as school closures. To fix an algorithm before the app is released, with no capacity for change, is to risk locking the app in a state of poor performance.

If an individual has erroneously received a notification instructing them to quarantine, either from malicious use or an index case reporting symptoms but later receiving a negative test result, it should be possible to send them an **early release notification**. If an individual has received a quarantine notice following contact with more than one index case, release should only be given if all index cases are confirmed to be uninfected.

2. High user uptake and adherence

Even at low levels of **uptake**, apps can reduce transmission and have a protective effect on the population, including those without smartphones [Hinch et al. 2020]. However, achieving epidemic control in the **absence** of other strong interventions will typically require uptake by around 60% of the population. Once installed, an app will only affect the epidemic if users follow the instructions it issues. Trust in the app and a positive user experience are therefore essential components for digital contact tracing to be effective. Any design choices which could hinder **adherence** should be avoided, such as frequent erroneous notifications. Further, in order to resume travel without increasing the risk of epidemic resurgence, apps should be **interoperable** across public health regions. Without this, users would have to install multiple apps, with a likely

detrimental effect on accuracy, uptake, and adherence.

3. Rapid notification

The time between onset of symptoms in an index case and the quarantine of their contacts is of key importance to COVID-19 contact tracing; any delay reduces its effectiveness [Ferretti & Wymant et al. 2020] and delays much beyond 2-3 days, reach the majority of contacts after they have infected others, removing all public health benefit entirely. Where a design feature introduces a short delay, such as awaiting test results, it should only be implemented if the delay is outweighed by other gains such as in specificity, uptake, adherence, etc. The relative impacts of these factors should be quantitatively compared at the design stage in open-source models such as [OpenABM-COVID-19].

4. Integration with local health policy

The advice given by an app notification should be adjustable to remain consistent with current local health policies. Ideally the app should be integrated within the full range of public health interventions such that it serves as a link to accessing further advice, medical care and testing.

5. Ability to evaluate effectiveness transparently

Users must be confident that notifications are based on the best available evidence. The contact tracing algorithm should therefore be transparent, auditable, under oversight, and subject to review. Any intervention in an epidemic should be evaluated, both alone and in combination with other measures. Multiple independent approaches should be used in these evaluations and the metrics of success and failure should be decided upon in advance.

Aggregated data (not linked to individuals) is essential for evaluating and improving the performance of the app. Although some such information could perhaps be gained via surveys, there are strong practical and ethical justifications for gathering these data via the app itself. These justifications are particularly concerned with the speed and scale of the epidemic, and the huge social and economic costs of failing to control it.

- **Instantaneous summary statistics**
  Summary statistics such as the numbers of index cases and contacts should be available. This data will be crucial for evaluating the app and rapidly identifying malfunctions or malicious use, as well as being extremely valuable for public health planning.

- **Geographical summary statistics**
  Knowledge of local uptake is vital for assessing the app's effectiveness and the reliability of its evaluation of individual risk. Without this, individuals in areas with low app uptake and high incidence of COVID-19 could erroneously be given the impression that they are at low risk, especially if there is an inverse correlation between epidemic growth rate and app uptake.

Summary of different proposed data architectures

Several architectures have been proposed for digital contact tracing, and there has been some controversy on which is the best choice. So far the debate has centred on privacy considerations which are an important - but not the only - ethical consideration in designing digital contact tracing [Parker et al. 2020].

There are broadly two classes of architecture, the so-called centralised architecture, and the decentralised architecture.

In the centralised architecture, the process of passing information from the individual who is ill to their contacts is mediated by a central server. The central server allocates IDs to individual phones. When an individual is diagnosed with either suspect or confirmed COVID, their contact history is uploaded to the server. The server performs some computations, and based on those computations, notifications are sent to some of the contacts.

In the decentralised architecture, the process of passing information from the sick individual to their contacts is done through direct broadcasts of lists of sick individuals over the phone network. Each phone then regularly does a computation to determine whether the phone has been associated with a risky contact with one of the sick individuals. Decentralised systems use a central server for updating tracing rules, and for collecting aggregate summary statistics.

Each specific implementation may use one or the other, and may mix elements of both. Broadly, the NSHX COVID-19 App uses a centralised architecture, and Google/Apple and D3PT both use a decentralised architecture. Google/Apple perform calculations in the operating system passing summaries to the App they support, whereas D3PT performs computations in the App.

Here, we propose an assessment based on our current understanding of these architectures. Proposed solutions are evolving and our assessment will continually update as more details appear.

Summary assessment

Many of the basic tracing functionalities are similar between the centralised and decentralised systems, including the ability to report aggregate data.

There are three important differences that favour the centralised model, and three that favour the decentralised model.

Benefits of centralised systems

1. Evaluation, learning and improvement of the notification process

Consider the following scenario. A passenger infected with COVID-19 gets on a bus and sits down next to someone. There are 15 other people on the bus. Some are sitting closer to the infected passenger, than others. This information will be recorded by the apps. Are all passengers at risk of being infected, or just the person sitting next to the infected passenger? Both the centralised and decentralised apps can be tuned to reflect different assumptions. However, by analysing the contact patterns submitted anonymously by all infected users, the centralised app will be able to answer this question so that future alerts can be more precisely targeted, thus reducing unnecessary notification, and also reducing infections, protecting hospitals and saving lives. The decentralised app will not be able to learn how to make alerts more or less targeted.

At present, not enough is known about the spread of COVID to make sure that the initial configuration of the apps (centralised or decentralised) strikes the right balance between notifying too many people (all the people on the bus in the example above), or too few (just the person sitting next to the infected person. Being able to evaluate, learn and improve the app, which can only be done by analysing a central database, will allow the app to quickly improve in the first few weeks of use.

One possible solution to this problem in the decentralised model is to ask many users to ‘donate’ their data on a regular basis. However this may result in more privacy problems than the centralised server, and is not currently possible with the architecture developed by Google and Apple, nor is it expected to be implemented in the future.

2. Safe notification based on self-reported diagnosis.

With the centralised model, it is possible to allow notifications to be sent based on self-reported diagnosis, and to later update these notifications based on testing results. This means that individuals could receive an Amber notification that they have been in contact with a suspect case of COVID. This would be upgraded if the case becomes positive, to a Red notification that the person has come into contact with a case, or removed if the person tests negative. Because

COVID is transmitted before people develop symptoms, the speed of notification is critical to the control of COVID.

In a decentralised system, self-diagnosis would present a security risk to the system, since it would be more open to abuse - anomalous patterns of self-reporting could not be detected and blocked in a timely manner and malicious attack could result in uncontrolled cascades of notifications

Given current turnaround times for testing, a decentralised system as proposed is predicted to contribute minimally to the control of COVID transmission.

3. Notifications based on the time since start of symptoms in the infected case.

People are not equally infectious throughout all stages of infection. People are most infectious near the start of onset of symptoms. Infectiousness typically starts about 2 to 3 days before and continues 2 to 3 days after (possibly longer in some individuals). In the centralised system, it is possible to rank contact events so that contacts that took place close to the onset of symptoms are more likely to result in notification. In the decentralised system, the notifications are based on the information about the person who is receiving the notification, not information about the person who has just been diagnosed. This means that unless this additional information is broadcast, in the decentralised system there is a greater risk of false positive and false negative notification.

Benefits of decentralised system

4. Support by owners of the operating systems

Google and Apple run the operating system that runs on most smartphones. Through their operating system and an API, Google and Apple will provide enhanced support for apps that use this API, and such apps must be decentralised. Decentralised apps will benefit from stability under upgrades of the operating system, and may benefit from improved support for bluetooth functionality. Use of Google and Apple systems may increase interoperability with other apps based on the same API. In the medium term, Google and Apple may include detection and recording of contacts in the operating system, such that an App would have access to a history of contacts even if only installed once symptoms of COVID initiate, which would potentially be a substantial benefit in terms of coverage.

5. Increased privacy

Increased privacy is an intrinsic benefit of the decentralised system, and may also have the additional benefit of increased uptake of the app. Increased uptake of the app will increase the efficacy of the intervention, provided the notifications can be sent in a timely manner (within 48

hours of symptom onset, ideally before), and provided the tracing algorithm is tuned to send notifications based on the correct measure of effective contact.

Summary of the trade-offs

6. Trade-offs in privacy, efficacy, and false positive notifications

In summary, the centralised system preserves information on contact events between individuals, one of whom is a diagnosed case of COVID. The information consists of pseudonymised IDs of both the case and their contact, as well as some information on the case: date of diagnosis, age band (to 10 years), time since onset of symptoms (in days), and first four letters of the postcode. The contact events themselves consist of signal strength for bluetooth.

For the decentralised system, similar data would be stored from cases, except for the details of the contact event. Analysis of the details of the contact event will be useful in stratifying risk of infection, and identifying infectious individuals before they transmit to others, with greater accuracy.

Both systems preserve much less information than manual contact tracing (where data is manually entered into a central database by a phone operator), though in manual contact tracing consent to record information about contacts is not obtained from both parties. Both systems can therefore be considered an improvement, in privacy terms, over established public health practices.

The loss of privacy associated with this extra information from the centralised system would be traded off for 1/ the capacity to send notifications quickly, which would reduce the risk of new infections and increase the chance of new infections 2/ the benefits of learning a better model of whom to send notifications to, reducing the number of false positive and false negative infection. Both of these capacities would increase the efficacy of the system in contributing to the control of the epidemic (saving lives), and to reducing the risk of people being falsely notified that they are at risk of being infected, and being asked to self-isolate or quarantine.

Both systems collect data, and distribute notifications that impose restrictions on civil liberties. Neither system is justifiable unless it’s benefit can be shown. Fatal flaws facing either system must be solved before release. For example, if decentralised contact tracing can only be initiated on a positive test, and test turnaround times cannot reach speeds necessary to trace the majority of people before they transmit to others.

Some unanswered questions that arise from this trade-off are
1/ whether any practical system exists that can be designed in a timely manner that combines advantages of both systems and has fewer disadvantages than either, or which of the existing systems can most quickly converge on this system;

2/ whether the benefits in terms of potential uptake for a decentralised system, and the operating system support offered by Google and Apple make up for the disadvantages of the decentralised system in terms of lack of control on false positives, false negatives, and delays introduced into the process of notification. For current testing turnaround times in the UK, these delays could make a difference between controlling the epidemic, or not, irrespective of uptake;
3/ conversely, whether the privacy concerns associated with the centralised model can be mitigated by appropriate oversight.

Detailed assessment

Receiving notifications that one has been in contact with someone infected with COVID is a disruptive event, it may cause concern, anxiety, and will lead to difficult and disruptive choices as to whether to self-isolate, work from home where possible, and reduce contact with family, friends, or engage in care of others. Similarly, not receiving notifications because a contact was judged low risk for transmission carries the potential for significant harm, since a person may infect others during the period when they could have been notified. The process of deciding who does and doesn't get notified given a history of contacts is central to contact tracing. Evaluating and improving this process is imperative to the correct function of contact tracing.

Concerns have been raised about the potential for digital contact tracing to invade the privacy of those who participate in contact tracing; for the process of contact tracing to be associated with coercive measures, e.g. compulsory use of the system, or coercive monitoring of individuals with the system; for digital contact tracing to leave a negative legacy of increased digital surveillance by states, especially through mission creep and misappropriation of identifiable data.

These are legitimate concerns. Two complementary approaches may be taken to mitigate these risks. One is to design systems that are less liable to being misused. The second is to engage a system of oversight, both by representatives of democratic institutions, by review and oversight boards, by compliance with data protection laws (generic and specific) and by choosing designs that minimise the risk of long-term harm from digital contact tracing.

Both centralised and decentralised digital-contact tracing in practice require both parties in a contact event to consent to it being recorded, which is not true of the central databases used to support manual contact tracing. A comparison with manual contact tracing is worthwhile. In manual contact tracing, individuals are interviewed by contact tracers after they become a case on the people they have contacted and the places they have been. These data are digitally centralised, including personal data on contacts who have not consented to be listed. Contact tracers then follow up on the information given, usually over several days, and pass on information to those at risk of having been infected. In sexual health medicine partner notification requires records to be kept of multiple linked-anonymised patient identifiers alongside extremely sensitive personal data. In comparison, digital contact tracing relies on the

informed consent of all parties involved, and even centralised versions of this system reduces the need to link patient identifiers with sensitive clinical and demographic data that routinely occurs in manual contact tracing. On the other hand, manual contact tracing will always be limited in scope due to its laboriousness, and it is easier for people to understand what the process is.

The issue of long-term legacy seems important too. Currently, there could be argued to be poor understanding of issues around privacy and consent on phone apps, and few people are aware of the collection, cleaning and re-sale of data that is widespread. There is some fear that digital contact tracing could lead to a continued deterioration of standards of privacy. Conversely, for many people, it may be the first time they explicitly consider the issue of data use at the population level, in which case it could be a welcome opportunity to have an overdue wider discussion of this important topic.

Done badly, digital contact tracing could set a precedent for increasing the use of digital technologies to reduce the privacy of individuals. Conversely it could be argued that exposing a wider proportion of the population to the issues of privacy, in the context of a public health intervention where data are pro-actively shared to help save lives, might facilitate a wider post-pandemic discussion of what is or isn't appropriate use of private data collected digitally, and help avoid the most egregious abuses of privacy that are already widespread.

The increased concerns around privacy of digital contact tracing compared to digitized manual contact tracing seem focussed predominantly on the ease with which relational data can be collected. Concerns about privacy need to be considered alongside the potential benefits of digital contract tracing and evaluated on grounds of proportionality. Specifically, digital contact tracing offers two benefits - (i) the intervention aims to save lives, and within an integrated public health system offers something that manual contact tracing cannot - and (ii) the intervention allows return of social freedoms , that blunter interventions, such as mass isolation (lockdown) have taken-away. Individual users, and society in general, will expect their data is used optimally in pursuit of these aims. In real terms this equates to accurate prediction of infection-risk, appropriate advice, and minimal amounts of disruption from inappropriate advice (eg, from false-positive and false-negative notifications).

The system should be as privacy-preserving as possible within these constraints.

Two broad classes of data architecture have been proposed for digital contact tracing, the decentralised model favoured by Google and Apple in their joint API, and the D3-PT consortium and being adopted by the public health systems of Switzerland, Austria and Germany and the centralised model favoured the public health systems of the United Kingdom, France, Norway and Australia. The purpose of this document to compare and contrast these different approaches to achieve maximum benefit, in terms of epidemic control with minimum disruption to society, in a privacy preserving manner

First, we address an issue of nomenclature: both these systems will operate ideally only when integrated within a public health response, such that messages can be reinforced by human operators, and such that the notifications can be reinforced and linked to a testing system. Furthermore, manual contact tracing will be needed to complement poor app uptake and address issues of digital exclusion. Both the systems will require many users to voluntarily upload data to a central database for the integration into a public health response. Both systems will need to be audited in terms of number of cases and average number of notifications per case to ensure correct functionality, and both systems will rely on parameters for the tracing algorithm that should be supplied by a central server so that the tracing policy can remain coherent with the local public health response. All of these points have been agreed for all the platforms under discussion.

Therefore the comparison is not one of one system having a central database in the public health system, and the other not, it is rather the question of which data are stored in the central database, and which data remain on the phones of users.

The principal difference between the two systems is that relational data, those that record parts of the contact graph, is proposed to remain on phones in the decentralised system, and is to be partly uploaded on the central server in the centralised system. In both cases, these relational graphs involve pseudonymized identifiers. In the decentralised system, no linked pairs of pseudonymized identifiers indicating a past contact event may be uploaded to the central database. Beyond that, both systems will rely on voluntary uploading of aggregated data about an individual's exposure to risky contacts for the correct functioning of the system and integration into the public health response.

Having discussed generalities, we now address a technical comparison of both systems in terms of how the notification procedure works, and how it may be evaluated and improved in both settings.

The setting

Consider a population of N individuals, labelled i=1,...,N, each equipped with a phone running a digital contact tracing app. Let $X_{i,t}$ be information about individuals that they can choose to share with a central server when requested. Examples include gender, partial postcode, nature and severity of symptoms, a score indicating relative vulnerability, test results, etc.

Let $G_{ij,t}$ be the history of contacts between individuals i and j up to time t. The history may be deleted with a rolling average. The history may be obfuscated using pseudonymized IDs.

Let $N_{ij,t}$ be the history of notifications received by individual i up to time t relating to prior contact with person j. The history may be obfuscated by using pseudonymized IDs.

Under all systems, the phone of user i stores the full list {i, $X_{i,t}$, $G_{ij,t}$, $N_{ij,t}$ for all j}. The list will be pseudonymised with respect to j, and j may have provided ephemeral pseudonymised IDs such that, from i's point of view, it will not be possible to see whether contacts are with the same or different people.

Identifier-stripped contact lists

Let $G_{ix,t}$ and $N_{ix,t}$ be the versions of $G_{ij,t}$ and $N_{ij,t}$ where the contact identifiers j have been removed and replaced with unlinked labels.

The system can only be called truly anonymised if it is not possible to reconstruct $G_{ij,t}$ and $N_{ij,t}$ from $G_{ix,t}$ and $N_{ix,t}$, which in general will not be the case.

Contact tracing

Centralised system

Under the centralised system, when the individual i is diagnosed with COVID at time t, they are asked to upload their history {i, $X_{i,t}$, $G_{ij,t}$ for all j} to the central server. At a minimum, the server contains the list of entries uploaded from diagnosed cases. An algorithm f_k regularly scans the database, with a functional operation we denote f_k({i, $X_{i,t}$, $G_{ij,t}$ for all j}) to generate a set of notifications at time t. The index k denotes that the chosen algorithm is one of many (probably infinite) set of possible algorithms. Running the algorithm generates a new set of notifications $N_{ij,t+1}$ that can be pushed or pulled to the relevant users, and are only stored on the central server.

Decentralised system

Under the decentralised system, when the individual i is diagnosed with COVID at time t, they are asked to broadcast their list of pseudonymised versions of their ID i to all other users, via a central server, for a specified amount of time. Let $D_{j,t}$ be their history of diagnosis that is stored on all other phones. Based on this broadcast, all other users i run an algorithm g_k({j, $D_{j,t}$, $X_{j,t}$, $G_{ji,t}$, $N_{ji,t}$ for all j}) to update their own notification $N_{ji,t+1}$, where g_k is one of many possible algorithms, which can be downloaded and updated from a central server at regular intervals.

Crude comparison with regards to privacy and effectiveness

The decentralised system does not rely on uploading any individual data to a central server, and is tuneable, in the sense that a central server can supply a function g_k.

In its basic form, the decentralised system, while maximally privacy-preserving, does not supply any information that would allow a public health system to know how many cases arose, how

many notifications were received, or to allow any assessment of the suitability of the tracing algorithm g_k, in terms of its two main functional benefits namely numbers of infections prevented and lives saved, and individuals freed from unnecessary isolation or lockdown.

The basic security and privacy assessments of these systems should be made by experts, which we are not. Centralised and decentralised systems each have different vulnerabilities to security breaches, explored elsewhere. Those are critically important considerations, but not the subject of this report.

An asymmetric epidemiological difference between the systems

The core algorithmic steps for the centralised system, $f_k(\{i, X_{i,t}, G_{ij,t} \text{ for all } j\})$ and decentralised system, $g_k(\{j, X_{j,t}, G_{ji,t}, N_{ji,t} \text{ for all } j\})$, are not equivalent, since the former uses covariates $X_{i,t}$ of the person who has been diagnosed with COVID, whereas the latter uses covariates $X_{j,t}$ of the person who has come into contact with the person with COVID.

(Use of the covariates of the contacts is not done in the centralised system, because only index cases upload contact data, not all individuals or notified individuals. This could be changed by asking notified individuals to donate data, but is not planned by any system as it would rapidly build a social graph without clear utility.)

So for example supposing Alice has been in contact with Bob 5 days ago, and Bob is now diagnosed with COVID. For the centralised system, the tracing algorithm can account for the age group, gender, time since symptoms, etc. of Bob, whereas for the decentralised system, the tracing algorithm can account for the age group, gender etc. of Alice. Which approach is better depends on which set of covariates is better for predicting whether transmission is likely to have happened or not. Currently, the three best established predictors of transmission are 1/ time before or since onset of symptoms of the index case, 2/ severity of symptoms of the index case and 3/ age of the contact. This gives a small accuracy advantage to the centralised system, since two of the three predictor depend on the transmitter, and the effect sizes are larger.

Data donation

Centralised system

Under the centralised system, individual i may choose to donate their history $\{i, X_{i,t}, G_{ij,tt} \text{ for all } j\}$ to the central server for additional assessments. It would also be possible to augment the variables $X_{i,t}$, though this is not discussed here.

Decentralised system

Under some configurations of the decentralised system, individual i could choose to donate their history {i, D_t , $X_{i,t}$, $G_{ix,\,t}$, $N_{ix,t}$ for all j} to the central server for additional assessments. It would also be possible to augment the variables $X_{i,t}$, though this is not discussed here. Most implementations would discourage this, or reduce it to aggregate statistics. (Individual specifications will vary, for example some systems may not allow time stamped histories $G_{ix,\,t}$ to be uploaded.)

Comparison with regards to privacy

The decentralised system is only privacy enhancing if the lists $G_{ij,\,t}$, $N_{ij,t}$ cannot be reconstructed from $G_{ix,\,t}$, $N_{ix,t}$. Critically, under the centralised system consent to donate data to a centralised system is obtained for all users of the system at enrollment (on-boarding) and no additional consent needs to be sought to ensure (and optimise) effectiveness and safety of the system. Conversely, if safety and effectiveness of a decentralised system is reliant upon a subset of individuals choosing to concede less privacy in order to benefit the system overall, this raises ethical questions of fairness.

Evaluation

General concept

Each time-step, each system issues a new set of notifications based on the tracing functions f_k and g_k , respectively, that process the contact history G_{ij} . These notifications are predictions that, if person i is diagnosed at time t, that person j is now likely to be infected (and so infectious), or not. A perfect system will issue notifications only to those infected, and not to those not infected. No system will be perfect, due to the stochasticity of virus transmission and imprecision in the risk calculation: Some individuals will be notified even though they were not infected, thereby disrupting their lives (false positives), and other individuals will not be notified even though they were infected, so missing an opportunity to inform them of the risk they pose to others (false negatives). A system that minimises false negatives can contain the epidemic if widely used, but may result in so many notifications that it becomes discredited (‘boy who cried wolf’) and may be less equitable and equally costly than rolling lockdowns (Hinch et al, Report 3). The credibility and utility of a tracing system therefore rests on being able to evaluate false positive rates and false negative rates (Table 1), that should be made public for transparency.

This issue is separate from other important evaluations which affect the outcome of the digital tracing policy, such as uptake of the app, user adherence, integration with testing and other public health measures. It is linked to engineering performance, conditional on any signal having been exchanged at all.

Centralised system

A possible schema for evaluating the tracing function f_k goes as follows. For each individual i declared a case at time t, for each subsequent time t+1, t+2, t+3,..., consider how many of the individuals j contacted by case i (non null entries in $G_{ij,s}$ s<=t) themselves become a case. Separate these into those that were notified and those that were not. Model the follow up process to correct for background infection and notification rates. Output is an estimate of the four entries of table 1 for the tracing function f_k at time t.

Decentralised system

Two possible schemas for evaluating the tracing function g_k , involving data donation, go as follows.

Schema 1. For each notified individual j notified of being a contact at time t, ask them to donate their history {i, $D_{x,t}$,$X_{i,t}$, $G_{ix,\ t}$, $N_{ix,t}$ for all x}. At a later date t+m, ask them whether they have developed symptoms (or perhaps more simply ask them to upload their own notification history $N_{jx,t}$), and to donate this information. With this information it is possible to estimate only the top row of table 1.

Schema 2. For a sample of individuals j ask them to donate their history {i, $D_{i,t}$,$X_{i,t}$, $|G_{ix,\ t}|$, $|N_{ix,t}|$ for all x}, and to donate information on whether they have experienced symptoms in the last m days, and whether they had been notified or not at the time of sampling. With this information, it may be possible to evaluate whether exposures to each diagnosed individual i led to a correct notification, with modelling to evaluate the case of multiple exposures, this leading to estimates of the whole of table 1. At this point, no clear schema for evaluation has been developed for evaluation of decentralised systems, and details of what is possible will vary depending on specific choices. In some variants of the decentralised system proposed, only the size of sets $G_{ix,\ t}$ and $N_{ix,,t}$ and not the linkage between them, which is needed to evaluate the tracing function.

Iterative improvement

By simulation, it is possible to find the desired sensitivity and specificity that maintains epidemic control (contributing to keeping R<1, maintaining the right to life) and that minimises the number of false positives (maintaining the right of movement). Note that in extremis, curtailing the right to movement may cost lives, such that trade-offs may need to be considered.

At this stage, it is possible to computationally test arbitrarily many functions to find the best next function f_{k+1} or g_{k+1} . Both centralised and decentralised systems will allow setting-specific and changing functions over time.

Rapid convergence will ensure closest adherence to values that keep R<1 whilst minimising false positives. Convergence is possible for the decentralised schema 1, but will be slower than the centralised schema or the decentralised schema 2.

Summary assessment

With the pure decentralised model, no evaluation or optimisation is possible using data internal to the system. Evaluation would involve detailed study of many users of the app, and construction of a new centralised database of app users, with extensive epidemiological questionnaires.

Unless other measures are in place to keep R<1, this architecture can be said to have placed the right to privacy above the right to life and the right to movement. If other systems are in place to keep R<1, the system is not needed, and privacy would be best preserved by not engaging in digital contact tracing. The system could be justified in marginal cases, where R>1, and some external metrics are kept to see that the number of notifications is e.g. comparable to manual contact tracing, and an external database is constructed e.g. by phone interviews, to see that the false positive rate is acceptable. For comparison, the false positive rate for contact tracing is approximately 85% to 95% of those notified.

With a centralised model, and with the decentralised model with large data donation (schema 2), it is likely to be able to rapidly optimise tracing functions f_k and g_k respectively. Both these schema involve central databases, the difference being that in one case the contact histories G_{ij} is pseudonymized, whilst in the other the contact histories are G_{ix} are anonymised with respect to one of the parties. Given that the point of decentralised systems is to enhance privacy, extensive data donation is likely not to be an allowed feature of such a system, and to date, no schemes have been developed for the evaluation and improvement of the function g_k.

Of the three schemas presented, the pure decentralised system clearly places higher value on the right to privacy than the right to life or the right to movement, which may seem acceptable only in the implicit but untested assumption that other privacy-preserving interventions exist to control the epidemic.

The decentralised model with data donation from only those individuals who are notified (schema 1) both constructs a central database and has poor convergence on good choices of function g_k , and so seems to have none of the benefits of any system.

Of the centralised system or the decentralised system with large data donation (schema 2), the centralised system has network information on fewer people, and less personal data stored on fewer phones. The benefits of anonymisation over pseudonymisation may be overstated when detailed contact histories are collected. Furthermore the centralised system allows tracing to proceed based on information from the source index case (time since symptom onset, severity of symptoms) which may plausibly predict infection risk better than using information from the recipient contact.

Precedent for the post-pandemic world

It is said that in a crisis, new norms are established which are hard to change. Democratic oversight of data use has functioned poorly in the recent past. There is a fear that surveillance to improve a pandemic response could lead to increased surveillance and coercive practices for future generations, and that increased data use for pandemic control could lead to increased data use for negative purposes. We propose that establishing the principle that governance should focus on a transparent process of balancing the right to privacy, the right to movement, and the right to life would be a step forwards in data governance.

The converse proposal, of prioritising the right to privacy over the right to life and the right to movement, only legitimises and reinforces the notion that digital solutions are not to be trusted to democratic oversight, and will be the exclusive purview of despots and autocrats, to all our detriment.

Conclusion

Digital contact tracing can contribute to the suppression of COVID-19 (maintaining R<1) whilst enabling more freedom of movement and economic activity than a lockdown. This is not a trivial task. The basic reproduction number of COVID is close to 3, and about half of transmissions occur before someone is symptomatic.

Success will depend on high uptake, trust in the beneficence of the system, that use of data and temporary loss of privacy has been commensurate with the task. It will also depend on engagement with the notifications issued by the app, and a widespread understanding that every effort has been taken to ensure that the system can realistically contribute to reducing transmission, and that the system that issues notifications is undergoing evaluation and improvement. Our aim here was to lay out some of the trade-offs in terms of these latter aims. Decentralised systems are a priori more privacy-preserving, and are currently supported by the operating system owners, which may enable larger uptake. Centralised systems may have a larger privacy cost, but for equal uptake, offer substantial benefits in terms of their intended potential public health benefits. Appropriate oversight of centralised systems may mitigate privacy risks.

Decisions taken at this stage will benefit from a clear understanding of the trade-offs between the three aims: preserving privacy, reducing infections, and minimising the number of people required to isolate.

The trade-offs could be reduced if a system emerges that combines benefits of each option and reduces drawbacks. At the present time, with sufficient oversight to ensure privacy is maintained in the centralised system, and if this oversight is transparent enough to encourage uptake, the centralised option will give more options to suppress COVID epidemic spread.

References

[Ferretti & Wymant et al. 2020] Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, David G., Fraser, C. (2020) Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, *Science* 2020.

[Morley et al. 2020] Morley, Jessica and Cowls, Josh and Taddeo, Mariarosaria and Floridi, Luciano, Ethical Guidelines for SARS-CoV-2 Digital Tracking and Tracing Systems (April 22, 2020). Available at SSRN: https://ssrn.com/abstract=3582550 or http://dx.doi.org/10.2139/ssrn.3582550

[OpenABM-Covid19] Available at https://github.com/BDI-pathogens/OpenABM-Covid19

[Parker et al. 2020] Parker MJ, Fraser C, Abeler-Dörner L, Bonsall, D, Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic, *Journal of Medical Ethics* Published Online First: 04 May 2020. doi: 10.1136/medethics-2020-106314

[Vaudenay. 2020] Vaudenay S, Analysis of DP-3T (April 8, 2020). Available at IACR: https://eprint.iacr.org/2020/399

[Abeler et al. 2020] Abeler J, Altmann S, Milsom L, Toussaert S, Zillessen H, Support in the UK for app-based contact tracing of COVID-19 (March 26, 2020). Available at OSF: https://osf.io/3k57r