Decentralized contact tracing protocols and a risk analysis approach to pandemic control

by

James Ian Mackie Petrie

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Po-Shen Loh Professor, Dept. of Mathematical Sciences, Carnegie Mellon University
Supervisor(s):	Stephen Vavasis Professor, Dept. of Combinatorics and Optimization, University of Waterloo
Internal Member:	Chris Bauch Professor, Dept. of Applied Mathematics, University of Waterloo
Internal-External Member:	Plinio Morita Professor, School of Public Health Sciences, University of Waterloo
Other Member(s):	Joanna Masel Professor, Dept. of Ecology and Evolutionary Biology, University of Arizona

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Section 2.2 contains excerpts from a white paper by Covid Watch (a non-profit that I co-founded) which was later published by IEEE Data Engineering Bulletin [1]. My contributions to this document include design of the decentralized architecture for exposure notification, numerical modeling of expected impact, framing of NPIs as a filtering problem, and writing the first draft of the document.

Chapter 3 and Appendix A (published in the Journal of Royal Society Interface as "The economic value of quarantine is higher at lower case prevalence, with quarantine justified at lower risk of infection" [2]) was a joint work with Prof. Masel. The breakdown of contributions as included in the article are: "J.P. conceived of the mathematical approach, carried out the mathematical analysis, created visualizations and supported writing. J.M. conceived of the method of presentation and example applications, supervised and administered the project and led writing."

In Section 4.2 and Appendix B, the manuscript "Realistic consideration of quarantine compliance suggests earlier PCR testing and shorter quarantine following SARS-CoV-2 exposure" was co-authored by myself, Anel Nurtay, Luca Ferretti, Christophe Fraser, and Joanna Masel. My contributions include the math and numerical implementation for policy optimization, proof that the optimal policy only depends on quarantine and isolation effectiveness via the ratio of these values, generation of figures, and providing a support role in writing.

Section 4.3 uses a disease model created in collaboration with Joanna Masel. The probability calculations, optimization, and a majority of the writing in the section are my work.

All other material in this thesis was written by me.

Abstract

Non-pharmaceutical interventions (NPIs) can protect against pandemic pathogens, but they depend on behaviour change, and so can impose costs on quality of life and civil liberties. With careful system design and risk analysis these tradeoffs can be improved, enabling more effective disease control at a lower cost. In this thesis, I propose a method for decentralized digital contact tracing that is fast, scalable, and cannot be used for mass surveillance. I show how targeted quarantine substitutes for broad social distancing and use this relationship to estimate the optimal quarantine risk threshold – finding that it strongly depends on disease prevalence. Using the joint distribution for infectiousness and test sensitivity, quarantine duration and test timing can be chosen to minimize the duration of quarantine without increasing expected transmissions. Decentralized digital contact notification apps were used by close to 100 million people during the COVID-19 pandemic and prevented a significant number of transmissions despite challenges with system robustness. Decentralized digital contact tracing combined with adaptive risk analysis can efficiently suppress infectious disease in the idealized (high participation) case, however more work is needed to design solutions that are both robust and socially acceptable.

Acknowledgements

I'd like to thank my supervisor, Stephen Vavasis, for his steady mentorship and many interesting conversations throughout my graduate degrees despite my frequent changes in research direction. I'd like to thank Joanna Masel for her dedication to improving the COVID-19 response, many interesting conversations, and for pushing me to improve. Thanks to the other members of my graduate committee: Plinio Morita, Chris Bauch and Po-Shen Loh for their useful feedback on this thesis and helpful advice on digital contact tracing and epidemiology research.

Thanks to Katerina Papoulia, Reza Hirmand, and Tim Thompson for all of their help introducing me to numerical fracture modeling and for many interesting discussions.

I'd also like to thank Chris Bakerlee for valuable advice on how to make my research more effective, and all of the people I've met through ELBI who've introduced me to their work in biosecurity.

I'd like to express my sincere appreciation for all of the people who worked hard to try to build our way out of the Covid-19 pandemic, especially those who I've been lucky enough to work closely with these past few years:

Tina White for taking the initiative to organise Covid Watch, and having the determination to push everyone (including herself) to produce the results we needed.

Rhys Fenwick, for taking responsibility for so many different things and having such an uplifting attitude.

Zsombor Szabo for bringing the expertise and dedication to build a high quality solution.

Scott Leibrand and Dana Lewis for their collaboration, great discussions, and their dedication to having a positive impact.

Peter Eckersley for adding a realistic perspective on contact tracing outcomes and connecting the digital contact tracing community.

Christophe Fraser's lab, with special thanks to Luca Ferretti for all of the helpful advice.

Jolene Elizabeth, Jeff Schwaber and others who stepped up to manage the many difficult parts of the non profit.

Manu Eder for contributing many good ideas on technology strategy, and his solid analysis of potential protocols.

Jack Gallagher for needed criticisms and discussion of important implementation details for the TCN protocol.

Henry de Valence for formalizing and writing much of TCN protocol description.

Ian Miers for interesting discussions on the TCN protocol and possible extensions to it.

The proximity team at Covid Watch; Mike Wirth, Dar Scott, Ramsey Faragher, Mark Ingle and others for their hard work on researching proximity estimation for digital contact tracing.

Akhil Veeraghanta for dedicating his time before final exams to build the first Android prototype of the TCN protocol .

Michelle Xie, Ian Brown, Rich Dandliker, Jesse Colligan, Tessa Alexanian, Isaiah Becker-Mayer, Sydney Von Arx, Cherry Wu and many others for their hard work with Covid Watch.

Jenny Wanger, Harper Reed, Andreas Gebhard and many others for bringing the TCN Coalition together.

Dan Kohn for his persistence in building LFPH and hard work supporting open source software for public health.

Amanda Wilson for her hard work with our collaborators at the University of Arizona.

Matthew McCaskey for great discussions on using digital contact tracing for influenza.

Martin Walsh and Denis Pitcher for their great questions and hard work bringing the exposure notification app to Bermuda, and also for being such welcoming hosts during our visit.

All of the Covid Watch advisors for donating their expertise in a broad range of fields.

Emergent Ventures and our other funders for the resources to support much of the work done by Covid Watch.

Travis Marsh and Rebecca Young for all of their help with Covid Watch, and also Rebecca for her encouragement and career advice.

Sameer Halai, Victoria Skelly, Pavlo Apisov, Devin McIntire, and everyone else from the WeHealth team for their focus on bringing an effective exposure notification system to end users with the features that they needed.

Sam Hoshyar for his help investigating UWB for digital contact tracing and being a supportive friend.

I would also like to thank the many others that aren't mentioned here who did their best to build solutions to reduce the negative impact of COVID-19.

Finally, I would like to thank my partner Oraya Srimokla, my friends, and my family for their continuous support and encouragement.

Table of Contents

List of Figures

Chapter 1

Introduction

Infectious diseases impose an enormous burden on human health globally, and the appearance of pandemics like COVID-19 have the potential for rapid devastation [5]. When vaccines or treatments are unavailable (as is typical in the early stages of a pandemic) Non-Pharmaceutical Interventions (NPIs) are the only option for disease suppression [6]. Fortunately, pandemics begin with a prolonged early period in which disease spread is concentrated to a very small fraction of the population[1]. So, in theory it is possible to contain a disease at a very low average cost per person by isolating infectious individuals [7]. As an example, contact tracing is an NPI that aims to quarantine recently infected contacts of known cases before they infect others. The number of potential contacts is proportional to the number of infected and not the total population, so contact tracing can be a highly efficient intervention when disease prevalence is low [8].

The focus of this thesis is on the development of technologies and strategies that can lower the cost (both in terms of dollars and quality of life) of disease containment by making use of heterogeneity in risk of infectiousness. We investigate several ways to do this, including collecting better information, improving risk analysis and making better decisions about risk on the basis of broader control dynamics. For example, the cost of quarantine can be reduced by using information from testing to shorten quarantine duration without increasing the expected number of transmissions. The decision about the tradeoff between health and disruption to daily life is a complicated ethical and political problem. Our aim here is to focus on increasing capabilities to enable better tradeoffs (essentially expanding the Pareto frontier [9, 10, 11]). Conversely, we aim to limit the enhancement

[1] Even after the number of cases has increased, broad social distancing could be used to reduce case numbers to a manageable level.

of capabilities for non health-related purposes such as mass surveillance. The intention is for tools we develop to be used broadly, so careful design is needed to ensure people are comfortable with how these systems can affect privacy and the role of government.

This work began in January 2020 with the design of digital contact tracing architectures in response to the COVID-19 pandemic. Chapter 2 provides an overview of my work with Covid Watch and the TCN Coalition on the development of the first decentralized digital contact tracing app. Such apps are capable of notifying people of exposure to an infectious individual without gathering information on the vast majority of users, and at the time of this writing (July 2022) have been used by over 90 million people worldwide through Apple and Google's Exposure Notification system.

For digital contact tracing, manual contact tracing, and a broad variety of NPIs, the decision about who to target for behaviour modification is crucial. The direct cost of an additional infection is relatively constant, but indirect costs strongly depend on the state of disease control in the broader population. For this reason, the risk of infectiousness that would make quarantine worthwhile varies depending on disease prevalence. Chapter 3 summarizes our work using a simple epidemiological model to estimate the 'optimal' quarantine risk threshold in a mostly immune-naive population that is actively social distancing. We find that when disease prevalence is low, preventing an additional infection can substantially reduce the burden of social distancing on the broader population. Compared to the implicit risk threshold of common policies, the optimal quarantine risk threshold is much more dynamic and could be several orders of magnitude lower.

While quarantine of potentially infectious individuals can provide enormous social benefit, it is also very costly to the individuals whose lives are disrupted. Besides supporting them through quarantine (logistically and financially), we can also minimize the duration of quarantine using well-timed PCR tests. Chapter 4 summarizes our work using the joint distribution for test sensitivity and expected daily transmissions (conditional on infection date and symptom onset date) to optimize test timing and quarantine duration. We find that the optimal policy doesn't depend on isolation effectiveness or quarantine effectiveness individually, but only on their ratio. For a reasonable value of this ratio (0.7), 6 days of quarantine with a test 5 days after exposure results in the same number of expected transmissions as the default policy (10 day quarantine with no test) with 4 days less quarantine.

The work on quarantine risk thresholds in Chapter 3 is for a special case – making decisions on the margin to offset social distancing. In Section 5.3, we outline how this approach could be expanded to choose better trajectories by borrowing concepts from Optimal Control [12]. In Sections 5.1 and 5.2 we explore how better classification algorithms

and better information increase the risk 'surplus' generated, which essentially means that the same number of transmissions could be prevented with fewer quarantines. In Section 5.4 we describe how to complete the definition of the overall policy optimization problem by assuming that a relationship between cost of border stringency and expected case importations can be estimated. This framing highlights the importance of reducing the expected cost of elimination per introduction and motivates the development of more efficient tools for disease suppression. Section 5.6 describes a more powerful (but still decentralized) digital contact tracing system that can perform backward/2nd degree and venue-based contact tracing as well as integrate with the manual contact tracing system.

There have been substantial real-world challenges with deploying targeted NPIs against COVID-19 – most significantly insufficient participation and fragility of processes to multiple points of failure. Section 6.2 investigates observed points of failure in the Exposure Notification (EN) system and Section 6.4 attempts to assess whether alternative protocols could improve the success rate. Finally, in Section 6.6 we discuss strategies for building NPIs that are better supported and more robust.

1.1 Background

1.1.1 Broad NPIs

Many modelling studies have demonstrated that NPIs can be a powerful tool against infectious disease [6, 13, 14, 15, 16]. The mechanism of action is simple - behaviour change is used to reduce either the frequency or transmission risk of interactions between infectious and susceptible people, thus reducing the rate of onward transmission. Gathering evidence of efficacy in the real world has been more challenging for many reasons (including ethics of study design, limited number of natural experiments, process and measurement noise, and the time-limited effect of interventions). Prior to Covid-19, there were a limited number of real-world studies examining the effect of NPIs. Hatchet et al. and Markel et al. showed that during the 1918 influenza pandemic, broad social distancing measures (such as closure of schools, churches, and theaters) temporarily reduced the number of transmissions [17, 18]. A major challenge with broad social distancing is that it can be enormously expensive (significantly impacting the daily lives of a large fraction of the population) and so it is difficult to maintain for a long time.

1.1.2 Targeted NPIs (Case Isolation and Contact Tracing)

Targeted interventions such as case detection and isolation as well as contact tracing function by finding people who are at (much) higher risk of being infectious than the general population and reducing their interactions with susceptible people. These targeted interventions are much more efficient than broad social distancing when disease prevalence is low [8]. The efficacy of contact tracing depends on several factors, including transmissibility of the virus, proportion of presymptomatic transmissions, delays in the tracing process, probability of case detection and effectiveness of isolation/quarantine [13, 19]. Contact tracing was used extensively during past epidemics (e.g. SARS-1 and Ebola), and was found to be effective but heavily dependent on public support [7, 20] .

1.1.3 Digital Contact Tracing

Prior to the Covid-19 pandemic there were several proposals to automate parts of the contact tracing process using digital devices. Salathe et al. used wireless communication between wearable devices [21] and Yoneki used Bluetooth handshakes and GPS on smartphones [22] to log proximity events between study participants. These studies enabled construction of high resolution contact graphs for more realistic simulation of disease spread. Farrahi et al. proposed the use of smartphone Bluetooth proximity detection for digital contact tracing using a system where all data is shared with a central server [23]. Nguyen et al. developed an app to log WLAN (Wi-Fi router) signals and detect proximity between users based on the similarity of their signal history [24]. Prasad et al. proposed a presence tracing system using temporary keys broadcast by stationary beacons [25]. While this approach allows for privacy-preserving notification, it depends on physical infrastructure that would be challenging to massively scale in an emergency. Another challenge is that it requires a large amount of data (around 30Mb) to be downloaded for each infected user, which would be prohibitive if thousands or millions of people were infected at the same time[2]. Approaches proposed prior to the Covid-19 pandemic were either not privacy preserving or would have been difficult to rapidly deploy, but they demonstrated the promise of accurate, rapid and scalable digital contact tracing.

[2]Despite these scaling challenges, the analysis of design objectives, threat model, and user preferences in [25] and [26] would have been very useful early in the Covid-19 pandemic, but unfortunately I wasn't aware of this work at the time.

Chapter 2

Decentralized Digital Contact Notification

2.1 Background

The following section describes some of my work on decentralized contact notification with Covid Watch and the TCN Coalition. Many design decisions were made to reduce time to deployment in early 2020 at the start of the COVID-19 pandemic. This section discusses the development process of the technology with the goal of providing context for why design decisions were made. See Section 5.6 for a more advanced version of decentralized digital contact tracing that I would advocate for were it not for lock-in to Apple and Google's Exposure Notification system (and Appendix D for a rough proposal for more robust disease control).

In January 2020, news from China warned of potential for COVID-19 to have a devastating impact on global health. China's extensive use of contact tracing and the success of contact tracing against other diseases suggested it could be an important tool against COVID-19[1]. However, the challenge to mobilize a large number of contact tracers in time in other countries seemed prohibitive, especially if the number of cases increased rapidly. Additionally, traditional contact tracing can be error-prone because it relies on people remembering who they have been near and being able to contact them[2]. With the recent

[1] And by the end of February 2020 more detailed modelling studies emerged predicting that it could be an effective containment strategy [27]

[2] Also, the speed of manual contact tracing is limited by several factors, which it turned out is especially

adoption of smartphones by a majority of people in some countries, we had an opportunity to make use of technology that was not available in previous pandemics.

It was clear that many of the containment strategies used in China would not be easily accepted in many other countries, especially anything perceived to have dual-use capabilities for traditional surveillance. The aim of my initial work was to develop a more scalable version of contact tracing that was politically acceptable in most countries. The first design considered was one where smartphones participating in the system would wirelessly communicate with other nearby devices to detect potential exposure events. In the simplest version of this system, the log of these contact events would be sent to a central server where the entire contact graph between app users could be assembled. Using this approach would allow people with access to the central server to monitor an almost real-time interaction graph between individuals participating in the system. If one of these individuals tested positive for COVID-19, their interaction history could be checked to find contacts that meet a sufficient risk condition for contact tracing (using a standard subset operation on this dataset). While this design is probably the simplest option, the potential for the system to be misused comes with a substantial acceptability cost in many countries.

The second design considered overcame most of the potential for system misuse at the cost of a moderate increase in complexity. Similar to the first design, potential contact events are logged when pairs of devices are within range of each other. The crucial difference is that logs of these contacts are kept on device, and whether a user has been exposed to infectious individuals is also checked on device. This can work by having each device wirelessly broadcast a random or pseudo-random temporary value (which we later chose to call a Temporary Contact Number, or TCN for short), which other devices detect and record. People who test positive can upload their TCNs from previous days to a publicly accessible server. All devices in the system can periodically download all TCNs from this 'infected' server and compare them to TCNs they had recently logged. If there were any matches, the app could alert the user of their potential exposure and make recommendations based on the assessed risk of infection.

The second (more decentralized) approach has a few key benefits: 1) no usable information ever leaves the device of a vast majority of users, which prevents use of the system for mass surveillance by either governments or third parties. 2) even the information shared by people who are infected is highly anonymous- just which random values they had broadcast previously (which only mean anything to people with logs from the exact places and times they were there). 3) the design is relatively simple to explain and could be implemented in

important for a fast-moving virus like COVID-19 [28] where delays of a day or two can allow for many more onward transmissions

a few weeks - no elaborate cryptography needed. 4) It is possible to use as a starting point for more sophisticated designs including second degree tracing, estimation of epidemiological parameters, or integration with manual contact tracing. If implemented properly such an app would be one of the most private apps on peoples' phones, while accomplishing the needed task of exposure notification.

For several years AdTech companies have been recording smartphone Wi-Fi broadcasts for location tracking purposes, which eventually caused smartphone companies to modify the OS to periodically randomize the MAC addresses emitted by smartphones [29]. This randomization prevented trackers from linking MAC addresses recorded at different times and locations to assemble the location history of individuals. I initially wanted to build the contact tracing system with these randomized MAC addresses acting in place of the previously described TCNs. Most smartphones broadcast 'Wi-Fi probe requests' with this unique MAC address 55-2000 times per hour [30], so if the contact tracing app could record them this would be frequent enough to detect most high-risk exposures. Such an implementation is a special case of the previously mentioned decentralized protocol with the key benefit that nearly every smartphone is already broadcasting the necessary information. Primary cases would still need to choose to upload data at the time of testing positive, but this would not depend on having previously downloaded an app. This would mean that the number of traceable transmissions could scale linearly with adoption of the app, rather than quadratically (since only the person receiving a notification would need the new software). Unfortunately, after speaking to smartphone hardware specialists in late January 2020, it became clear that this type of Wi-Fi logging is not possible to do with a standard app- it would require operating system level permissions.

It seemed like the best way to get this system built was to convince either Apple, Google or governments to implement it. In early February I began writing a whitepaper to explain the privacy model and present results of an epidemiological simulation of digital contact tracing applied on a real-world network [31]. Three weeks later I saw a forum post by Tina White advocating for the use of smartphones to build a COVID-19 risk heatmap [32]. I reached out and we decided to collaborate on tech development for COVID-19 mitigation, and the collaboration soon developed into the Covid Watch non-profit. In early March we talked to the founders of a group named CoEpi [33] who were starting to build a contact tracing app with a focus on grassroots adoption and symptom reporting. They were investigating Bluetooth for device-device communication, which seemed to have fewer restrictions than Wi-Fi. We decided to collaborate on several parts of the project including an implementation of the previously described decentralized protocol (CEN protocol, later renamed TCN protocol). Two weeks later we had built a proof-of-concept app and changed our focus back to advocating for this technology by writing the Covid Watch white paper

(posted March 20, 2020 and published by IEEE Data Engineering Bulletin in June 2020 [34]).

2.2 Excerpts from Covid Watch White Paper on Decentralized Digital Contact Notification (March 20, 2020)

Existing mobile apps without a privacy focus have been an effective intervention to reduce the spread of COVID-19. However, invasive interventions carry significant human rights costs, including the temporary loss of personal freedom and fears around whether that freedom will be restored. A mobile app with a strong privacy model may also have greater efficacy because people will be more likely to share accurate data if they know that data is safe. Ensuring privacy prevents COVID-19 patients from being ostracized or socially harmed on account of inadvertent potential data exposures. Also, mobile apps with poor privacy models may further undermine public confidence in responses and exacerbate existing mistrust. In contrast, our mobile app research has focused on developing a strong privacy model while still providing effective intervention. Using the app, users can alert recent contacts without anyone being able to trace the information back to them. We believe this intervention has the potential to slow or stop the spread of COVID-19 and save lives.

2.2.1 Current Mobile Phone Interventions

South Korea and China have demonstrated two successful systems for containing COVID-19 that make extensive use of technology. The results they have seen match well with predictions from numerical models: with a sufficient diagnosis rate and contact tracing accuracy COVID-19 can be contained. China was the first to create a mobile app intervention [35]. Their app uses GPS history and other data to assign a risk score. This score is then used to control which individuals are allowed to move freely. China's intervention appears to have been successful, but required far-reaching state surveillance that, by the standards of most liberal democracies, would be considered highly invasive, likely unlawful, and politically unpalatable. South Korea publicizes a large amount of information collected from the cellphones of infected patients so that others can determine if they had been in contact [36]. South Korea's success has been attributed mostly to (1) widespread testing (2) contact tracing and (3) case isolation. However, their mobile alert solutions do

not effectively anonymize patient data. They gather location data from interviews, mobile phone GPS history, surveillance cameras, and credit card records then send text alerts with the location history of patients. Much like the intervention in China, this appears to be effective, but takes a similarly high toll on personal privacy.

We've built a privacy-preserving version of these successful interventions that we believe would have a regulatorily, publicly, and politically viable adoption process within the United States and other Western countries. The system as designed complies with existing regulations around medical information in the United States and does not reveal identifying patient information. Making Interventions More Efficient

2.2.2 Making Interventions More Efficient

Non-pharmaceutical pandemic interventions fundamentally make a trade-off between two important social goods: (1) loss of life from the pandemic and (2) economic impact, which influences health and well-being outcomes indirectly. Mobile app interventions are a powerful public health tool because they can improve this trade-off. In general, non-pharmaceutical approaches to infectious disease control have the following components:

- Filtering (picking a subset of the population)
- Intervention (modifying the behaviour of these people)

For example, quarantining patients with a positive diagnosis applies a filter based on testing and then applies the quarantine intervention. Other examples include travel restrictions for at-risk areas, cancelling public events in a specific city, or encouraging more handwashing in an entire country. Some of these interventions, especially self-isolation [37], are highly effective at preventing the transmission of infectious diseases like COVID-19. The downside is that they can also be costly to use.

The quality of filtering plays a crucial role in determining the trade-off between loss of life and economic impact. If filtering is poor, a correspondingly larger economic impact will be needed to achieve the same loss of life reduction. Without good filtering, broad quarantines and social distancing are needed, incurring a huge cost in the form of negative impact on people's lives. Unfortunately, traditional approaches to filtering, such as contact tracing, are labor intensive and don't scale well. So we expect filtering (and, correspondingly, the trade-off between loss of life and economic impact) to degrade in quality as a pandemic grows. But automated contact tracing solutions have the potential to be more

scalable – and potentially even more accurate, with access to higher quality information than traditional contact tracing. This may allow for a better trade-off to be maintained in the midst of a pandemic.

The system proposed here is intended to be used as part of a broader campaign to combat COVID-19 immediately and in the long term. These methods focus on gathering and disseminating the information needed to perform targeted interventions.

2.2.3 Bluetooth Contact Tracing

Contact Tracing Background

Non-pharmaceutical methods focused on social distancing reduce the spread of COVID-19. These methods are based on reducing contact between infected and susceptible people, even when it isn't known who is infected. In the simplest form, this is being achieved by reducing all social events and increasing precautions like handwashing. This is an effective measure because the number of new infections is roughly proportional to the number of contact events with infectious people. A more targeted approach employed by many health agencies is contact tracing. This system works by finding and monitoring contacts of patients that have been diagnosed. Individuals that are thought to be infected are then put into isolation to prevent further transmission, and individuals who have previously been in contact with infected individuals are quarantined.

Hellewell et al [27] analyzed the effectiveness of contact tracing as a method to contain COVID-19 at the beginning of an outbreak. Their findings are promising: with 80% contact tracing accuracy and a mean detection time of 3.8 days after symptom onset, containment is likely. In the context of contact tracing there are three parameters that models show [38] can strongly impact results:

1. Reduction in overall transmission through social distancing
2. Testing rate and time to diagnosis
3. Contact tracing accuracy

Model Description

Mobile phones are carried by a majority of people in several countries with an estimated 3.5 billion users worldwide [39]. They are extremely common in Western society, with

over 70% of the US population estimated to own one [40]. Bluetooth is a radio protocol that can be used to wirelessly communicate between nearby mobile devices and the signal strength can be used to estimate distance. Mobile devices can be made to proactively record contact events with other nearby devices by sending bluetooth signals. By measuring the signal strength and discrete number of contact events, the duration and distance of contact between two phone users can be estimated. By recording all contact events, a high accuracy list of at-risk individuals can be generated automatically when a new person is diagnosed. These individuals can then be immediately notified to ensure they self-isolate before infecting more people. Bluetooth proximity may be the most accurate crowdsourcing method for approximating close contact to perform contact tracing.

While GPS data is a more well-known general technology, there are significant advantages Bluetooth has over GPS in terms of accuracy for contact tracing. With Bluetooth, proximity can be approximated by signal strength that is reduced by obstructions like walls; therefore, it more accurately reflects functional proximity in high-risk environments for close contact: inside buildings, in vehicles and airplanes, and in underground transit. Bluetooth communication also occurs directly between mobile devices. This means a decentralized system can be built more easily and with stronger privacy protection than other crowdsourcing data types like GPS trajectories.

We are also pursuing research in developing inexpensive external bluetooth devices under the same automatic contact tracing alert system for use in communities with fewer smartphone users. These methods would face much steeper adoption challenges, but if mobile app users and external device users could be integrated under a single system, outcomes could be further improved over more global communities.

The Bluetooth contact tracing system can be structured in a decentralized and anonymous way using randomly generated and locally stored 'Contact Event Numbers'. This allows the system to function fully without any private information being stored or transmitted. By generating a new random number for each contact event, the system is able to operate without storing or transmitting any personal information. This method is designed so that only the phones involved in a contact event are able to identify messages on a public database.

The only authentication required is the permission number provided by a public health authority. This permission number is used so that malicious actors cannot send false alarms. After authentication, the permission number is deleted from server memory. The contact event numbers are random and only known by the message recipients and the message sender, so the database can be made public without risk of sensitive information being discovered. While our current intervention is based on permission numbers, in regions

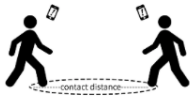
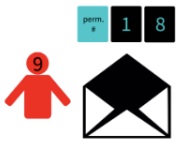
Every time two phones are close a contact event is detected	
Each phone broadcasts a randomly generated contact event number. The number is updated every few minutes	
All phones keep a local log of every number they have transmitted or received	
If one of the phone owners is diagnosed positive, they are given a permission number by health authorities.	
This person sends a packet to the public database with the permission number and their history of contact event numbers (both transmitted and received).	
If the permission number is valid, the contact event numbers are stored in the database and transmitted to all other phones.	
Each phone compares the publicly posted contact event numbers against their own history. If there are any matches this means they were close to an infected individual and are given instructions on what to do next.	

Figure 2.1: Steps for notification process

where widespread testing is unavailable, a well-designed symptom sharing questionnaire may perform a similar function, with a higher number of false positives. Research in this direction is currently being done by the CoEpi team [33]. The most effective form of this intervention would occur in communities that implement widespread testing and where permission numbers are shared with the mobile app by public health departments.

The specification for the database is very simple: it is shared across all installations of the app and stores anonymized Contact Event Numbers. If protections against hoaxes are required, permission numbers can be used, but they are not required for the privacy model. If the database grows too large, it can also be fragmented based on general location. The code and a more in depth discussion of architecture are available on the open source Github repository [41].

2.2.4 Implementation

Bluetooth contact tracing is implemented via background processes on iOS and Android. The approach currently being investigated utilizes BLE functionality for background advertisement and scanning. Due to different system requirements for Android and iOS, the protocol works differently depending on the operating systems of the devices involved. The key challenges are:

- iOS devices acting as "peripherals" in the background can only be found by "centrals" that are scanning for their specific service UUID. These peripherals must establish a connection to transfer any data.
- Android devices have several unfixed bugs where subsequent connections with many devices can cause the Bluetooth system to lock up.

The current solution is a hybrid model that is asymmetric for communication between iOS and Android. All devices will simultaneously act as peripherals and centrals, but only some devices will be able to detect others, and only some devices will need to establish a connection to exchange data. An extended description of the communication model and the code are available on Github here for iOS [42] and Android [43]. This model has been successfully implemented as a proof-of-concept, as shown here [44].

2.2.5 Quantitative Analysis of Impact

The impact of this technology will depend largely on the state of the world around it. Numerical models and ongoing campaigns [36] suggest that with extensive testing, accurate contact tracing, and isolation of suspected cases, outbreaks can be contained. For intermediate testing and contact tracing detection rates, a system like this would likely need to be used in combination with continued social distancing measures and manual contact tracing. However, the measures suggested by Ferguson et al. [37] could potentially be relaxed if supplemented with sufficient targeted interventions. Of the three target parameters (tracing accuracy, detection rate, base transmission), the potential influence of the app on tracing accuracy is the simplest to quantify. Any two users of the app who are at the same location at the same time will register a contact event. In theory, all transmission events except those by fomites would be detected. This includes all types of contact classified as being "high risk" by the CDC. Preventative measures encouraged by the app such as avoiding high-risk areas and increased precautions would reduce overall transmission rate, however it is difficult to quantify this impact. The current design of the app may also increase detection rate by informing users of symptoms to watch for and how to get tested. Ongoing research is being conducted on how to allocate testing resources to maximize the expected value of secondary cases detected.

To have a significant impact the technology will require a significant adoption rate. To detect contact events using bluetooth both individuals must have the app installed at the time of transmission. Assuming 'P' percent of the population uses the app, 'A' percent of transmission events between app users are detected, and 'T' percent of infected are tested, then $T * A * P^2$ percent of total transmissions are detected (if app users are homogeneous in the population). Figure 2.2 examines the percentage of detectable transmission events in countries with different testing rates as a function of app usage. The analysis assumes that these events are uncorrelated, which may significantly under-estimate detected transmissions due to missing chain-reaction testing. Ongoing modelling work is being done to investigate the dynamics of rapid tracing and testing along infection chains.

In order to investigate disease dynamics with clusters of app users, the model developed by Hellewel et al was modified to account for two sub-populations, one that uses the app and another that doesn't. The modified implementation is available on Github here and currently assumes 90% app tracing accuracy and 50% non-app tracing accuracy (assuming health systems are partially overwhelmed). This model was used to estimate the average number of infections caused in 6 weeks due to a single imported case. Results show a significant reduction in risk for clusters of app users that are partially distanced from non app users. This is important because it provides an additional incentive for individuals to

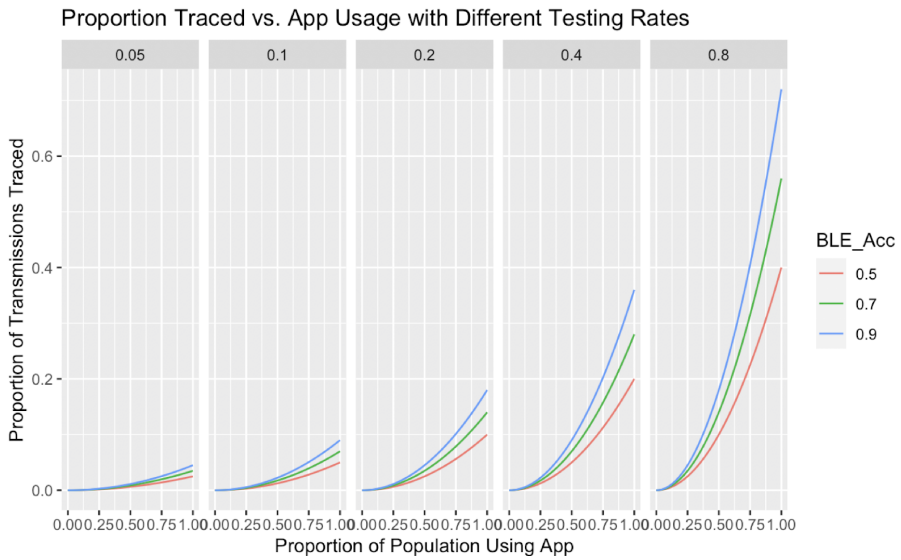


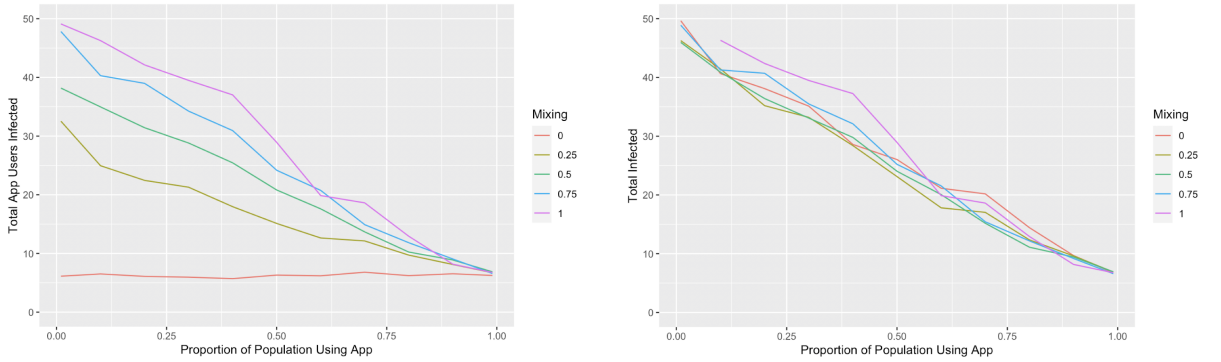
Figure 2.2: Proportion of transmissions traced depending on app usage (x axis), detection accuracy (colour), and test usage (subfigure labels)

use the app even when the overall adoption rate is low. As the proportion of the population using the app increases the risk for the entire population is reduced and most outbreaks are contained (Figure 2.3).

What we've most wanted to know is the answer to this question: "Can an effective contact tracing program reduce local transmission so that sustained local spread does not occur?" The answer looks like yes. With a comprehensive testing program, high mobile app contact tracing accuracy, and self-isolation of diagnosed individuals, our models predict that each new case could cause on the order of 10 other cases before the outbreak is extinguished. Also, even in parameter regimes where automated contact tracing alone is not enough, this technology can be used in combination with existing methodologies to provide greater protection with lower social cost.

2.2.6 Conclusions

Mobile technologies can provide instantaneous and high accuracy contact tracing, even between strangers at low social and economic cost. Instead of requiring thousands of healthcare workers to do this manually, the process will be essentially cost-free. Because



(a) Expected transmissions within app-using population

(b) Expected transmissions within non app-using population

Figure 2.3: Expected transmissions after 6 weeks within app-using and non app-using populations depending on relative sizes and mixing.

the system will be so accurate, a majority of people can resume living their lives normally without the need for indefinite social distancing. We are developing this technology as a high-quality filter to be used for the pandemic optimization problem. Combined with a comprehensive testing program, our filter may be powerful enough to stop further spread of COVID-19.

2.3 Minimizing Unintended Capabilities

The intention of the previous design was to build a base protocol so unobtrusive that a vast majority of people would be comfortable using it (which we have seen is very important for success). One way to frame system design decisions is based on the capabilities they enable. Here the goal is to build the most useful system for public health while not enhancing undesirable capabilities - either information or control given to system administrators, other users, or third parties that they did not previously have. Besides the primary goal of contact notification, there are several extended features that can be built on top of a decentralized framework - these are described in section 5.6. There are many threat models which other authors describe in more detail [45], but 3 primary uses we were concerned with were: passive surveillance by system administrators, tracking by third parties, and disruption through false notifications.

2.3.1 Passive tracking by server owner

In many alternative designs, access to the server would allow third parties (most likely governments) to passively track the location and/or interaction graph of all system users. Here, passive tracking is distinguished from active tracking because passive tracking would allow for mass surveillance of all users without detection. Some designs (but not the one considered here) could enable active tracking where attempts to gather information from users would be apparent to those users.

In the simplest centralized design, all data (interaction and/or location) is frequently shared with a single server, which then is able to check for exposures. Access to this server would allow near instantaneous tracking of all users all the time. Under a more sophisticated design (used by Singapore [46], PEPPT [47] and others) devices would broadcast information that a central server can associate with their identity or contact information. To find exposed users, the infected user would share their previously detected information with the server, allowing the server owner to decipher the identity of recent contacts. Passive surveillance using this system is more difficult but still possible. Devices placed in locations could record the information broadcast and use it to assemble the identities of nearby users. With access to many of these devices, real-time location tracking is possible.

Our decentralized design prevents this because the information emitted by devices has no meaning to observers or anyone with server access, and uninfected users don't share any information with the server.

2.3.2 Third party tracking via broadcast linking

Broadcasting persistant (or linkable) values would allow for oberservers with multiple devices to track the location of specific users over time (similar to how adtech companies used the persistent MAC addresses broadcast by smartphones to track their location). The same solution for MAC addresses can be used here- periodic randomization or pseudo-randomization (via a one way cryptographic operation) of broadcast values can be used to ensure eavesdroppers cannot link location traces over time. Most proposed protocols used some version of this strategy. At a high level the strategy is fairly simple, but care must be taken around implementation details - e.g. ensuring broadcast values are modified on the same schedule as the MAC address (to prevent intermediate linking). With our system, users who upload their infection status do share information linking their broadcast values over time. While the ability for a well-resourced adversary to link the shorterm movement of infected users is not ideal, it is limited in time and would only impact a small fraction

of users. Furthermore, such a risk is inherent to the notification process. Even with a centralized protocol it is not possible to completely prevent because the purpose of the system is to notify devices that have been in proximity to an infected user, and infections will be relatively rare events.

2.3.3 Disruption with false notifications

With unrestricted ability to anonymously notify others of exposure, malicious actors could cause many people to needlessly quarantine or get tested (perhaps to disrupt elections or similar). Our solution to this was to require proof of a positive test result in order to send a notification. When people test positive they are given a "permission number" which allows them to upload the information needed to notify their recent contacts [3]. Optional exposure sharing (which would allow notified users to share the details of their exposure with public health - see Section 5.6 for more details) could be used to detect if a single device had notified an unrealistically large number of other devices. This would alert system administrators if a high powered device were being used to falsely notify people and allow them to retract that notification.

2.4 Protocol Data Size Reduction

Soon after publishing the white paper I proposed an improvement to the protocol that could reduce data bandwidth requirements by 2-3 orders of magnitude. One of the main challenges with the decentralized approach is that to do any exposure checking, information has to be sent to every device. In scenarios with large numbers of infections and large numbers of app users, this could potentially reach an amount of data that would be prohibitive for people with only data plans and no regular access to Wi-Fi (relatively common for low-income smartphone users). Additionally, server costs for app maintainers could become too expensive for a non-profit or lower budget health department. The amount of data downloaded from the server daily is equal to $M * N^2 * P * D$, where M is the number of regions, N is the number of app users per region, P is the fraction of app users that test positive daily and D is the amount of data that needs to be transmitted per positive case. One way to reduce this number is through location fragmentation, where users

[3] While it is important to prevent abuse of the system and maintain users' trust, this step adds a large amount of friction and likely prevent many users from notifying others. CoEpi [33] wanted to allow users to notify others of unverified symptoms.

only downloaded exposure data for the region they were in - effectively lowering N and increasing M. However, this requires more fine-grained location data that people may not be comfortable sharing. My proposal was to reduce the size of the data packets by deriving the Bluetooth broadcast data from a private key using a one-way cryptographic operation. This private key could be regenerated on a time period T longer than 15 minutes, giving a reduction in data size proportional to $T/15$ minutes [4] (3000x reduction for T = 1 month, or 100x reduction with T=1 day which Apple and Google choose to use). The excerpt below contains the original description of this proposal that I shared on March 23, 2020.

Decentralized, low bandwidth, less potential for any tracking

- Each device randomly generates a monthly key and stores it locally
- The device takes the current time (rounded down to 15 minutes) and hashes it with the monthly key to generate a CEN (Contact Event Number)
- The device broadcasts this CEN and records all other CENs
- If the user is diagnosed positive they post their monthly key
- All other devices download this monthly key
- These devices locally compute the sequence of hashes for the last month using this key and compare against their records of detected CENs
- Matches indicate a contact event

Drawbacks

- The hashes need to be recomputed locally for every infected user
- This requires symmetric communication of some form (since they only know their own key, not those of the detected CENs)

Benefits:

- Decentralized and private
- Quite low bandwidth (1 million infected users would just be 16 MB download)

[4] with a maximum possible reduction of D/15min, where D is the duration contact tracing shared for. More detailed calculation here [67]

2.5 Influence

Shortly after our white paper was published[5], several other groups proposed similar decentralized protocols [48, 49, 50, 51]. In an effort to prevent fragmentation, we collaborated with several of these groups to build the TCN protocol [41].

Our low-level BLE design was based on an asymmetric communication scheme making use of both broadcast and paired BLE modes to work around bugs in Android and restrictions imposed by iOS (these choices made our app one of few that avoided a Bluetooth vulnerability found in many other apps later [52]). This design allowed communication between all pairs of devices except two iPhones that both had the app in the background. We spoke with Apple about this challenge and presented our contact tracing design in late March, but the response from them was cryptic, so we kept attempting to find workarounds (including a partial solution that uses nearby Android devices as a communication channel between iOS devices). Both the cryptography and Bluetooth protocol are described in much greater detail at [41] including a reference implementation for iOS and Android. A high-level simulation of information flows between devices, servers, and health authorities with the exposure sharing capability discussed in Section 5.6 is available at [53].

On April 9th, the MIT PACT team lead by Ron Rivest posted a preprint advocating for tech companies to implement an approach that they noted was nearly identical to the one proposed in the Covid Watch white paper [54]. On April 10th, Apple and Google announced the Exposure Notification system [55, 56] , which is built using a decentralized Bluetooth protocol very similar to the one we were the first to propose. By May 2021, over 90 million people in 38 countries had downloaded an app using the Exposure Notification API [57, 58]. Evaluation of the system has been challenging (for technical and non-technical reasons), however a study of the NHS app in England estimates that from September 2020 to December 2020, app usage prevented between 100,000 and 900,000 infections [59]. Sections 6.2, 6.3, and 6.4 contain a retrospective on system performance, compare the Exposure Notification system to alternative designs, and suggest potential improvements.

[5] and prior to this, the Singapore app team and other groups read earlier drafts of the white paper

Chapter 3

The economic value of quarantine is higher at lower case prevalence, with quarantine justified at lower risk of infection

3.1 Overview

Some infectious diseases, such as COVID-19 or the influenza pandemic of 1918, are so harmful that they justify broad scale social distancing. Targeted quarantine can reduce the amount of indiscriminate social distancing needed to control transmission. Finding the optimal balance between targeted vs. broad scale policies can be operationalized by minimizing the total amount of social isolation needed to achieve a target reproductive number. Optimality is achieved by quarantining on the basis of a risk threshold that depends strongly on current disease prevalence, suggesting that very different disease control policies should be used at different times or places. Aggressive quarantine is warranted given low disease prevalence, while populations with a higher base rate of infection should rely more on social distancing by all. The total value of a quarantine policy rises as case counts fall, is relatively insensitive to vaccination unless the vaccinated are exempt from distancing policies, and is substantially increased by the availability of modestly more information about individual risk of infectiousness.

3.2 Introduction

Control of SARS-CoV-2 transmission has not been achieved by interventions such as improved ventilation alone, but has also required social distancing. Social distancing can take the form either of population-wide measures, ranging from extreme, mandatory lockdowns to more modest voluntary behavior change, or it can be targeted via effective testing, tracing, quarantine, and isolation. Social distancing to control the spread of SARS-CoV-2 has imposed immense social costs, and is difficult to maintain for long periods. Targeting quarantine to those at higher risk of being infectious has the potential to achieve the same benefits while reducing the total harms of social distancing on the economy and mental health of the population. Here we calculate which combination of population-wide social distancing plus targeted quarantine will minimize harms while controlling transmission to the same degree. We operationalize this question by comparing the harms of different policy combinations that all achieve the same target value for the effective reproduction number. Figure 3.1 demonstrates how quarantining a greater fraction of infectious individuals allows overall social distancing to be reduced while holding the number of expected transmissions constant

Our approach, at the interface of infectious disease epidemiology and economics [60], solves a number of problems previously encountered at that interface. First, it is difficult to decide on an exchange rate between social restrictions / livelihoods versus lives [61]; compounding this is the fact that selfish agents may count only the cost to their own life, and not that of others they infect [62]. Difficulties determining an exchange rate lead some to present only idealized curves without a concrete recommendation [63, 64]. Idealized curves can lead to abstract insights, but are of limited use to practical decision-making.

We avoid the need for an exchange rate between social restrictions and infections by instead comparing one form of (population-wide, partial) social restrictions with another form of (targeted, more stringent) social restrictions. This allows the lowest cost strategy to be selected without having to monetarize the cost of health outcomes, the latter being held constant in our analysis. The policy optimization problem is thus framed as a minimization of a cost function across different degrees of isolation for different individuals, while holding the benefits of reduced disease transmission constant.

Second, much prior work integrates across an entire sweep of a pandemic modelled using an SIR approach [65]. These projections have often been inaccurate. We instead focus on a moment in time, an approach that is enabled by a control theory view. At the beginning of the COVID-19 pandemic, populations experienced rapid exponential growth, e.g. doubling in as little as 2.75 days in New York City [66]. Through some combination of top-down

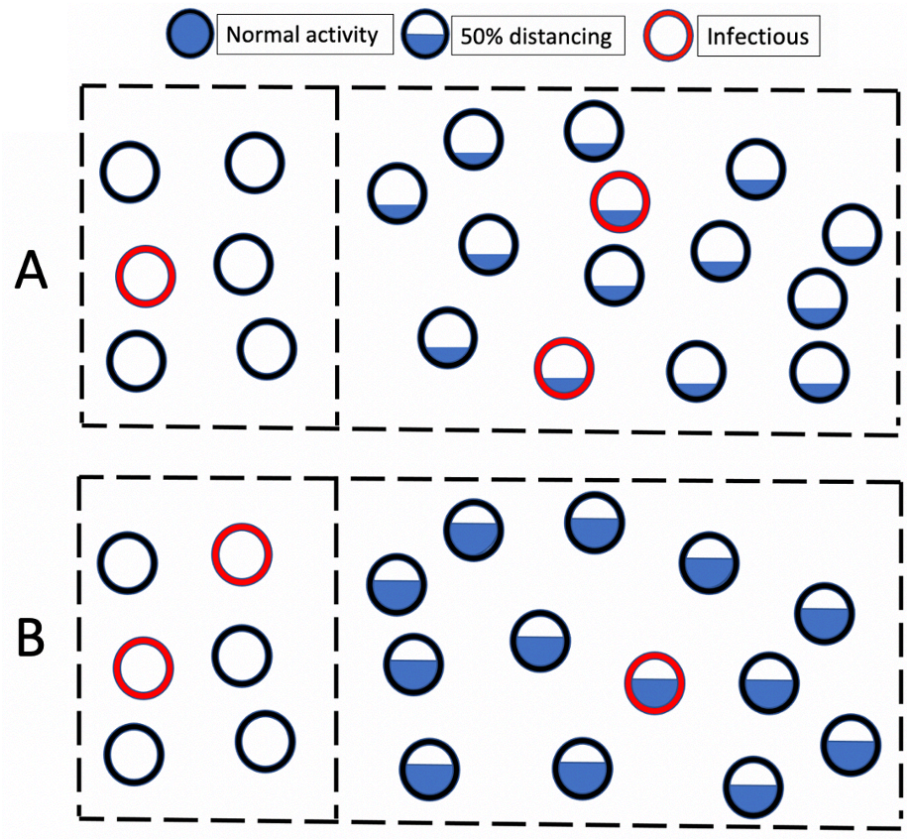


Figure 3.1: Two scenarios of quarantine and social distancing that give rise to the same number of expected transmissions. A) One out of three infectious individuals is either quarantined or isolated, and the general population is at 1/3 of normal social activity. B) Two out of three infectious individuals are either quarantined or isolated, and the general population is at 2/3 of normal social activity. By finding and quarantining infectious people, social distancing imposed on the general population is significantly reduced.

control measures and individual behavior modifications, many populations subsequently achieved relatively flat case counts, i.e. an effective reproduction number remarkably close to 1 [67, 68, 69]. Even locations that subsequently lost control of the pandemic have tended to eventually issue stay at home orders leading to exponential decline. Failing that, individuals tend to modify their behavior to reduce personal risk once healthcare systems are overwhelmed. Over the very long term, a geometric mean of the effective reproduction number not much greater than 1 is inevitable (see Appendix A.2). This implies that there is a form of control system (whether intentional or not) using social distancing to regulate transmission rate.

Third, many models consider decision-making that is informed by perfect, instantaneous knowledge [65]. This is problematic given that key indicators such as hospital usage or even positive tests have a marked lag relative to infections. Our approach focuses instead on making optimal use of available, probabilistic information. Past economic approaches to modelling the transmission of SARS-CoV-2 treat more stark examples of information, e.g. testing to find out who is infected and should isolate, as opposed to who is exposed to what degree and should quarantine [65, 70]. But quantitative information about individual risk is available from a variety of sources. For example, the risk of infection with SARS-CoV-2 can be estimated based on proximity and duration of contact with a known case, and their estimated infectiousness as a function of timing relative to symptom onset date [71]. Similarly, setting local rather than global shutdown policies, in the light of differences between regions, can lower costs [72].

Our work ultimately describes the value of information, specifically information about who is at high enough risk of being infected to quarantine strictly rather than merely to conform to population-wide restrictions. Given an individual estimated to be infectious with probability 'r', we propose a method for deciding whether to quarantine this individual. We do so by weighing the cost of quarantine against the degree to which indiscriminately applied social distancing would need to be increased to achieve the same reduction in transmission. Our approach informs choice of the lowest cost strategy needed to achieve epidemiological targets. In particular, we formalize the intuition that populations with low prevalence should widen the net of who they quarantine.

3.3 Model

3.3.1 Optimal Risk Threshold

Consider a well-mixed population of size P, of which I people are currently both infected and infectious, and S people are susceptible. According to a standard continuous, deterministic SIR model approach [73], the effective reproduction number is defined to equal $R_t = R_0 * S/P$, i.e. it depends on the the basic reproduction number R_0 and the fraction of people S/P who are still susceptible. Here we define R_0 and R_t to explicitly exclude interventions such as social distancing or quarantine, because these are the interventions whose magnitudes are being optimized. However, R_t is intended to include fixed cost interventions such as improved ventilation.

Population-wide social distancing is parameterized as a value D that varies between 0 and 1 such that the reproductive number is proportional to $(1 - D)$. I.e., a value of $D = 0$ indicates normal social contact, while $D = 0.5$ indicates that total social activity is reduced such that expected transmissions are halved. Reductions in D come from a variety of behavior changes including working from home and reducing social contact, and also include measures to mitigate the danger given contact, such as wearing masks, and meeting outside and/or at greater physical distance.

Let Q_i denote the number of infectious people who are quarantined or isolated and Q_n denote the number of non-infectious people who are quarantined. The benefit from a targeted quarantine policy depends on $\frac{I-Q_i}{I}$, the fraction of infectious cases who are not successfully quarantined and then isolated. The effective reproduction number R_t' after both social distancing and quarantine interventions are applied is given by:

$$R_t' = R_t * (1 - D) * \frac{I - Q_i}{I}. \tag{3.1}$$

R_t' gives the expected number of onward transmissions per infected case in the general population. Equation 3.1 is undefined when the number of locally transmitted cases $I = 0$. Even for a low but non-zero value of I, a different, stochastic treatment warranted - this is discussed in Appendix A.3.

Consider an individual reducing their social contact by a factor of x. Here $x = 1$ for complete loss of all social contact and $x = D$ for the average individual not in quarantine or isolation. Let the function $f(x)$ indicate the cost of such reduction. I.e. if $f(1) = 3f(0.5)$, this means that the average individual finds a day of total quarantine to be an equivalent burden to 3 days of half-quarantine. The exact functional form of $f(x)$ is unknown, but

in Model Section 3.3.2 we show that the results have a bounded dependence on its form. The total cost to a population is given by Expression 3.2, based on $Q_i + Q_n$ people having no social contact, and $P - Q_i - Q_n$ people with contact reduced by D.

$$f(1)(Q_i + Q_n) + f(D)(P - Q_i - Q_n) \tag{3.2}$$

The amount of social distancing needed to meet a target reproductive number, R_{target}, given fixed values for Q_i and Q_n, is given by:

$$D = \max\Big\{0, 1 - \frac{R_{target}}{R_t}\frac{I}{I - Q_i}\Big\}. \tag{3.3}$$

This value is bounded at 0; if $R_t\frac{I-Q_i}{I} \leq R_{target}$, the target has already been met and no social distancing is required.

In the real world, it is not possible to adjust D instantaneously or exactly in response to changes in the impact of quarantine policy. However, in Supplement Section 2, we show that Equation 3.3 is insensitive to time and control uncertainty, so long as caseload and control measures are fairly steady over time, as has been the case in many regions some substantial periods of time [67, 68].

Combining Equations 3.2 and 3.3, the total cost of a quarantine policy, in combination with the degree of population-wide distancing that is necessary to achieve R_{target} given that policy, is given by:

$$J(Q_i, Q_n) = f(1)(Q_i + Q_n) + f(\max\Big\{0, 1 - \frac{R_{target}}{R_t}\frac{I}{I - Q_i}\Big\})(P - Q_i - Q_n). \tag{3.4}$$

This equation can be used to compare the overall cost reduction achieved by different quarantine policies, given in terms of the total reduction in social distancing that is enabled.

We assume that an estimated risk of infectiousness is available when considering whether to recommend quarantine to an individual. Risk could be estimated using proximity and duration of contact with a known case as described in [71]. Or it could be estimated for members of a subpopulation like a workplace by rapid testing of a random sample of that subpopulation and projecting the proportion positive onto the remainder. When a person with risk of infectiousness r is quarantined, Q_i is incremented by 1 with probability r, and Q_n is incremented by 1 with probability $1 - r$. Quarantining this person is worthwhile if the expected cost J if they quarantine is less than the expected cost if they do not, giving the inequality:

$$rJ(Q_i + 1, Q_n) + (1 - r)J(Q_i, Q_n + 1) < J(Q_i, Q_n) \tag{3.5}$$

Solving the corresponding equation for r, the optimal risk threshold is given by:

$$r_{thresh} = \frac{J(Q_i, Q_n) - J(Q_i, Q_n + 1)}{J(Q_i + 1, Q_n) - J(Q_i, Q_n + 1)} \tag{3.6}$$

To apply Equation 3.6, we first consider the simplest equation for the cost of loss of social contact, namely the linear function $f(x) = x$. Using this equation in a situation where some social distancing is necessary ($\frac{R_{target}}{R_t} * \frac{I}{I-Q_i-1} < 1$), Equation 3.6 simplifies (as shown in Appendix A.1) to:

$$r_{thresh} = \frac{I - Q_i - 1}{P - Q_i - Q_n - 1}. \tag{3.7}$$

In this simple case, the interpretation of Equation 3.7 is straightforward and intuitive. The optimal risk threshold for quarantine is equal to the disease frequency in the community, excluding from that community both those already isolated or quarantined, and the focal individual. This implies that in regions with very low prevalence rates, much stricter quarantine requirements should be used, making the general population less impacted by disease-control measures. In populations with higher disease prevalence in the community, there is less benefit to quarantining a single individual at low risk, because they represent a smaller fraction of the overall disease risk to the population.

3.3.2 Sensitivity to Cost Function

We do not know the shape of the cost function $f(x)$ as the extent of reduction in social contact x varies. However, some non-linearity is expected. E.g., to achieve total rather than partial quarantine can require a specialized quarantine facility in addition to a more extreme loss of individual utility. We define D to be constructed such that benefits in reducing disease transmission are linear in D. When considering a reduction in the extent of social activity, corresponding benefits are likely somewhere between linear and quadratic [64]. When people seek out places that have variable levels of crowding, benefits are quadratic; when they seek out people who are still willing to meet with them, benefits are linear. By constructing D to have linear benefits, we pass this uncertainty in human behavior to the cost function. Disparate considerations all suggest a concave-up shape for $f(x)$.

In Appendix A.4, we show that the deviation from Equation 3.7 is relatively small for reasonable choices of $f(x)$. In particular, if the threshold r_1 is computed using $f_1(x) = x$

and the threshold r_2 is computed using any $f_2(x)$ that is strictly increasing and concave up, then $r_2 \geq r_1$ and $\frac{r_2}{r_1} \leq \frac{\max_{x\in[0,1]} f_2'(x)}{\min_{x\in[0,1]} f_2'(x)}$. In practice, the value of r_2 is often much closer to r_1 than indicated by this upper bound. For example, while the ratio between optimal risk thresholds for the two cost functions shown in Figure 3.2 is bounded at a 10-fold difference, we numerically find it to be at most 4.1-fold.

Both the costs and benefits of reduced social contact can vary among individuals. For example, the cost is higher for essential workers than for those that can easily work from home, while the benefit is higher for those whose jobs expose more and/or more vulnerable people. Equation 3.7 can be modified to accommodate knowledge about an individual. If refraining from social contact is m times more expensive for this person than for the general population, and the potential danger to others if they are actually infectious is d times that of an average case, then the risk threshold for this person should be modified to $r_{thresh} * m/d$. Note that this approach differs substantially from suggestions to target isolation directly to the elderly and other high medical risk individuals [64]. We instead recommend increasing the stringency of quarantine among their contacts.

For a given individual, the cost of quarantine is likely to depend more than linearly on quarantine length, due both to logistical and to psychological factors. To accommodate this, the cost of quarantine could be modelled as time-dependent by scaling $f(x) * m(t)$. The benefit of quarantine is also not equal for each possible day of quarantine, given a probability distribution for incubation period and a quantitative timecourse of infectiousness [71]. Consideration of these two factors will lead shorter recommended quarantine durations, targeted to the most infectious days. Testing individuals in quarantine could further shorten its duration [74].

The remarkable thing about Equation 3.7 is its insensitivity to R_t and R_{target}. To determine the degree to which this might generalize, in Appendix A.4 we derive Equation 3.8:

$$\frac{1}{r_{thresh}} \frac{f(1) - f(D_0)}{1 - D_0} \approx \frac{1}{r_{pop}} f'(D_0) \tag{3.8}$$

Equation 3.8 shows that r_{thresh} is chosen so that the risk-weighted cost of quarantine is equal to the marginal cost of distancing for the general population. In general, r_{thresh} depends indirectly on R_t and R_{target} because they influence the set point, D_0. However, for a linear cost function, the slope ($f'(D_0)$) and average slope ($\frac{f(1)-f(D_0)}{1-D_0}$) are constant, and therefore r_{thresh} only depends on the disease prevalence in the general population. In other words, when the cost per individual is a linear function of distancing, then the cost is the same whether that distancing is all directed at making one individual quarantine completely

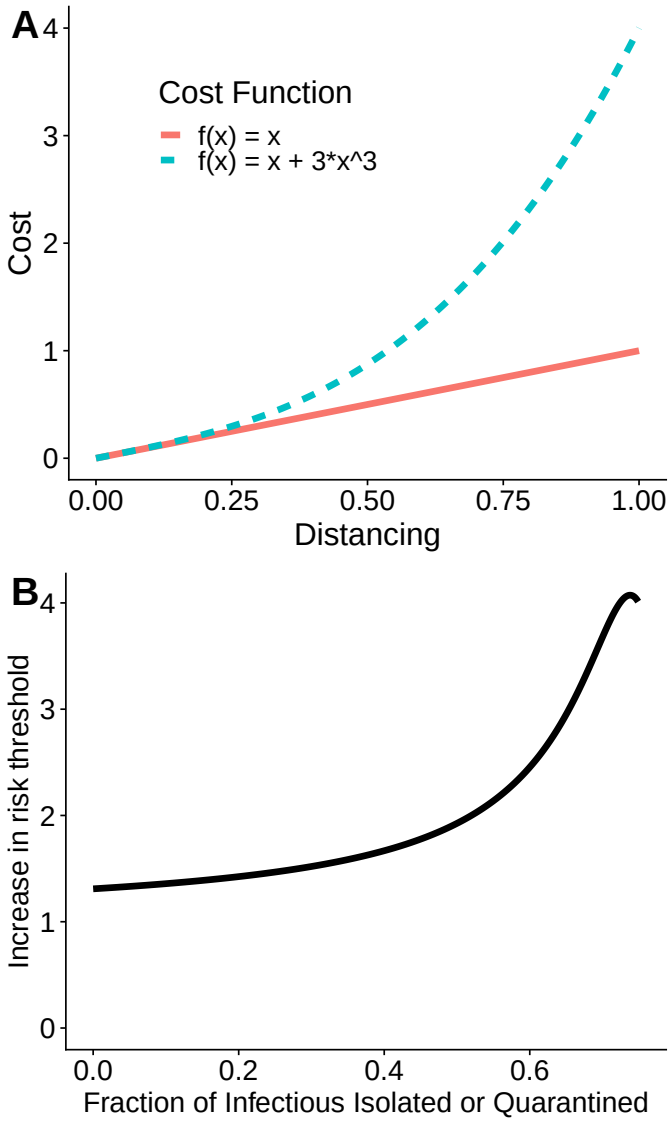


Figure 3.2: Non-linearity in cost as a function of the degree of loss of social contact (A) leads to a modest four-fold change in optimal risk threshold (B), or less, depending for example on the proportion of infectious people quarantined or isolated (Q_i/I). Risk thresholds were computed using Equation 3.6 with $J(Q_i, Q_n)$ defined by substituting the two functions shown in (A) into Equation 3.4.

instead of following current distancing norms, versus distributing the additional distancing among all individuals. For a non-linear function, these two costs are no longer equal. D_0 is proportional to R_{target}/R_t, and so Equation 3.8 describes the degree to which non-linearity makes r_{thresh} depend on these terms. This dependence will be most pronounced when the local slope of $f(x)$ around current distancing levels and the mean slope from current distancing to $D = 1$ are substantially different. However, because quantitatively, the optimal risk threshold is not very sensitive to the cost function, we focus our analysis on the more convenient linear cost function (with solution given by Equation 3.7).

3.4 Results

3.4.1 Applications in time and place

In a large population with many cases, the risk threshold for quarantine r_{thresh} is closely related to the base rate of infection in the population. Here we give several examples in which we apply this finding, to illustrate how dramatically this causes optimal policy to vary across different places and times.

In British Columbia in October 2020, rapid SARS-CoV-2 testing was widespread, but only when symptoms were present, making official case counts of around 100/day a good estimate of symptomatic cases, such that total cases can be estimated in proportion. (When testing is less adequate, estimating the true number of cases is more complex, involving projections from death rates and/or testing rates, but has been attempted e.g. at [67] or [75].) Assuming another 20 asymptomatic or other undiscovered cases per day, each with a 10 day infectious window, yields 200 non-isolated infectious individuals on any given day. In addition, we estimate 4 infectious days per discovered case prior to isolation, yielding another 400 non-isolated infectious individuals on any given day. The base rate of infectiousness, conditional on not having been isolated, is thus of order 600/5,000,000=O(10^{-4}). From Equation 3.7, a quarantine risk threshold of 10^{-4} is appropriate, or a little higher if a non-linear cost function is assumed in Equation 3.6. This implies that quarantine recommendations for entire schools or workplaces are worthwhile if they are expected to isolate one more infected person than would be achieved via traditional contact tracing.

In contrast, North Dakota in October 2020 had around 400 detected cases per day and perhaps another ∼400 undetected in a population of ∼800,000. This yields ∼5000 undetected infectious individuals on a given day, or a 0.6% base rate and corresponding risk threshold. Several studies have found the secondary attack rate to be on the order

of 1/10 - 1/100 [76, 77, 78], so this risk threshold aligns well with standard guidelines for the minimum contact considered to be close for contact tracing purposes. We note however that a 14 day quarantine is more than is needed to put the conditional probability of infectiousness, given lack of symptoms, of a non-household close contact below the base rate [71].

Imported and exported cases introduce complications. Policy is best set for regions that are relatively self-contained, and do not e.g. cut off commuter communities from one another. When imported and exported cases make up only a small fraction of total cases, then subtleties in using case counts to estimate R'_t (discussed in Supplement Section 3) have negligible effects. When local transmissions are abundant enough such that elimination is not a near-term goal, then incoming individuals can therefore be quarantined by comparing their infectiousness risk to r_{thresh} on the basis of base rates of infection in the location they arrived from, plus the risk of travel. These considerations apply e.g. to the need to quarantine following travel from one part of the U.S. to another.

Some locations, e.g. British Columbia, achieved very low levels of local transmission, but with a degree of economic connectedness with harder hit locations that made infeasible sufficiently strict quarantine to put the risk from incomers below r_{thresh} as described above [79]. This is because the marginal cost of quarantine applies not just to the quarantined individuals, but to trucking routes and other aspects of the economic system. At this point, some significant expected number of imported cases acts as a forcing function outside the exponential dynamics of R'_t. In this case, it might be better to set a target value for the expected number of cases per day, and then to set R_{target} to achieve this in combination with imported cases.

If local elimination is achieved, including successfully quarantining all imported cases, then R_{target} can be relaxed up to R_t, removing all social distancing, with huge social gain. The disproportionate benefits warrant careful attention to avoid even a single imported non-quarantined case. How best to achieve this is best quantified by a stochastic model, outside the scope of this paper. Qualitatively, we note that the cost of letting a single case slip through, and hence having to return to social distancing, is higher for larger populations. The harm is also magnified by delays in realizing that an outbreak is underway, and hence the extent of harm depends strongly on local surveillance and contact tracing capabilities.

When different regions come under shared political control, importation risk can be controlled at the source rather than the destination. An example is Vietnam in March 2020. Viewed as an entire country instead of as smaller regions, it is optimal to quarantine at a threshold of $1/10^6$. This would motivate quarantining entire cities so that the rest of

the country doesn't have to substantially modify their behaviour. This is similar to the approach Vietnam actually took, with up to 80,000 people in quarantine at a time from regional lockdowns and aggressive contact tracing [80].

3.4.2 Quantifying the Value of Quarantine

The value of a test, trace, and quarantine policy can be quantified in terms of the net reduction it enables in full or partial person-days of avoiding contact with others. By definition, the marginal value of quarantining an individual of marginal risk r_{thresh} is 0. This implies, for a cost function that is uniform across individuals and neglects complications due to compliance, that quarantining one definitely infectious person for one day yields a benefit of C/r_{thresh} where e.g. C is the assessed cost of quarantine. Taking into account the cost of quarantining the focal individual, the net benefit is $C/(1/r_{thresh} - 1)$.

In many countries, compliance not just with social distancing, but also with quarantine has been low. This is to be expected, given the uncompensated cost of quarantine to individuals, who are currently asked (or in cases coerced) to sacrifice for the public good. We advocate that governments align incentives, by guaranteeing e.g. 150% the individual's normal daily income. This approach also serves to convert the costs and benefits calculated here, whose units are of person-days of elimination of social contact, to a dollar value for quarantine. On the basis of median income, and taking into account that some individuals are able to work from home and so will not draw the full 150%, this would in an affluent nation amount to something on the order of USD$150 per day of complete quarantine. This is the implied utility sacrificed by either total quarantine or equivalent social distancing D. I.e., a dollar value on the order of USD$150 per day is what one would need to pay people in order to incentivize their compliance, i.e. to achieve neutrality among preferences. Because this neutrality between quarantine plus payment vs. no payment and no quarantine does not include the dangers posed by COVID-19 itself, the existence of such payments would not incentivize deliberate exposure. Redistributive payments move costs from the quarantined individuals to the taxpayer.

With this cost applying to both indiscriminate distancing and targeted quarantining, we can assess the value of a quarantine policy from its ability to reduce the total amount of distancing required to achieve R_{target}. Given the changing nature of a pandemic, we calculate how this total value depends on case prevalence and on the extent of population immunity.

When case prevalence is low, quarantining a single individual makes a larger proportional difference to the risk to others, enabling a greater relaxation of social distancing.

We see in Figure 3.3B that in the pertinent range of 0.1%-5% case prevalence, the value of effective quarantine policies is high, and increases as case counts drop. Results are similar for the differently shaped distributions of assessable risk shown in Figure 3.3A. For lower case prevalence, the value flattens out because the benefit from quarantining one infectious person rises at the same rate that the number of total infectious people drops, while the number of infectious contacts quarantined per index case remains almost constant.

Vaccination has enormous value in reducing the amount of distancing required to achieve R_{target}. But unless the immune are exempted from distancing norms, the marginal value of quarantine, in helping sustain R_{target} with even less distancing, is unchanged. The effective reproductive number before including distancing, R_t, is proportional to the susceptible fraction, S/P. Under a linear cost function, we see from Equation 3.7 that the threshold for quarantine is insensitive to R_t. This means that the value of a quarantine policy is also insensitive to R_t and hence to this change in the vaccination status of the population. However, if people who have immunity are exempt from social distancing, not only is the amount of social distancing per person reduced, but also the number of people who have to distance. This has an additional impact on r_{thresh}: the effective size of the population is reduced from P to S. Reducing the effective size of the population increases r_{thresh} by replacing P with S in the denominator of Equation 3.7.

In Figure 3.4, the value of a contact tracing system is shown as a function of the susceptible fraction of the population. If the immune are required to socially distance, the benefit of the program is independent of the proportion immune. If the immune are exempt, the benefit of the program is nearly proportional to the susceptible fraction. Intuitively, this is because quarantining infectious cases only allows an increase in social activity for the susceptible.

3.4.3 Marginal Value of Surveillance Testing

With a method available to assess the value of a quarantine policy, we can now assess the value of information in enabling better quarantine/isolation policies. We first consider random surveillance using a test with 50% sensitivity and near-100% specificity (the latter achieved via a follow-up PCR test). The value per test is given by the product of the probability of testing positive, expected number of additional infectious days in isolation, and the daily benefit of each additional isolation day. If a positive test leads to 3 extra isolation days, and assuming a large enough population to neglect the -1 term in Equation 3.7, the direct benefit (not including onward contact tracing) per test is $\frac{I-Q_i}{P-Q_i+Q_n} * 0.5 * 3 * \$150/r_{thresh}$. With $r_{thresh} \approx \frac{I-Q_i}{P-Q_i+Q_n}$, the per-test benefit is insensitive to the risk

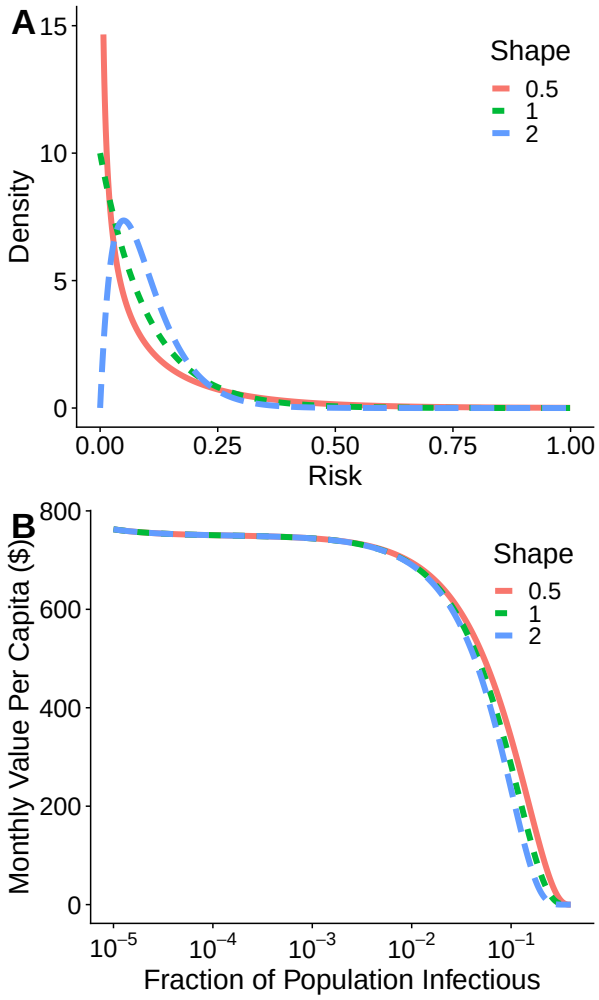


Figure 3.3: The benefits from quarantine policies are highest when case prevalence is low. In this example, each case one serial interval prior has on average ten contacts, of whom one is infected, such that $R'_t = 1$. Higher case prevalence corresponds to high r_{thresh} and so lower D. As D decreases, the expected number of contacts each un-quarantined case has increases, but the number of positive cases in quarantine increases so that the overall expected number of contacts stays constant (maintaining $R'_t = 1$). We assume that contact tracing succeeds in reaching 10% of all contacts. A) Assessable risk among contacts is modeled as a gamma distribution, in all cases with mean 0.1, and with different shape parameters to explore the importance of resolution among low-risk individuals. B) $Q_i = 0.4I, Q_n = 4Q_i, C = \$150, R_t = 4, R_{target} = 1, f(x) = x$. While $R_t = 4$ is higher than some estimates for SARS-CoV-2, other estimates place R_0 as high as 5.7 even during early spread in China [3], and higher for new variants.

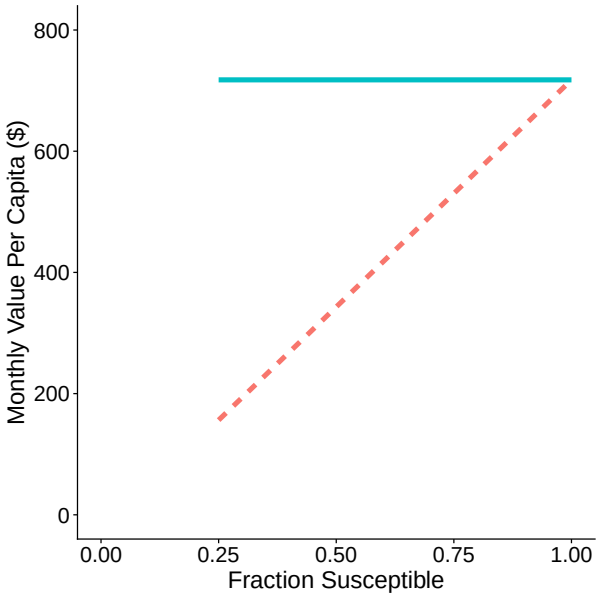


Figure 3.4: With vaccinated people exempt from distancing, the benefit of a contact tracing system is approximately proportional to the susceptible fraction (dashed line). When distancing is required, with the immune following the same distancing norms as the susceptible, the benefit does not depend on the susceptible fraction (solid line). $I/P = 0.0055, Q_i = 0.4 * I, Q_n = 4 * Q_i, C = \$150, R_0 = 4, R_{target} = 1$ Assumptions about contact tracing are as for Figure 3.3. A linear cost function is shown, but the approximately proportional relationship also holds for a non-linear cost function, with immunity adjusting the effective population size.

threshold and equal to ≈ $225. With a non-linear cost function, the benefit might be up to ten times less.

This makes random surveillance clearly worthwhile for rapid testing technologies that can be made available for as little as $5 per test [81]. The benefit of surveillance testing increases with when close contacts are traced and then tested and/or quarantined. The benefit is reduced if a fraction of people who test positive do not effectively isolate. Note that previous estimates of the value of random surveillance supported less intensive measures [82], because they did not account for the benefits in terms of population-wide relaxation of other measures.

When testing is too expensive to be used for random surveillance, or merely too slow to roll out at sufficient scale, it can achieve higher marginal value if targeted to populations at higher risk. We note that it can be hard to justify strict quarantine on individuals of modestly above-average risk when the risk threshold is low in absolute terms, especially when elimination does not yet seem within reach. An alternative is daily testing instead of quarantine. The economic benefit from testing increases with risk of the focal individual relative to the average, and can thus justify more expensive and more sensitive daily tests.

3.4.4 Marginal Value of Distinguishing among Exposures

Current guidance from the World Health Organization [83] and the Center for Disease Control and Prevention [84] calls for quarantine for those within 1 meter or 6 feet, for 15 minutes, from 2 days before an infected individual's symptom onset date to 9 days after. This guidance does not correspond to a consistent risk threshold, because it combines three binary thresholds rather than combining three risk factors prior to applying a threshold [71], because the infectiousness window does not reflect current science but instead a paper for which a formal correction has been issued [85, 86, 87], and because additional information such as ventilation in an indoor space is not taken into account. In addition, once a consistent risk assessment can be produced, we have shown here the benefits of adjusting the threshold based on local conditions.

While complex risk assessment can be difficult to apply in real-time case investigation interviews, exposure notification apps [34, 28] can easily do so [71, 88, 89]. Here we consider the economic value of using a better rather than a worse risk assessment algorithm. Specifically, we consider as an example the economic benefit that would arise from modifying the Google/Apple Exposure Notification protocol so that it could sense whether the user was indoor vs. outdoor, and differentiate risk between the two.

We assume a 50:50 distribution between indoor and outdoor exposures, each of which have a gamma distribution of risk with shape parameter 2 as in Figure 3.3, and a 15-fold difference in mean risk [90]. We set the mean risk across both to be 0.06. When the risk threshold is low, there is no economic benefit from distinguishing indoor vs outdoor, for the simple reason that quarantine is recommended to anyone with any recorded exposure. However, when the risk threshold is high, knowing that an exposure occurred outdoors can spare an individual from costly quarantine, while also quarantining indoor exposures that occurred at greater distance.

With a 2% risk threshold, for the average exposed person, a fraction of whom are told to quarantine, the net benefit is 2.32 quarantine-days averted with the information and 2.05 without. These units of quarantine-days averted include both targeted quarantine and indiscriminate social distancing, which can be summed under a linear cost function.

Consider a population of 10 million, 2% of whom test positive over the time period of interest. We assume 20% use the app, of whom 50% enter their diagnosis into the app, thus notifying 20% of contacts that the app scores as above threshold risk, who go on to infect half as many people as they would were they not contacted. We assume the cost of quarantine is similarly only half what it would be given complete compliance. We assume that each has 16.7 contacts (set so that $R_t = 1$) following the distribution described above.

We assume that the cost of quarantine is \$150 per day, and that the average quarantine is 10 days long. This gives an expected cost of quarantine of \$1500 per individual that a particular configuration succeeds in getting to quarantine. At the margin, this indicates the value that society implicitly places, i.e. that removing a 0.02 risk is worth \$1500. The benefit of quarantine is calculated in terms of "excess risk" above 0.02 having a value where each 0.02 of excess is worth \$1500. Under these assumptions, net benefit comes out to \$116 million with information about indoor/outdoor contact (i.e. \$11.60 per capita or \$58 per app user), and \$105 million without. This value of \$58 per app user can also be used to calculate a marginal return on further investment in marketing for higher adoption, noting the non-linearity that will tend to produce accelerating rather than diminishing returns on investment.

We note that the value of an exposure notification app would be dramatically higher if, instead of a blanket 14 day quarantine duration, a shorter quarantine, targeted to the days of greatest risk, were used [71]. Further gains would come from using negative test results to shorten quarantine.

3.5 Discussion

Here we showed that the socially optimal policy is to quarantine an individual if their risk of infectiousness is even mildly above that of the average person in the population who is not under quarantine. How much above depends slightly on the non-linearity of the cost of isolation with the strictness of isolation. While only order of magnitude calculations are sometimes possible, given that case prevalences vary over several orders of magnitudes across different populations, our rough calculations are nevertheless instructive. Some tools, like exposure notification apps, have quantitative sensitivities that can be tuned in real time in ways informed by such calculations. With more information about risk, quarantine policy can make a greater contribution to returning to a more normal life. The value of quarantine policy in doing so is higher for low case prevalence. It depends on vaccination prevalence only if the vaccinated follow different distancing norms than the unvaccinated, or if population immunity on its own is extensive enough to achieve $R_t < 1$.

To simplify the optimal trade-off between targeted quarantine and broad social distancing, we have neglected two complications. First, we have based recommendations as to who should quarantine on the assumption that they will do so completely. In reality, quarantined individuals often reduce rather than eliminate non-household contact, let alone all contact. More nuanced messaging regarding the degree of quarantine may be difficult to manage.

Second, how to choose R_{target} is not specified here. In the Introduction and Supplement Section 2, we discuss a control theory perspective on the fact that social distancing adjusts to keep the long-term geometric mean of R_t near 1. A well-functioning control system is one with a negative feedback loop such that the error in achieving a desired outcome feeds in as a correcting input [91]. Feedback can come either from top-down policy or from changes in individual decisions in response to conditions.

A control system can be described as a combination of proportional, integral, and derivative control [91]. Derivative control would set social distancing policy based on whether exponential growth has resumed, and if so with what doubling time. This is unlikely to occur spontaneously, but rather requires government action in response to epidemiological reports not yet posing significant personal risk to the average individual. Spontaneous control by self-interested individuals responding to local conditions would under ideal circumstances follow proportional control, i.e. reflect current case counts. More likely, individuals will respond to the integral of case counts over time, e.g. as reflected in hospital occupancy rates. Targeting hospital occupancy rates (a combination of proportional and integral control) is also the explicit policy of some governments, e.g. the

State of Arizona [92]. If the goal is a managed approach to naturally acquired population immunity [93], hospital usage is a reasonable target. But since the amount of social distancing (whether mandated or voluntary) needed to keep R_t near 1 is the same against a background of low case counts as with high case counts, there are obvious health benefits to doing so with low case counts. Integral control will lead to significant fluctuations with a high total case burden.

We recommend mostly derivative control, with sufficient proportional control to ensure return to low case counts following fluctuations. The optimal extent of proportional control depends on whether shorter, sharper action vs. moderate decline with more social freedom achieves the least harm in transitioning between different case prevalences. The choice of R_{target} to control the speed with which case numbers are brought to low levels must also place a value on illness and death in comparison with social restrictions, a difficult problem.

Regardless of the form of feedback, compliance with transmission control measures such as size limits to gathering, compulsory masking etc. can be hard to predict, and hence the degree of top-down control is limited. This underlies our decision in this manuscript to focus on marginal costs given prevailing policy and behaviors. Our approach exploits the fact that some unspecified form of negative feedback keeps R_t, or at least its geometric mean, near 1.

If testing and tracing improves, then less population-wide distancing will be required to achieve the same value of R_{target}, and if the same level of distancing is maintaining, this will manifest as a drop in R'_t. This might on occasion lead to a reconsideration of R_{target} to a lower level once that seems achievable at lower cost, but otherwise, reductions in social distancing should again ideally be triggered via derivative control, reducing the burden on the population.

While some form of negative feedback can be inferred from the empirical tendency of R_t to stay near 1, positive feedback can also occur: as case counts rise, contact tracing becomes more difficult [94], potentially accelerating the outbreak. Consciously adjusting the risk threshold that triggers quarantine, whether regarding incoming travellers, or risk thresholds in an exposure notification app, contributes to this positive feedback loop. However, conscious recognition of this positive feedback loop, and the limits it creates, can focus contact tracers' attention where it can do the most good, in a manner that is under straightforward real-time control.

Because of the positive feedback associated with contact tracing workload and optimal quarantine policies, control of R'_t via negative feedback, e.g. to bring it back down to its target given a rise due to seasonal conditions or simply pandemic fatigue, is better

exerted through population-wide measures to affect D. Our work shows that this is the socially optimal approach, and indeed if performed adequately, should be accompanied by a perhaps unintuitive relaxing of quarantine. Our formal development assumes a perfect control system, and focuses on one-individual perturbations to an equilibrium. It is likely that D will not reliably adjust in a completely timely manner, but a proactive approach can improve the chances of this occurring.

While contact tracing apps enable some precision in setting risk thresholds, more broadly our findings are primarily useful for quickly assessing the direction in which policy should change. For example, if the secondary attack rate amongst quarantined contacts is 10%, but the base rate in the general population is $\ll 10\%$, our work implies that it would be greatly beneficial to cast a wider net in who qualifies as a contact and how aggressively to trace them. The theoretical framework of the paper can thus be used to inform policy even without aspiring to achieve precise optimality.

Many extensions to our approach are possible. Treatment of stochastic effects at low case counts, especially related to border policy, would expand the scope of this work to enable optimization of quarantine and testing decisions when local elimination of the disease is possible. Another way to strengthen the model would be to introduce sub-populations with different propensities for social distancing. Finally, the model presented is intended for periods when a new outbreak is being suppressed to wait for better vaccines and treatment and to reduce the burden on hospitals. Permanent social distancing is not a desirable policy, so an analysis that depends on a trade-off between social distancing and quarantine has less validity in the endemic phase. A better approach for the endemic phase would be for decisions about quarantine and broader non-pharmaceutical interventions to trade off against health outcomes.

The value of quarantine policies is highest when case counts are low. However if low cases counts are a consequence of immunity (whether from natural infection or from vaccination), then the value of a quarantine policy will go down when those with immunity follow different distancing norms from those without. There is a direct benefit from relaxing distancing for the immune, whether this behavior is self-regulated or whether it comes in the form of a more official "immunity passport". Immunity passports were once problematic because they would have set up an adverse incentive in favor of contracting SARS-CoV-2. However, with increasing vaccine availability, the benefits from immunity passports have risen at the same time as an adverse incentive to get infected has been converted into a pro-socially aligned incentive to get vaccinated. Our approach can be used to merge immunity history, quarantine policies, and even recent negative test results into an integrate metric of "risk to others". The socially optimal policy is to use all available information to target restrictions to where they will do the most good, and restore as much

social activity as possible while still controlling transmission.

Chapter 4

Optimization of Test Timing and Quarantine Duration

4.1 Background

This section describes two methods to choose quarantine duration and test timing after exposure to COVID-19. The first (Section 4.2) aims to compute the optimal fixed-duration quarantine policy. In this case, we aim to choose the best test timing and quarantine duration following a single exposure on a given day where the risk of infection is unknown. This problem is posed as a constrained optimization where we attempt to minimize quarantine duration without increasing expected transmissions compared to a default policy.

The second method (Section 4.3) considers the case where exposures are measured over one or more days, each with an estimated risk of infection conditional on not already being infected. The decision of quarantine and test timing is posed as an optimization problem with a weighted sum of expected transmissions and quarantine duration as the objective function. This optimization problem is then solved using two-stage stochastic programming with test scheduling as the first stage and the quarantine decision for the current day as the second stage (depending on symptoms and test results).

4.2 Fixed Quarantine Policy

4.2.1 Overview

We calculate the optimal timing of PCR testing following exposure to SARS-CoV-2. We minimize the length of post-exposure quarantine, constrained not to increase expected transmission above that of a reference policy of quarantine without testing. Given equally short quarantine, we minimize expected transmission. Unlike previous optimizations, our calculations account for the shared dependence of test sensitivity and infectiousness on the individual's incubation period. To achieve this, we infer time-dependent test sensitivity directly from longitudinal test data, rather than indirectly from viral load trajectories. The absolute degree to which a quarantine recommendation induces behavior change does not affect optimal policy, which instead depends only on the ratio of behavior change magnitudes following notification of exposure (quarantine) vs. a positive test (isolation). Given a ratio inferred from U.K. survey data, testing on day 4 or 5, and then ending quarantine after a negative result comes back 1-2 days later, is sufficient to keep onward transmission lower than for a 10 day quarantine without testing. The use of a second PCR test allows quarantine to be shortened by one more day. With zero behavior change induced by quarantine, day 4 is the optimal date for a single PCR test, or days 3 and 4 for two PCR tests, earlier than the CDC's current recommendation for one test no earlier than day 5. With more recent variants having shorter incubation periods than the variants studied in our test sensitivity data, both testing and quarantine release shift correspondingly earlier.

4.2.2 Significance

Risk analysis that minimizes both infectious disease transmission and quarantine days, while making optimal use of testing, can reduce total harm from outbreaks. This work directly quantifies how the false negative rate for a SARS-CoV-2 test depends on its timing, develops new methods to account for incubation time as a confounding variable within mathematical optimizations, and uses these new tools to make practice-changing recommendations.

4.2.3 Introduction

Some public health agencies, including the Centers for Disease Control and Prevention (CDC), recommend PCR testing after exposure to SARS-CoV-2 [95]. While other agen-

cies instead recommend serial testing with rapid antigen tests, as a means to avoid quarantine, Omicron infections can be undetectable by antigen test for the first several days of significant viral load [96], and variants that escape detection by antigen tests have been observed [97]. Returning to PCR testing therefore also remains an important contingency for jurisdictions currently using antigen tests during or in lieu of quarantine.

Long quarantines are required to avoid all post-quarantine transmission with high confidence [98], because the distribution of incubation periods for most infectious diseases has a lognormal shape [99], meaning that a small but not non-negligible fraction of individuals take a long time before they develop symptoms. However, quarantine causes financial, logistical, and psychological [100] harms, and so tests can be used to shorten or even eliminate quarantine. Shorter quarantine recommendations might also increase compliance [101], which has been low from early in the COVID-19 pandemic [102], and falling [103].

Here we consider optimal policy for quarantine length and testing during quarantine, following exposure to SARS-CoV-2. The CDC initially recommended a 14 day quarantine, and then reduced it to 10 days in the absence of testing, or 7 days if a negative test is obtained within 48 hours of the end of quarantine [104]. The CDC later changed this to recommending an exit test on day 5 for individuals not up to date with vaccinations. The World Health Organization continues to recommend 14 days quarantine [105]. We calculate the shortest quarantine that, in combination with an optimally timed RT-PCR test that comes back negative, will lead to an equivalent number of onward transmission events as a 10 day or 14 day quarantine without testing. We also explore the benefits of a second RT-PCR test during quarantine.

The objective of the CDC recommendation for testing within the last 48 hours of a 7 day quarantine was to minimize post-quarantine transmission [106]. This choice of objective implicitly assumes 100% compliance with quarantine, by ignoring the benefits of shifting contacts earlier from quarantine to isolation. Other work has assumed the same [107]. While a positive test result does not guarantee total compliance with isolation recommendations, a recommendation to isolate following a positive test is nevertheless likely to stem transmission better than a recommendation to quarantine after a mere exposure. This is supported by Norwegian data in which compliance on at least one day with quarantine/isolation without symptoms was only 28%, but rose to 71% with symptoms [108]. Similarly, in the U.K., compliance with quarantine on all recommended days was only 10.9%, rising to a substantially higher if still low value of 18.2% for isolation [109]. Here we optimize for testing regimes that will minimize total transmission, assuming a transition from quarantine to isolation following either the onset of symptoms or a positive test result. This is different from assuming perfect compliance and minimizing only the post-quarantine transmission component. We also consider the potential for a negative

test result to reduce compliance with remaining quarantine.

Test sensitivity depends on the day that test is administered, but there has been little data to inform this relationship. CDC models [110] rely on the model of Clifford et al. [111] that is based on data from Kucirka et al. [112]. But if quarantined individuals move to isolation as presumptive COVID-19 cases upon the onset of any symptom, then the test sensitivity that matters is that which applies pre-symptomatically, for which Kucirka et al. had data only from a single patient (from [113]), who never tested positive by nasopharyngeal swab on any day (only by endotracheal aspirate). Extrapolation from such limited data is inappropriate. Apparent inference of presymptomatic test sensitivity therefore primarily reflects model assumptions, with significant bias toward underestimating presymptomatic sensitivity. More recent work has abandoned this approach to instead simulate viral load trajectories, and make assumptions about test sensitivity as a function of viral load [114, 115]. This is more straightforward for rapid antigen tests than for PCR tests. Here we instead use two longitudinal datasets to obtain the first substantive inferences of presymptomatic test sensitivity: the Big 12 conference subset of the larger National Collegiate Athletic Association (NCAA) quarantine dataset collected by the COVID-19 Collegiate Athlete Testing Group and described by Atherstone et al. [116] but not made public, together with the publicly available dataset of Hellewell et al. [117].

Careful treatment of the non-independence of symptom onset timing, test timing, and time-dependent infectiousness is required to make accurate recommendations. Specifically, the probability of infection from a given exposure, as assessed x days after that exposure, falls when it is made conditional on the lack of symptoms until day x [118]. The probability of infection is similarly conditional on a negative result for a test administered on day t. However, these two conditionalities are interdependent; a test on day 5 might yield little additional information about infection probability on day 10, because conditional on lack of symptoms on day 10, the incubation period if infected was likely to be sufficiently long such that a day 5 test would have very low sensitivity. The CDC model assumes that both infectiousness and test sensitivity peak 5 days after exposure, independent of incubation period, which they assume follows a lognormal distribution [106]. But analysis of pairs with known transmission timing rejected this hypothesis in favor of peak infectiousness being closely associated with date relative to symptom onset, not date relative to exposure [85]. Here we rigorously treat the joint dependence of the timing of symptom onset, test sensitivity, and infectiousness.

Optimal schedules for testing and quarantine may depend on the estimated degree of compliance in different situations, the distribution of incubation periods, the timing of infectiousness, the fraction of cases that are asymptomatic, and the length of delay between testing and receiving the test results. We explore these parameters, in particular

when they vary or are poorly characterized, in order to recommend robust optimal testing and quarantine regimes.

4.2.4 Results

Time-Dependent Test Sensitivity

We infer RT-PCR test sensitivity as a function of both infection date and symptom onset date. Figures 4.1A and 4.1B show curves for different incubation periods, with the thick black line showing the weighted average across the distribution of incubation periods. Figure 4.1C illustrates the 90% credibility interval associated with our inference, taking as an example the case of the median incubation period of 5 days. This shows that we have sufficient data for reliable inference.

Optimal Test Timing

We compare to a reference policy of $\pi_{\{10,\emptyset\}}$: quarantine until 10 days after exposure, without testing unless symptoms develop, and if symptoms develop (either during or after quarantine), enter isolation (and request a test). By making use of testing, we find the shortest quarantine length that does not increase the total number of expected transmissions (not just post-quarantine transmissions). If there are multiple test regimes that allow equally short quarantine, we select the one with the least onward transmission.

In Appendix B.1, we show that optimal policy does not depend on the absolute level of quarantine or isolation compliance, but only on their ratio C_Q. Note that "compliance" here describes the reduction in transmission due to any behavior change ranging from staying at home to mask wearing, whether or not those behavior changes are recommended as policy.

With $C_Q = 1$ (i.e. quarantine compliance just as high as isolation compliance), $C_N = 0$ (i.e. no drop in quarantine compliance after a negative test result), and test turnaround $T_{delay} = 1.5$ (assuming testing in the middle of the day yields results in time to act on the day after next), the regime that best matches the post-quarantine transmission from 10 day no-test quarantine is 7 day quarantine with an exit test on day 6 (Figure 4.2A). This means that our improved estimates of time-dependent test sensitivity, and more rigorous treatment of conditional probability, do not cause the resulting optimization to deviate from that derived from CDC models. However, our recommendations are sensitive to assumptions about compliance. For the more realistic case of $C_Q = 0.7$, we instead recommend 6 day quarantine with an exit test on day 5 (Figure 4.2B).

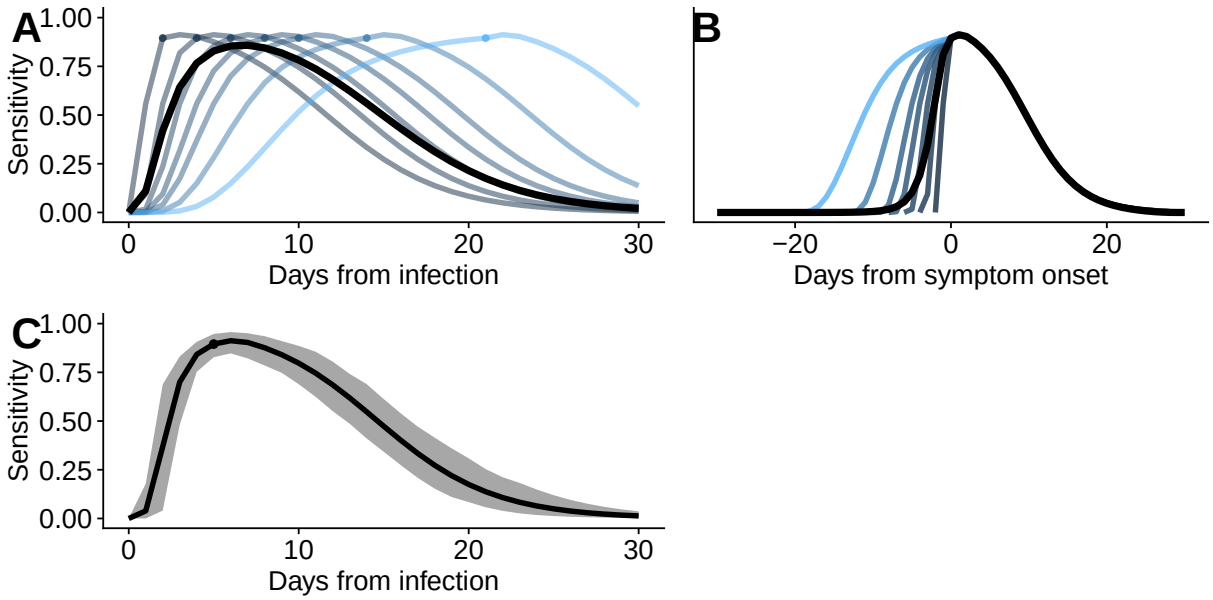


Figure 4.1: Inferred time-dependent test sensitivity. In A) and B), lines are shown for incubation periods of 2, 4, 6, 8, 10, 14, and 21 days, and the black line represents the weighted average over incubation periods. Dots indicate the day of symptom onset in A). Lines start on the day of infection in B). Line in C) shows the mean sensitivity across the range of parameters values sampled during our Monte Carlo approach, given the median incubation period (5 days) and the shaded area represents the 90% credibility interval for the posterior distribution produced by our Monte Carlo inference method.

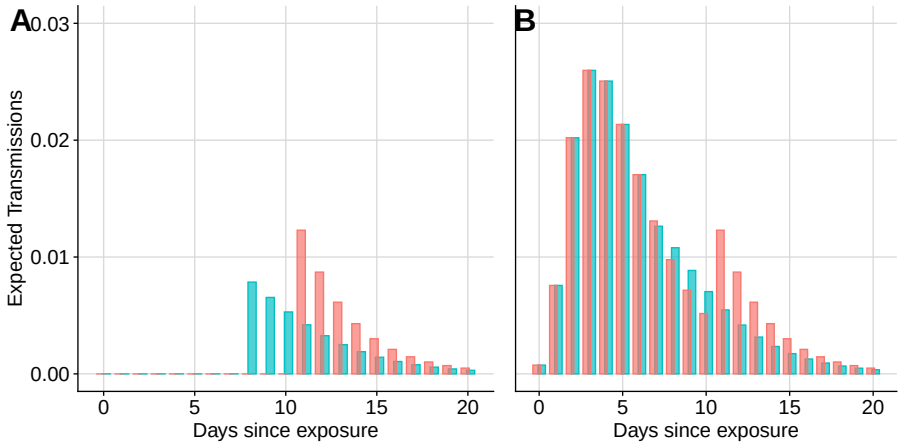


Figure 4.2: Expected transmissions per day for default 10-day no-test quarantine policy (red), and shortened quarantine with test (blue). A) Isolation and quarantine compliance are both 100%, quarantine is shortened to 7 days with a test on day 6. Test turnaround time $T_{delay} = 1.5$ makes this an exit test. B) Isolation compliance is 100%, quarantine compliance is 70%, and quarantine is shortened to 6 days with an exit test on day 5. In both cases, shortened quarantine regime is chosen such that it leads to the same or fewer expected transmissions. $C_N = 0.5$. Results are shown for $R_t = 1.0$; y-axis values can be proportionately scaled for other values of R_t.

We explore C_Q more broadly in Figure 4.3. When quarantine is taken nearly as seriously as isolation (high C_Q), as in Figure 4.2A, the risk reduction provided by a single test can shorten quarantine by 3 days (Figure 4.3A, right). For lower quarantine compliance, as in Figure 4.2B, we recommend earlier testing, again with quarantine ending as soon as negative test results are received (Figure 4.3A, middle). But when quarantine compliance is extremely low, the risk reduction from a single test on day 4 can completely replace any quarantine (Figure 4.3A, left).

Importantly, in the absence of quarantine, the optimal date for a single test is 4 days after exposure, in contrast to current CDC recommendations that contacts who are up to date with recommended vaccines (who are not required to quarantine, only to mask) should test 5 days after exposure [119]. The day 4 testing recommendation is the result of a trade-off between earlier testing ensuring results before the period of highest transmission risk (around day 5), and later testing reducing the risk of a false negative due to lower test sensitivity. A negative test on day 4 should obviously be followed by isolation and a second test should symptoms develop.

To make a practical policy recommendation based on Figure 4.3, an estimate of C_Q (quarantine compliance relative to isolation compliance) is needed. Data on self-reported compliance with isolation and quarantine was collected by “” in the UK between “” and “”. For quarantine, 11% reported full compliance, 54% reported partial compliance, and 35% reported no behaviour change. For isolation, compliance was slightly higher, with 40% reporting full compliance, 40% partial, and 30% none. To use these numbers to estimate C_Q in units of expected transmissions, we we assume that full compliance eliminates transmission, that no compliance leaves the number of transmissions unchanged, and we explore the effect of partial compliance as 25%, 50%, and 75% in scenarios 1, 2, and 3 respectively. This leads to a reduction in transmissions due to isolation of 47.5%, 55%, and 62.5%, and a reduction due to quarantine of 24.5%, 38%, and 51.5% in the three scenarios. These yield C_Q values of $24.5/47.5 = 0.516$, $38/55 = 0.691$, and $51.5/62.5 = 0.824$ in scenarios 1, 2, and 3 respectively, which are shown by the three vertical lines in Figure 4.3. We note that Quilty et al.’s [114] baseline scenario of values for both compliance rates maps to our $C_Q = 0.75$, within the higher end of the range we consider.

Depending on which scenario is most accurate, we recommend 5-6 days of quarantine with corresponding exit test on days 4-5, a narrow range reflecting limited variation among interpretations of the same U.K. data. In a more quarantine-compliant population than the U.K., quarantine might be 7 days, while in a less quarantine-compliant population, no quarantine might be recommended. For contacts who are exempt from quarantine, e.g. due to vaccination, we recommend testing on day 4.

However, optimal policy will depend on SARS-CoV-2 variant characteristics, e.g. the Delta variant has higher transmission risk and may have a shorter incubation period than previous strains [120]. The Omicron variant has also been suggested to have a shorter latent period [121]. While our recommended policies are independent of absolute transmission risk R_t (because risk would increase proportionally in both the test and default policy), they are not independent of the incubation period of the virus. Dashed lines in Figure 4.3B show the recommended policy with a mean incubation period of 4.4 days instead of 5.4 days. With the shortened incubation period, the recommended quarantine duration and test day are generally shortened by one day.

When we use the World Health Organization reference policy of $\pi_{\{14,\emptyset\}}$ (Figure 4.3C, dashed lines) instead of the CDC reference policy of $\pi_{\{10,\emptyset\}}$ (Figure 4.3C, solid lines), quarantine remains equally short for low quarantine compliance, but is lengthened as quarantine compliance approaches isolation compliance. However, even for $C_Q = 1$, quarantine is not lengthened by the full 4 days, but instead rises from 7 days to 10 days.

If we increase the PCR test turnaround time from $T_{delay} = 1.5$ to $T_{delay} = 2.5$, this rarely changes the date of testing (Figure 4.3D, red dashed vs. solid lines), but instead adds one day to quarantine (Figure 4.3D, blue lines). This supports the CDC's shift in framing, from recommending that quarantine end on a certain day subject to a negative test in the last 48 hours [106], to specifying date of testing, and allow quarantine to end as soon as a negative test result is received. As well as being more robustly optimal, framing advice in this way is also simpler than inviting individuals to research current test turnaround times in order to decide when within a 48 hour period to get tested. Our results points to days 4-5 as suitable dates for testing for early variants, or earlier for currently circulating strains with shorter incubation periods.

While testing can increase compliance when test results come back positive, it can also decrease compliance when test results come back negative. Recommendations in Figure 4.3 are based on $C_N = 0.5$, i.e. halving quarantine compliance following receipt of a negative test, which results in policies that recommend quarantine exit shortly after the receipt of test results, and thus little opportunity for false negative results to increase transmission. However, results are the same for $C_N = 0$.

While the tightness of our credibility intervals in Figure 4.3A is reassuring, the data from which we infer time-dependent test sensitivity might be inaccurate in ways not captured by our Monte Carlo simulations. We therefore explore the scope for new data to change our recommendations. In Figure 4.4, we adjust both the height and width of our inferred time-dependent test sensitivity curve. Shorter quarantine, enabled by earlier testing, is made possible both by tests with higher peak sensitivity and by tests that remain sensitive

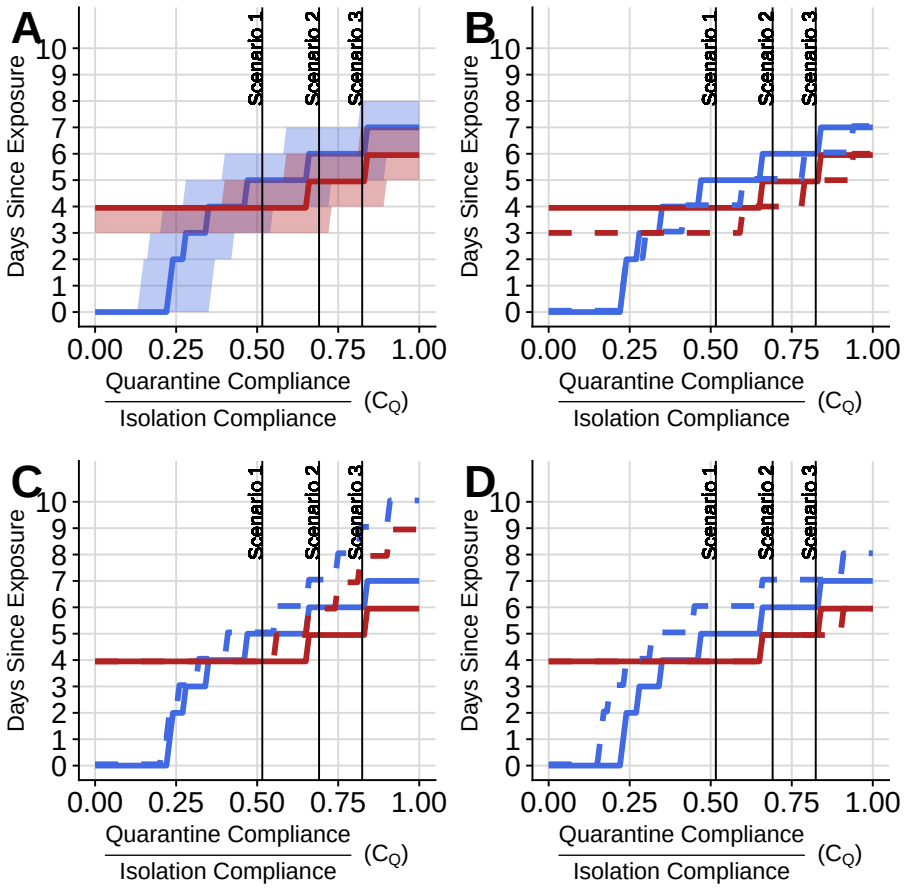


Figure 4.3: Shorter quarantine and earlier testing are recommended when quarantine compliance is much less than isolation compliance. Quarantine duration (blue) and test day (red) depend in all figure panels on the ratio C_Q of quarantine to isolation compliance. Vertical lines correspond to three different interpretations of survey results on the compliance ratio, as described in the main text. $C_N = 0.5$ throughout, i.e. quarantine compliance is reduced by half following a negative test result, but results are identical for $C_N = 0$. A) Shaded areas reflect 90% credibility intervals from posterior distributions produced by our Monte Carlo runs. This shows that we have sufficient data on time-dependent test sensitivity to make recommendations that are precise to +/- one day. B) A variant such as Delta with an estimated 4.4 day mean incubation period [4] (dashed lines) calls for testing and quarantine release one day earlier than for the original SARS-CoV-2 strain with a mean incubation period of 5.4 days (solid lines). C) Benchmarking against a 14-day no-test quarantine (dashed lines) rather than a a 10-day no-test quarantine (solid lines) results in less than the 4 days between the recommended with-test policies. D) Test delays of 1.5 days (solid lines) and 2.5 days (dashed lines) result in the same recommended test day, with slower test return.

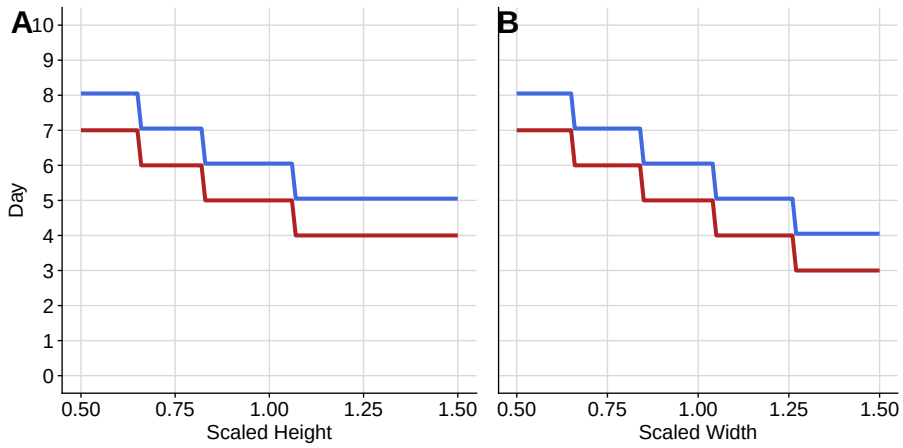


Figure 4.4: More sensitive tests across larger date ranges allow shorter quarantine with earlier testing. Optimal quarantine duration (blue) and test date (red) with $C_N = 0.5$ and $C_Q = 0.691$ (matching scenario 2) for modified test sensitivity curves. A) Test sensitivity function is scaled vertically (with a ceiling of 1.0). B) Test sensitivity function is scaled horizontally around the symptom onset date (higher values of 'scaled width' lead to a longer period of high sensitivity).

over a greater range of dates. However, results are not overly sensitive to the exact height or width of our inferred time-dependent test sensitivity curves. Note that the use of a single antigen test, rather than a single RT-PCR test, likely corresponds to narrower and lower sensitivity curve, making it less effective at shortening quarantine, although shorter T_{delay} can help compensate.

Antigen tests provide the possibility of daily rapid testing [115]. But even with PCR testing, multiple tests during quarantine are also a viable option today. Figure 4.5A shows the optimal timing of quarantine with two tests. Testing twice allows quarantine to be one day shorter than when testing once. One test is timed to receive negative results for quarantine release, and a second test is generally timed for the following day. Two tests are almost never recommended for the same day (with the exception of very high C_Q in Figure 4.5A). We recommend that the unquarantined individuals are tested either on day 4, or on days 3 and 4 following exposure, or earlier for strains with shorter incubation periods. Again, this is earlier than current CDC advice for unquarantined individuals to wait at least 5 days after exposure before testing.

Our optimization procedure yields shorter quarantine with earlier testing than current

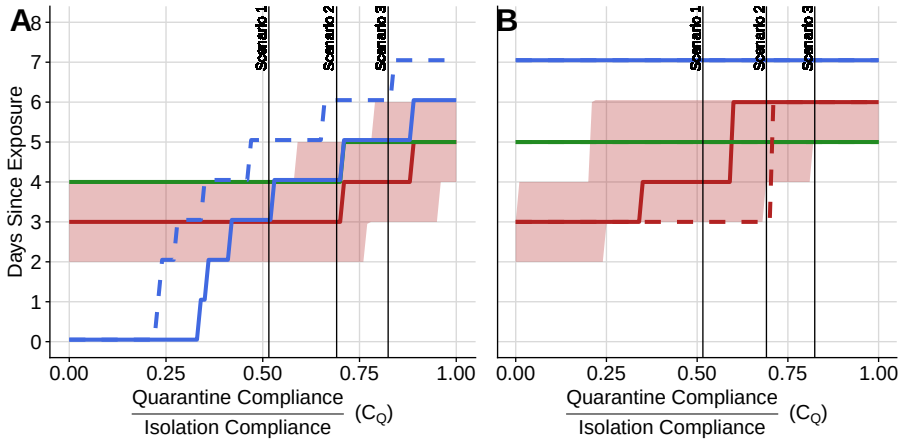


Figure 4.5: Recommended quarantine length and dates of 2 tests . Quarantine duration in blue, two tests in red and green, shaded area reflects a 90% credibility interval on the timing of the earlier test on the basis of the posterior distribution produced by our Monte Carlo runs. A) With all choices free. Dashed blue line shows longer quarantine duration with only one test. B) With a fixed quarantine length of 7 days and test on day 5, plus an additional test with optimized date. Dashed line indicates $C_N = 0$, solid line for $C_N = 0.5$.

CDC suggestions that post-exposure quarantine ends upon receipt of a day 5 test [106]. Given that public health authorities in the U.S. may feel constrained by CDC suggestions, we therefore optimize for the timing of one test while holding the other constant on day 5, with $T_{delay} = 2.5$ allowing quarantine end after 7 days. When quarantine compliance relative to isolation compliance is low, the first test is recommended on day 3 to reduce the risk of transmission during quarantine. As quarantine compliance increases, the additional test is allocated to day 6 to minimize transmissions after release from quarantine.

In Figure 4.5B, with quarantine duration held fixed, we see for the first time how the effect of negative tests on compliance can change recommendations. When compliance is insensitive to negative tests ($C_N = 0$, dashed line), earlier testing becomes possible. The reason $C_N = 0$ and $C_N = 0.5$ results were indistinguishable in earlier figures is that quarantine length was not held constant. When earlier test dates enable earlier release, we see no difference between recommendations based on $C_N = 0$ versus $C_N = 0.5$.

4.2.5 Discussion

We estimated time-dependent, presymptomatic RT-PCR sensitivity curves from longitudinal data, and used them to optimize testing and quarantine regimes relative to a quarantine-without-testing benchmark policy, while accounting for the joint dependence of test sensitivity, symptom onset, and infectiousness on incubation time. A key mathematical finding of our work is that the optimal choice of test/quarantine regime depends not on the absolute degree of to which notification of exposure prompts transmission-reducing behavior change ("quarantine compliance"), but on the ratio between quarantine compliance following contact tracing and isolation compliance following positive test or symptoms. Our calculations recommend shorter quarantine and earlier SARS-CoV-2 testing than the CDC does, especially when quarantine compliance is low relative to isolation compliance, and especially for variants such as Delta and Omicron that seem to have shorter incubation periods. We recommend PCR testing 4-5 days after exposure, or earlier for strains with shorter incubation periods, with quarantine ending upon receipt of a negative result. Use of a second RT-PCR test can shorten quarantine by one more day. Individuals exempt from or unlikely to comply with quarantine are best tested on day 4 following exposure, or days 3 and 4 in case of two tests, or earlier for strains with shorter incubation periods.

Our work overcomes three key limitations of previous studies that have considered the question of how to optimize RT-PCR test timing within quarantine [122, 123, 124]. First, Wells et al. [122] and Johansson et al. [124] focus exclusively on post-quarantine transmission, ignoring the reduction in transmission achieved by moving individuals from

quarantine to isolation. This focus is appropriate for enforced border quarantine scenarios such as entry into an oil well [122], or to a location that has achieved local elimination of SARS-CoV-2, but not for post-exposure voluntary quarantine within a population.

Interestingly, despite including during-quarantine transmission, our model still recommends testing just in time for quarantine exit (in cases where quarantine is recommended at all). We note that our treatment of the shared dependence of infectiousness and test sensitivity on incubation period is necessary to capture the degree to which late testing gives more significant information about post-quarantine transmission.

A second limitation with previous studies is a lack of data on presymptomatic test sensitivity, which we directly rectify. As discussed in the Introduction, studies such as Johansson et al. [124] that rely on data compiled by Kucirka et al. [125] effectively use only a single presymptomatic patient, and thus primarily reflect model assumptions, with significant bias toward underestimating presymptomatic sensitivity. As an alternative, Wells et al. [122] fit a relationship between post-symptomatic transmission and test sensitivity, and then used it to infer pre-symptomatic test sensitivity from pre-symptomatic transmission. The problem with this is that infectiousness and transmission cannot be assumed proportional in this way; post-symptomatic transmission is reduced by behavior changes to a level lower than that expected from infectiousness. This method is therefore also expected to underestimate presymptomatic test sensitivity. The same applies to the method used by Johansson et al. [124] to model antigen test sensitivity, namely assuming directly proportionality with transmission, which was assumed proportional to infectiousness whether or not symptoms were present.

A third limitation with some past work [122, 124] is the assumption that test sensitivity depends only on time since infection, not time of symptom onset. The consequence of the neglect of shared dependence of infectiousness and sensitivity on incubation period is to overstate the amount of information about late symptom onset infection obtained from early tests, and thus to bias the recommendations toward testing on the inferred day of maximum test sensitivity, given that it is inappropriately assumed to give the same amount of information independently of incubation period. Johansson et al. [124] also treat infectiousness as independent of incubation period. This will further overstate the amount of information about late infectiousness obtained by an early test.

Given these limitations, the best approach prior to our work, used by Quilty et al. [123] and Ferretti et al. [126] has been to simulate viral load trajectories and from those apply functions for both infectiousness and sensitivity. But the relationship between viral load and RT-PCR sensitivity is far from clear. We more directly infer test sensitivity from longitudinal testing data.

A limitation of our approach is that new strains might have different parameters, which can only be estimated from contact tracing efforts such as the data used here. Unfortunately, contact tracing is rarely conducted with the same intensity as occurred earlier in the COVID-19 pandemic. It is important both to collect and to rapidly share data that informs key parameters for decision-making [127]. This includes de-identified line list longitudinal testing data on contacts, together with the timing of their exposure relative to symptom onset in the primary case.

Another limitation of our study is that while we quantitatively modeled isolation and quarantine compliance, we assumed full compliance with testing. This is motivated by the assumption that those who do not comply with testing are sufficiently unlikely to comply with quarantine such that precise recommendations for this population are irrelevant. However, it is still useful to consider strategies that make testing as convenient as possible, in order to increase compliance.

Here we used a fixed quarantine duration of 10 or 14 days as a reference policy, and found the best policy that leads to no more transmission than that. However the socially optimal policy would effectively make the reference policy depend on case prevalence [128], and on probability of infection as estimated from exposure dose [129] and immunity status. There are strong ethical arguments for minimizing quarantine recommendations by targeting them to where they will stem the most transmission [130]. Note that when taken from this perspective of risk analysis, there is likely more information in exposure history than in travel history. Outside of zero-COVID countries, test regimes for travel are perhaps best targeted at lowering the probability of travel-associated superspreading events, rather than minimizing the probability of introduction to a population where SARS-CoV-2 is already established.

Our recommendations for earlier testing and quarantine end following exposure are conservative for several reasons. Firstly, new strains seem to have shorter incubation periods. Secondly, we did not in this study include the benefits of earlier case detection for the purpose of tracing more contacts, either by forward contact tracing from the focal case or by backward contact tracing to detect superspreader events. Including these benefits might shift recommended testing dates slightly earlier.

4.2.6 Methods

Incubation period

The day of infection is designated day 0. The time until peak viral load, which corresponds for symptomatic patients to the day of symptom onset, follows a lognormal distribution. Following [131], we assume

$$t_{symp} \sim \text{LogNormal}(\mu = 1.58, \sigma = 0.47) \tag{4.1}$$

We convert this into a discrete distribution through integration to get a random variable for the day of peak viral load / symptom onset, D_{symp}. For convenience, we truncate at 30 days and renormalize.

Time-dependent test sensitivity

To estimate test sensitivity as a function of day of test, we consider screenings from employees of care facilities who were monitored longitudinally [117] and similar data from college and university attendees [116]. Each participant included into this analysis was tested for SARS-CoV-2 intermittently using RT-PCR tests and had at least one positive test result. Along with 334 positive and negative test results from 97 individuals, presence and absence of COVID-19 symptoms were recorded.

We estimate the sensitivity of a PCR test from the observed test results of infected individuals using Bayesian model. The main data is dated test results: $y_i = 1$ for a positive test and $y_i = 0$ for a negative test result, $i \in [1, 334]$. We fit Bernoulli distribution to logit-transformed chance parameter which depends on the timing of infection:

$$y_i \sim \text{Bernoulli}(\text{logit}^{-1}(f(t))). \tag{4.2}$$

We examined different parametrisation of $f(t)$, and report the one which explicitly infers the time of symptom onset, the infection time, the time when the sensitivity peaks, and the coefficients of the curve itself. Inverse logit-transformed $f(t)$ is the sensitivity of PCR test, and shows the probability of detecting an infection in an infected individual.

Expected daily transmissions from presymptomatic and symptomatic cases

$T(x)$ is a random variable giving the number of transmissions on day x. We take its expectation from Ferretti et al. [85] as a function of the day that symptoms begin, z:

$$E[T(x)|D_{symp} = z, Asymp = False] \propto \begin{cases} \frac{e^{-((x-z)*\tau/z-\mu)/\sigma}}{(1+e^{-((x-z)*\tau/z-\mu)/\sigma})^{\alpha+1}}, & \text{if x <z} \\ \frac{e^{-((x-z)-\mu)/\sigma}}{(1+e^{-((x-z)-\mu)/\sigma})^{\alpha+1}}, & \text{if x} \geq z \end{cases} \tag{4.3}$$

where $\tau = 5.42, \mu = -4, \sigma = 1.85$, and $\alpha = 5.85$.

We set the proportionality constant so that $R_t = \sum_x \sum_z P(D_{symp} = z) * E[T(x)|D_{symp} = z] = 1.0$, merely to make values of T(x) more concrete. Not that the sum of expected transmissions R_t will be lower for vaccinated and/or recovered individuals than for immunologically naive individuals. If the individual is in isolation or quarantine, expected transmission is dampened by a factor given by the level of compliance, C_I or $C_Q C_I$, respectively, as discussed below.

Asymptomatic cases

A fraction of cases p_{asymp} never develop noticeable symptoms. We assume that their viral load profile over time nevertheless follows the same shape as symptomatic cases, and D_{symp} represents the time until peak infectiousness. Because this does not trigger symptoms and thus isolation, we also need to model infectiousness after this time. The estimates of Ferretti et al. are not suitable for this, because they are for post-symptomatic transmission that is affected by behavior modification as a consequence of symptoms. Instead, we use the TCID50 results of Bullard et al. as a more direct indication of infectiousness. We assume that peak infectiousness continues at a plateau for days 0, 1, and 2 relative to peak viral load, and that infectiousness then declines geometrically by a factor of 0.631 per day.

$$E[T(x)|D_{symp} = z, Asymp = True] \propto \begin{cases} \frac{e^{-((x-z)*\tau/z-\mu)/\sigma}}{(1+e^{-((x-z)*\tau/z-\mu)/\sigma})^{\alpha+1}}, & \text{if x <z} \\ \frac{e^{\mu/\sigma}}{(1+e^{\mu/\sigma})^{\alpha+1}}, & \text{if x} \geq z \text{ and } x \leq z+2 \\ \frac{e^{\mu/\sigma}}{(1+e^{\mu/\sigma})^{\alpha+1}} * 0.631^{x-z-2}, & \text{if x >z + 2} \end{cases} \tag{4.4}$$

We normalize total expected transmissions from asymptomatic individuals to equal the same value of $R_0 = 2.5$ as we do for symptomatic individuals. This corresponds to a lower peak infectiousness and a higher proportion of onward transmission occurring post-peak.

Quarantine Policy

A quarantine policy $\pi_{\{m,t\}}$ consists of a recommended test date, t, and quarantine length, m. Infection occurs on day 0, and a 10-day quarantine means quarantine on days 1 through 10 inclusive, with release from quarantine on day 11. The parameter T_{delay} specifies the number of days after testing that results are returned. We assume that symptoms develop in the morning, and that symptomatic individuals immediately enter isolation. We generally set $T_{delay} = 1.5$, so that a test is taken midway through the day, and positive test results are received at the end of the next day or early morning the day after, again triggering isolation from the beginning of the day. E.g. when a test is recommended for day 6 of a 7 day quarantine, a positive test result would trigger isolation beginning at the start of day 8, smoothly taking over from the quarantine that just ended. If no testing is used in a given policy, t is set to $\emptyset$. $Test(t)$ is a random Boolean variable giving the boolean test result, r, on day t, with probability dependent on time since infection and symptom onset date, as described in Section 4.2.6. $Asymp$ is a random variable, set to true for asymptomatic cases. The number of expected transmissions on a given day can be computed based on the quarantine policy, infection date, symptom onset date, whether the symptoms are noticeable, and test result. In the absence of isolation or quarantine, the expected number of transmissions on a particular day is given by Equation 4.3 for symptomatic cases or Equation 4.4 for asymptomatic cases. Equation 4.5 gives the daily transmission adjustment factor based on quarantine/isolation status (and indirectly on symptom status, test results, and quarantine policy).

$$
\begin{aligned}
&h(x, \pi_{\{m,t\}}, Test(t) = r, D_{symp} = z, Asymp = a) = \\
&\begin{cases}
1 - C_I, & \text{if } !a \text{ and } x > z \text{ (noticeably symptomatic)} \\
1 - C_I, & \text{if } r \text{ and } x > t + T_{delay} \text{ (tested positive)} \\
1 - C_I/2 - C_Q * C_I/2, & \text{if } !a \text{ and } x == z \text{ (first day of symptoms)} \\
1 - C_I/2 - C_Q * C_I/2, & \text{if } r \text{ and } x == t + T_{delay} \text{ (day received positive test result)} \\
1 - C_Q * C_I * (1 - C_N), & \text{if } !r \text{ and } x > t + T_{delay} \text{ and } x < m \text{ (negative test + in quarantine)} \\
1 - C_Q * C_I, & \text{else if } x < m \text{ (in quarantine)} \\
1, & \text{otherwise (released from quarantine and no symptoms)}
\end{cases}
\end{aligned}
\tag{4.5}
$$

Under a two-test policy, behavior changes after the first test, there is no second test if the first test is positive, and behavior following first a negative then a positive test is the same as for any other positive test.

Transmission depends on levels of compliance (i.e. degree to which social contact is reduced) captured by the C terms above. We assume that quarantine compliance is a fraction C_Q of isolation compliance C_I. In Appendix B.1, we show that because $1 - h(x, \pi_{\{m,t\}}, r, z, a)$ is always proportional to C_I, we can assume, without loss of generality, 100% compliance with isolation. To capture the effect of a negative test result on quarantine compliance, we use a linear expression in which $C_N = 0$ corresponds to no effect and $C_N = 1$ corresponds to the complete collapse of quarantine compliance.

Across symptomatic and asymptomatic cases, all possible days of symptom onset / peak infectiousness, and either positive or negative test results for day t, we obtain the expected transmission g_x on day x:

$$\begin{aligned} g_x(\pi_{\{m,t\}}) = \sum_{z \in 0:30} \sum_{r \in 0,1} \sum_{a \in 0,1} \{ & P(D_{symp} = z, Test(t) = r, Asymp = a) * \\ & h(x, \pi_{\{m,t\}}, Test(t) = r, D_{symp} = z, Asymp = a) * \\ & E[T(x) | D_{symp} = z, Asymp = a] \} \end{aligned} \tag{4.6}$$

The first term gives the probability of a scenario, the second gives the degree to which transmission is reduced in that scenario, and the third term from Equations 4.3 and 4.4 gives the expected number of transmissions in the absence of quarantine or isolation. Equation 4.6 can be easily expanded to the case of two tests on days t_1 and t_2 with results r_1 and r_2.

The joint probability of peak infectiousness date, test result for day t, and asymptomatic status is:

$$\begin{aligned} & P(D_{symp} = z, Test(t) = r, Asymp = a) \\ & = P(D_{symp} = z) * P(Test(t) = r | D_{symp} = z) * P(Asymp = a) \end{aligned} \tag{4.7}$$

Under a two-test regime, the middle factor is repeated for each of the two tests.

For a given quarantine policy, the total number of expected transmissions, summed over the 30 days following exposure, is given by $g(\pi)$:

$$g(\pi_{\{m,t\}}) = \sum_{x \in 0:30} g_x(\pi_{\{m,t\}}) \} \tag{4.8}$$

4.3 Dynamic Quarantine Policy

4.3.1 Introduction

In section 4.2 we used conditional distributions for test sensitivity, expected transmissions, and incubation period to choose a static quarantine and testing policy (after an exposure on a known day) that minimized the quarantine duration while holding expected transmissions below a specified value. In this section we will solve a similar problem - how to choose quarantine duration and test timing following exposures on one or more days each with an estimated transmission risk (conditional on not already being infected). Here we will minimize a weighted sum of expected transmissions and quarantine days to be consistent with the objective function introduced in 3. To somewhat simplify the problem, we will assume that quarantine and isolation are equally effective.

An exposure notification app is well situated to automate this inference and provide updates to quarantine recommendations immediately as new information is entered. The risk threshold (from section 3 describes a preference ratio between additional transmissions and an additional day of quarantine. This preference can be used to compute the expected value of a scenario in units of quarantine days. Using two-stage stochastic programming, an optimal test date can also be recommended to the user. Combining both of these optimizations is numerically shown to reduce the cost of quarantine by 30% in typical scenarios.

Section 4.3.2 builds a computation that takes probability of infection on given days (conditional on not already being infected) and a day up to which symptoms have not been noticed, and outputs the expected transmissions on future dates. This expected transmissions calculation is then extended to incorporate dates of negative tests using a timing-dependent test sensitivity function. Using the expected daily transmissions function, Section 4.3.3 shows how to build a quarantine recommendation system that recommends quarantine on days when the expected transmissions exceed a certain threshold. Finally, Section 4.3.4 shows how to optimally choose test timing and quarantine days to minimize a weighted sum of expected transmissions and quarantine duration by posing it as a two-stage stochastic programming problem.

4.3.2 Conditional Infection Probability

Conditional on Symptoms

Consider the case where someone is known to have been exposed on certain days. This individual receives a notification a few days later warning them of their risk. Using duration, proximity and infectiousness of the index case, the probability of infection from each exposure event can be estimated as described in [118]. The individual is estimated to have been infected with probability $p_{inf}(d)$, for each day of exposure, d. If multiple exposures with risks, p_i, occurred on the same day, the combined risk for that day is calculated as $1 - \prod_i (1 - p_i)$. (This occurs automatically under GAEN v2, in which expected doses, in the form of "Meaningful Exposure Minutes", are summed). The person can be instructed to presume SARS-CoV-2 infection and self-isolate immediately if they are symptomatic or if they become symptomatic. In theory, this eliminates post-symptomatic transmission risk, and only pre-symptomatic and asymptomatic transmission need to be considered to make a decision about quarantine. To do so, it is useful to estimate the expected number of transmissions that would occur on a given day if they did not quarantine.

Our aim is to compute the expected number of transmissions on day x given that no symptoms have been experienced up to and including day w (typically $w = x-1$). This can be done by constructing the joint probability distribution of infection date and symptom onset date, then multiplying by the expected transmission curve to compute the expected risk on a given day.

Permissible values for D_{inf} and D_{symp} are the natural numbers and $\varnothing$ (to represent the scenario where the event does not occur). These permissible values are represented by the set $A = (\mathbb{N} \cup \varnothing)$. In practical implementations, days more than a month away from x can be neglected. The expected value of $T(x)$ given symptom onset date and infection date is given by equation 4.9. Summing over all possible values of D_{inf}, D_{symp}, and $Asymp$ yields the expected value of $T(x)$ in the absence of other information, as shown in equation 4.10. Finally, the expected number of transmissions can be adjusted based on the knowledge that symptoms have not occurred by day w, as shown in equation 4.11.

$$E[T(x)|D_{inf} = y, D_{symp} = z, Asymp = a] = \begin{cases} 0, & \text{if } y = \varnothing \\ T_{symp}(x - y, z - y), & \text{if } a = False \\ T_{asymp}(x - y, z - y), & \text{if } a = True \end{cases} \quad (4.9)$$

$$E[T(x)] = \sum_{i \in A} \sum_{j \in A} \sum_{k \in \{0,1\}} \{E[T(x)|D_{inf} = i, D_{symp} = j, Asymp = k] \\ * P(D_{inf} = i, D_{symp} = j, Asymp = k)\} \tag{4.10}$$

$$E[T(x)|Asymp = True \text{ or } D_{symp} > w)] = \\ \sum_{i \in A} \sum_{j \in A} \sum_{k \in \{0,1\}} \{E[T(x)|D_{inf} = i, D_{symp} = j, Asymp = k] \\ * P(D_{inf} = i, D_{symp} = j, Asymp = k|Asymp = True \text{ or } D_{symp} > w)\} \tag{4.11}$$

In order to compute the conditional expected value in equation 4.11, $P(D_{inf} = y, D_{symp} = z, Asymp = a|Asymp = True \text{ or } D_{symp} > w))$ needs to be evaluated. First, the joint probability $P(D_{inf}, D_{symp}, Symp)$ will be constructed, and then the distribution will be adjusted to handle the truncation of the D_{symp} domain.

The joint probability can be rewritten as $P(D_{inf}, D_{symp}, Asymp) = P(D_{symp}, Asymp|D_{inf}) * P(D_{inf})$. We assume an individual can be infected at most once (within the time period considered - less than a month). Relaxing this assumption is for the time being not tractable. Our logic is based on the idea that each live viral unit is independently capable of initiating infection. When multiple infection events happen, the incubation time is likely to be set by whichever infection event would on its own lead to earlier symptom onset, e.g. as sampled independently from the same lognormal distribution of incubation times, potentially beginning from different infection dates should there be exposure on more than one day. We neglect the complications arising from multiple infection initiation events because our exponential dose-response curve is not grounded in discrete numbers of virus - this kind of Quantitative Microbial Risk Assessment is not yet available for SARS-CoV-2. Indeed, we calculate infection probability not even from dose in arbitrary units but from expected dose, with significant but poorly quantified variation in dose. While this approximation performs acceptably for realistically low infection probabilities, our poor quantification of variance in dose precludes rigorous treatment of multiple independent infection events.

Instead, we assume only a single infection event, with preference given to events associated with earlier exposure. In this case, the variable D_{inf} follows a geometric distribution with different daily infection probabilities, $p_{inf}(d)$, as shown in equation 4.12. In this analysis, symptoms are assumed to only be caused by infection with SARS-CoV-2. Anyone who develops symptoms is of sufficiently high risk that they should self-isolate. Those that develop symptoms for other reasons are excluded from the population under consideration.

The probability of first experiencing symptoms on a given day is conditional on the date of infection, as shown in Equation 4.13. If the infection date is after the symptom date, the probability is 0. Otherwise, the probability is computed using the incubation distribution and scaled by a factor to account for asymptomatic cases.

$$P(D_{inf} = y) = \begin{cases} p_{inf}(y) * \prod_{i=1}^{y-1} (1 - p_{inf}(i)), & y \in \mathbb{N} \\ \prod_{i \in \mathbf{N}} (1 - p_{inf}(i)), & y = \varnothing \end{cases} \tag{4.12}$$

$$P(D_{symp} = z, Asymp = a | D_{inf} = y) = \begin{cases} 1, & \text{if } y = \varnothing \text{ and } z = \varnothing \text{ and } a \\ 0, & \text{if } y = \varnothing \text{ and } (z! = \varnothing \text{ or } !a) \\ 0, & \text{if } y! = \varnothing \text{ and } z < y \\ (1 - p_{asymp}) * t_{symp}(z - y), & \text{if } y! = \varnothing \text{ and } !a \\ p_{asymp} * t_{symp}(z - y), & \text{if } y! = \varnothing \text{ and } a \end{cases} \tag{4.13}$$

The truncated distribution can be evaluated as shown in Equation 4.14, where $I()$ is an indicator function, and $P(D_{symp} \leq w)$ is computed by summing over all allowed scenarios as shown in equation 4.15.

$$\begin{aligned} &P(D_{inf} = y, D_{symp} = z, Symp = s | Asymp = True \text{ or } D_{symp} > w)) = \\ &\frac{P(D_{inf} = y, D_{symp} = z, Symp = s) * I(z > w \text{ or } a)}{1 - P(Asymp = False \text{ and } D_{symp} \leq w)} \end{aligned} \tag{4.14}$$

$$P(Asymp = False \text{ and } D_{symp} \leq w) = \sum_{i \in A} \sum_{j \leq w} [P(D_{inf} = i, D_{symp} = j, Asymp = False)] \tag{4.15}$$

By combining Equations 4.11, 4.14 and 4.15, the expected number of transmissions on day x given no symptoms noticed prior to day w can be computed. Figure 4.6 shows a scenario after two exposure events 10 days apart. The expected transmission risk is computed for each day and adjusted based on recent information about symptoms. Updating with information about lack of symptoms on the current day significantly reduces the expected risk compared to the prior risk curve.

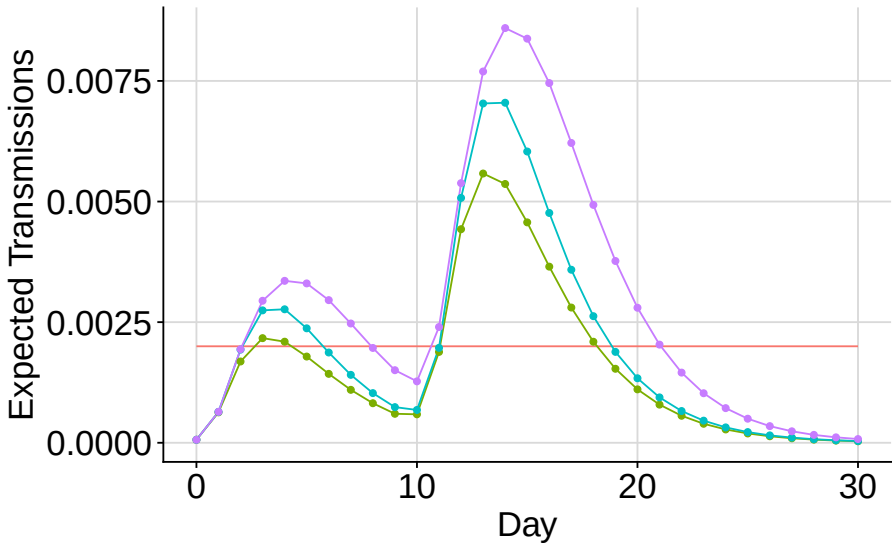


Figure 4.6: Expected daily transmissions following exposures with risk 0.01 on day 0 and risk 0.025 on day 10. The purple curve is independent of symptom information. The blue curve includes the condition that no symptoms were experienced up to the previous day, and for the green curve no symptoms were experienced up to the current day.

Conditional on Symptoms and Test Result

The risk estimate from the previous section can be improved significantly with knowledge of a COVID test result. Using Bayes theorem, the posterior risk of infection can be updated as shown in equation 4.16.

$$
\begin{aligned}
&P(D_{inf} = y, D_{symp} = z, Asymp = a | !Test(q)) \\
&= \frac{P(!Test(q) | D_{inf} = y, D_{symp} = z, Asymp = a) * P(D_{inf} = y, D_{symp} = z, Asymp = a)}{P(!Test(q))} \\
&= \frac{P(!Test(q) | D_{inf} = y, D_{symp} = z, Asymp = a) * P(D_{inf} = y, D_{symp} = z, Asymp = a)}{\sum_{i \in A} \sum_{j \in A} \sum_{k \in \{0,1\}} P(!Test(q) | D_{inf} = i, D_{symp} = j, Asymp = k) * P(D_{inf} = i, D_{symp} = j, Asymp = k)}
\end{aligned}
\tag{4.16}
$$

As before, the distribution can be truncated based on days where no symptoms have been detected.

$$
\begin{aligned}
&P(D_{inf} = y, D_{symp} = z, Asymp = a | !Test(q), (Asymp = True \text{ or } D_{symp} > w)) \\
&= \frac{P(D_{inf} = y, D_{symp} = z, Asymp = a | !Test(q)) * I(z > w \text{ or } a)}{1 - \sum_{i \in A} \sum_{j \leq w} P(D_{inf} = i, D_{symp} = j, Asymp = False | !Test(q))}
\end{aligned}
\tag{4.17}
$$

Using Equations 4.16 and 4.17, the expected number of transmissions on a day can be conditioned on test results and symptom status by summing over all values of D_{inf}, D_{symp}, and $Asymp$ as shown in Equation 4.18.

$$
\begin{aligned}
&E[T(x) | !Test(q), (Asymp = True \text{ or } D_{symp} > w)] \\
&= \sum_{i \in A} \sum_{j \in A} \sum_{k \in \{0,1\}} \{E[T(x) | D_{inf} = i, D_{symp} = j, Asymp = k] * \\
&P(D_{inf} = i, D_{symp} = j, Asymp = k | !Test(q), (Asymp = True \text{ or } D_{symp} > w)\}
\end{aligned}
\tag{4.18}
$$

Figure 4.7 shows the daily transmission risk following an exposure on day 0 with 5% estimated infection risk. Two curves are shown: one with with no additional information, and another with knowledge that no symptoms were experienced the previous day, and a negative test occurred on day 8 (with a 2 day reporting delay). Both pieces of information significantly reduce the expected transmission risk. In the next section we will show how intelligent risk estimation using testing and symptom data can be used to reduce expected quarantine lengths while maintaining efficacy.

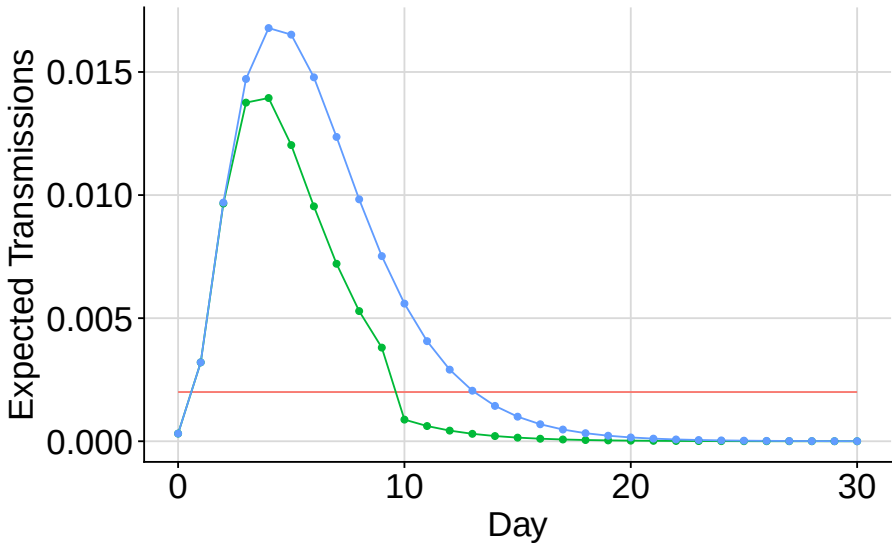


Figure 4.7: Expected daily transmissions following an exposures with 5% probability of infection on day 0, and a test on day 8. The red curve shows a typical risk threshold, the blue curve shows the unconditional risk, and the green curve shows the risk conditional on no symptoms the previous day, and the negative test result (with a 2 day reporting delay).

4.3.3 Quarantine Policy

In Chapter 3, we show that the optimal quarantine policy can be chosen to minimize the total amount of social distancing required to meet a target reproductive number, and that this policy is highly dependent on the disease state within a population. The quarantine policy can be represented as a risk threshold giving the probability of infectiousness that is required to quarantine an individual. In Chapter 3 the risk of an infectious individual is assumed to be evenly distributed over the period of infectiousness. This assumption can be improved using more detailed information to compute expected daily transmissions as was done in Section 3.

In economic terms, the risk threshold is the point of indifference between expected daily transmissions and the cost of an additional quarantine day. Assuming that the cost of a day of quarantine is fixed, quarantining any individual with expected daily transmissions exceeding the threshold generates a surplus.

The quarantine decision for a given day, x, can then be represented as the optimization problem in Equation 4.19. Here, r_{thresh} gives the risk threshold in units of expected transmissions per day, and $Data_x$ denotes all information available on day x. This includes any test results and the measurement of symptoms, both delayed by a set number of days. π_x is a boolean variable taking the value of 1 if the decision is to quarantine the individual on day x, and 0 otherwise.

$$\pi_x^*(Data_x) = argmin_{\pi_x}[\pi_x + (1 - \pi_x)E[T(x)|Data_x]/r_{thresh}] \tag{4.19}$$

This quarantine decision process could be incorporated into an exposure notification app that evaluates exposure risk, queries for daily symptom status, and records test results. Using Equation 4.18, expected daily transmissions can be calculated to give immediate and projected quarantine recommendations consistent with the regional risk threshold.

An intelligent quarantine policy can significantly reduce the overall cost of disease mitigation. Individuals with higher risk of infectiousness will be asked to quarantine for longer, and individuals with lower risk may only be asked to quarantine for a few days (or none at all). Compared to a fixed quarantine policy, the adaptive approach substantially reduces the overall cost of quarantine and social distancing. As a rough example, consider a risk threshold of 2e-3, exposure on a single day with uniform probability of infection between 1% and 20%, and comparison with a fixed 14 day quarantine policy. The optimal approach has an average cost of 14.9 and the fixed approach has an average cost of 18.9 (in units of quarantine days). Valuing the cost of a single day of quarantine at $150, the smart policy saves an average of $600 per person quarantined.

4.3.4 Optimal Test Timing

In Section 4.2, we explore the choice of fixed quarantine length and test timing after a single exposure. Here we generalize the procedure to account for multiple exposures with estimated risk, and the incorporation of an adaptive quarantine decision using equation 4.19. In addition, the test timing problem is posed in a framework consistent with the quarantine optimization: choose the test date that minimizes the expected value of equation 4.19 over the days in consideration (e.g. the 30 days following the latest exposure).

One approach is to pose this decision as a two stage stochastic programming problem: in the inner optimization choose which days to quarantine, and in the outer optimization choose which day to test. The most general (but least computationally efficient) form of this problem is solved by integrating over all possible scenarios as shown in Equation 4.21. The optimal value of q can be found by computing the cost for all reasonable test days and choosing the lowest.

$$min_q \sum_x E[(1 - \pi^*_x(Data_x(q))) * T(x)/r_{thresh} + \pi^*_x(Data_x(q))] \tag{4.20}$$

$$\begin{aligned} &min_q[\sum_y \sum_z \sum_{t_q} \sum_a P(D_{inf} = y, D_{symp} = z, Test(q) = t_q, Asymp = a) * \\ &\sum_x \{\pi^*_x(Data_x(D_{symp} = z, Test(q) = t_q, Asymp = a)) + \\ &(1 - \pi^*_x(Data_x(D_{symp} = z, Test(q) = t_q, Asymp = a))) \\ &* E[T(x)|D_{inf} = y, D_{symp} = z, Test(q) = t_q, Asymp = a]/r_{thresh}\}] \end{aligned} \tag{4.21}$$

There are computational approaches faster than a search over all possible test dates and disease scenarios. However, the above approach is tractable for problems of interest if the quarantine decisions are cached instead of recomputed each time in the loop. For example, if test and symptom dates are considered over 30 days, and there are 3 possible exposure dates, then 30*3*30*2*2*30 = 324000 function calls are needed. This is easily compute-able on a smart phone. The main benefit of this approach is that it reduces implementation complexity; more nuanced optimization and expected value calculations do not need to be coded and debugged.

In Figure 4.8, the expected cost of each possible test day is shown in terms of quarantine time and transmissions for an example problem with an exposure on day 0. This example

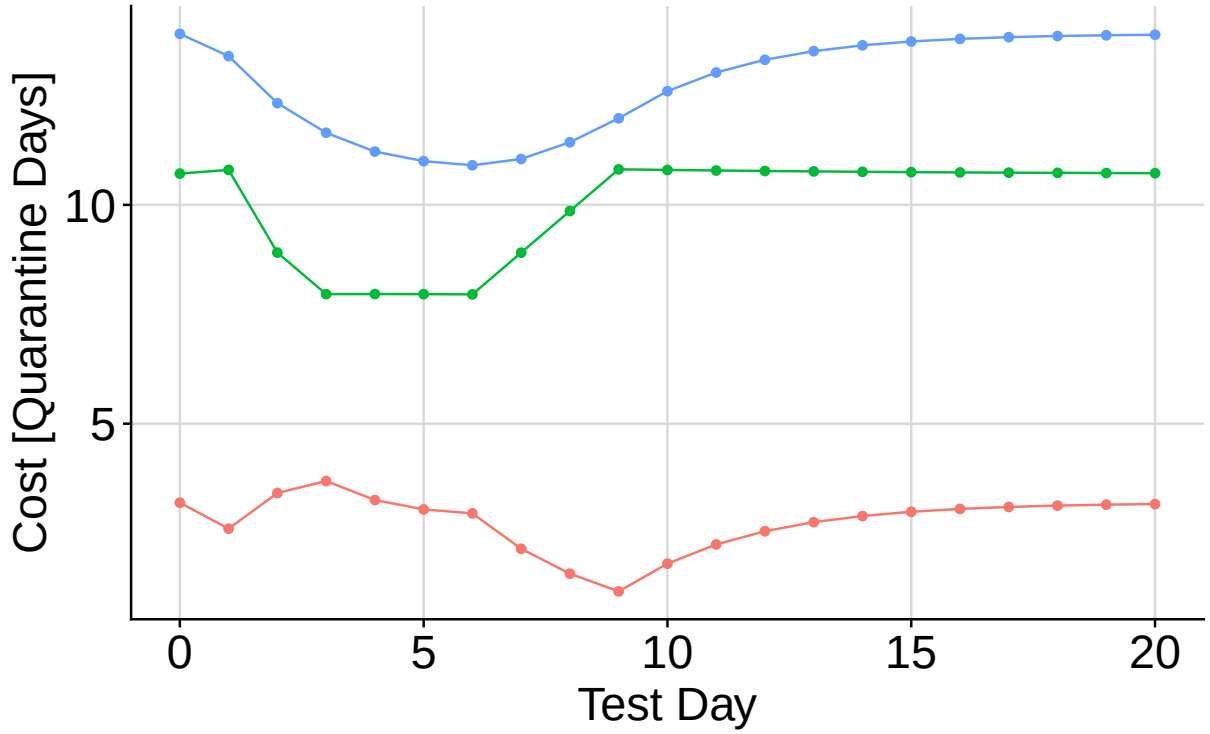


Figure 4.8: Expected quarantine cost as a function of test day, given an exposure with infection probability 5% on day 0. The green curve is the quarantine cost, the red curve is the transmission cost (converted to units of quarantine days), and the blue curve is the total cost. The optimal test date in this scenario is day 6.

has only one exposure day for easier interpretation, but the method works for an arbitrary number of exposure days. The transmission cost is put into units of quarantine days using the risk threshold as a conversion ratio. From the plot it is clear that testing on day 6 would minimize the overall expected cost. In this scenario, the difference in cost between no test and the optimal test date is equivalent to 3.0 quarantine days. Compared to a fixed quarantine policy of 14 days with no test, there is an expected reduction of 5.4 quarantine days.

4.3.5 Conclusion

Quarantine plays an important role in mitigating the spread of Covid-19, but it is very costly to individuals and society as a whole. By incorporating all available information about exposure risk, symptoms, and test results, an adaptive risk assessment can be used to reduce quarantine lengths when possible. Further gains can be made by optimizing the test date to reduce the expected cost of quarantine using a form of two stage stochastic programming. In combination, these methods reduce the expected cost of quarantine by several days for each individual.

Chapter 5

Improvements for Idealized Case

Previous chapters explored specific technologies and strategies to reduce the cost of disease control. The aim of this chapter is to focus on the idealized case (with high system participation) to investigate what closer to optimal strategies for disease containment could look like. Finally, in Chapter 6 we will discuss some of the real-world challenges faced when deploying these systems.

5.1 Building Classifiers to Predict Infectiousness

Manual contact tracing and digital contact tracing make use of rules to determine whether an exposure is of sufficiently high risk to warrant quarantine and/or testing for a contact. This risk assessment can be refined by training classification algorithms on real world exposure and infection data. More accurate risk assessment would make quarantine more efficient (less disruption needed to prevent the same number of transmissions) and allow for the optimal quarantine risk threshold to be set based on disease status in the general population.

Manual contact tracing rules are typically based on heuristics for risk such as thresholds on duration, proximity, and timing of exposure relative to symptoms [132]. Some health departments recommend contact tracers subjectively adjust risk assessment based on other factors like airflow, activity, and PPE worn [133], however a quantitative approach to combine these risk factors is lacking. Risk scoring for digital contact tracing is usually based on duration of contact, RSSI (Bluetooth signal strength as a rough proxy for distance), and timing of the exposure relative to symptom onset or date of positive test. Rules to

combine this information are typically chosen to mimic manual contact tracing (15 minutes at 2 metres) [134], or by quantitatively combining different risk factors using parameters typically estimated early in the pandemic using intensive contact tracing [118]. But these datasets are often small, not updated, and can be slow to surface.

While heuristics help simplify decisions, they can also lead to sub-optimal results. Classification rules for high-risk contacts that depend on specific boundaries for individual variables don't treat risk consistently. For example, someone who spent 3 hours 2.5 metres away from someone infectious is likely a higher risk contact than someone who spent 15 minutes 2 metres away, but standard rules would make the opposite classification. By treating the risk factors multiplicatively this difference could be captured, as described in our work on EN risk assessment [129]. However, even more advanced risk assessment approaches could be challenged by uncertainties about disease transmission and propagation in a contact network.

There have been many data-driven studies aiming to understand transmission risk (e.g. the timing of transmission [85], contact mode and viral load [135]). However little work has been done to develop classification algorithms that take as input information related to exposure to a known case and predict the probability of the contact becoming infected (likely due to the challenge of getting access to enough high-quality data). A review of data driven methods for contact tracing from May 2022 found no articles on estimating the risk of transmission based on details of contact [136]. One paper since then (June 2022) used test results from 1130 close contacts to train a Naive Bayes classifier on 6 variables: contact age, relationship to patient, contact situation, wearing a mask, distance to patient, and contact time [137]. While promising, the data set could be much larger and more precise by making use of digital contact tracing, and this could be done much earlier in the pandemic[1] to actively guide policy and minimize costly quarantine.

We can use real-world data to evaluate the performance of existing decision rules and to see if statistical algorithms perform better. As an illustrative example, Figure 5.1A shows a fabricated ROC curve [138] which could have been generated by training a classifier on exposure data and eventual test results. The classifier it is based on could be a number of black box algorithms that are frequently used in low dimensional classification problems, such as random forest or gradient boosting [139].

One thing we can use this curve for is to choose the optimal classification boundary based on a set risk threshold (ideally computed using methods from Chapter 3). For this section, we will assume that the risk threshold is unchanged by quarantine choices - essentially a linearization around an operating point. This approximation is valid if the

[1] and regularly updated to account for new variants and behaviour change

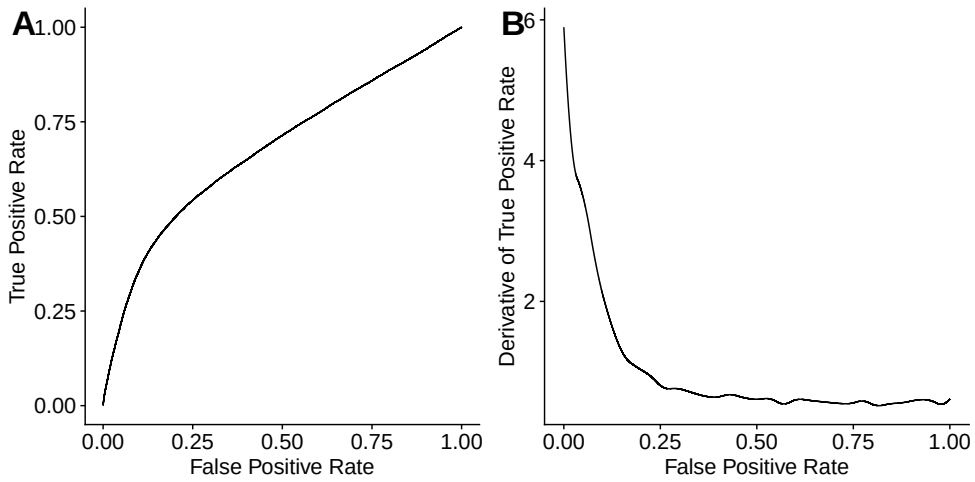


Figure 5.1: A) Example ROC curve for classification of infectiousness. B) Numerical derivative of ROC curve

subpopulation considered contains a small fraction of the total number of infectious people (which unfortunately is the case with the subpopulation of people who received an Exposure Notification due to low adoption). It is possible to make this choice in a derivative-free way by posing the problem as a linear optimization with the solution constrained to lie on the ROC curve. However, using the approximate derivative of the ROC curve gives a better illustration of risk distributions and how they change with better classifiers and information. We will use the numerical derivative of the ROC curve (using smoothing splines as an intermediate step) as shown in figure 5.1B. In general, the derivative of the ROC curve should be non-increasing, as for any dataset, if the ROC curve isn't concave down it can be replaced by a strictly better classifier that does produce a concave down ROC curve[2].

The derivative has units of (True Positives / Total Positives) / (False Negatives / Total Negatives), which if multiplied by (Total Positives / Total Negatives) gives the marginal increase in positives quarantined per unit increase in negatives quarantined. This is almost the same as marginal risk when the marginal risk is low (which is almost always the case in areas of interest). To be safe however, we can transform our risk threshold to this format

[2] "You can always create a model whose ROC curve is the convex hull of the original ... The new model is created by looking at pairs of adjacent vertices of the convex hull, and creating random combinations of them at different probabilities" [140]

using $r' = r_{thresh}/(1-r_{thresh})$, which 'subtracts away' the true positives in the denominator.

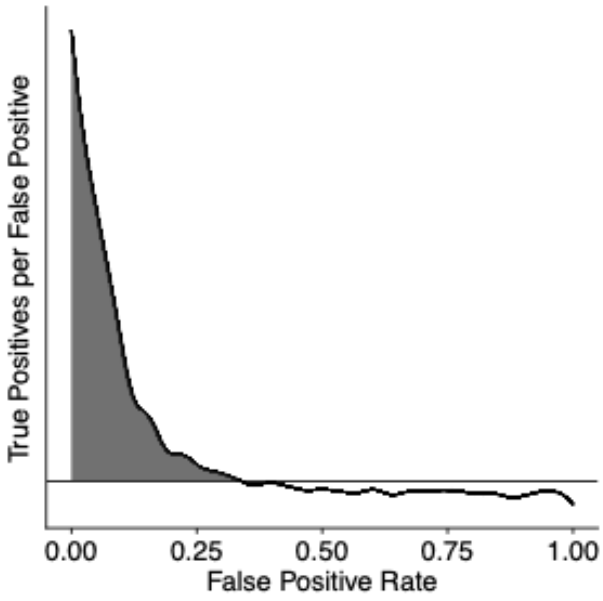


Figure 5.2: Ratio of true positives to false positives and shaded 'surplus' for a given risk threshold

As argued in Section Chapter 3, the risk threshold (in this case r') is our point of indifference, which we can draw as a horizontal line on the graph in Figure 5.2. Further, using reasoning from microeconomics [141], any area above this threshold represents a surplus (willing to trade x days of quarantine for y prevented transmissions, but instead we get y + s). So, for a given risk threshold, we would like quarantine to apply to everyone left of the point of intersection with the r' line. The benefit (or rather, reduction in cost) we get from this intervention is the shaded area above the risk line- representing the benefit from infectious people quarantined minus the cost of quarantine for the false positives.

For a given risk threshold, the optimal classifier is the one that provides the best decision-relevant risk information, which is computed as the area between the two curves. Making this data driven allows us to objectively compare candidate decision rules and choose the best one (with the highest 'surplus' for a given risk threshold). This calculation can be used to compare several classification approaches; e.g heuristics from manual contact tracing, quantitative estimates currently used by EN apps, or with standard classification algorithms from machine learning. If there are patterns in the data that are not used by

the hand-designed rules, then black-box classification algorithms may perform better.

To summarize, applying the classification algorithms to a test data set allows us to estimate the marginal risk of infectiousness based on the relative ranking within the classification. Using this marginal risk of infectiousness, we can make quarantine recommendations that are consistent with our target risk threshold. Using a similar (but separate) training data set, we can train black-box or partially black-box classification algorithms to search for un-modelled patterns in the data. For a given classification algorithm and risk threshold we can compute the 'surplus' risk captured, allowing us to choose the best between a set of candidate algorithms. Quarantine decisions are extremely costly in the real world, so we should do as much as we can to optimize their accuracy.

5.2 Improved Information

The data used to make quarantine recommendations and inform other NPI decisions is not fixed. The utility of better information can be visualized by how its inclusion affects estimated risk distributions, and the benefit can be quantified by the change in area of the surplus region. Which information is valuable is thus a function of the risk threshold.

More practically, this means that the value of information depends on whether it will be acted on in a useful way. As an example, at the beginning of the COVID-19 pandemic we (Covid Watch) were considering developing location-based proximity alerts at a low enough spatial resolution that they could not be used to identify people. Unfortunately, the low spatial resolution meant that the number of people that would be notified per index case was quite high- potentially thousands. There was little willingness to act on risk at this low of a level at the time (outside of some places like Vietnam), so this was one of the main reasons that we focused on Bluetooth proximity tracing.

In machine learning, the decision about what types of information to include in a model is called feature selection [142, 143]. Typically, feature selection for contact tracing risk assessment has been done based on expert consultation [144]. Perhaps the most useful 'feature' added to Covid contact tracing was consideration of ventilation during exposure [145]. However, the inclusion of this data in public recommendations was significantly delayed in many countries. In future outbreaks, automated collection of exposure information for a portion of the population that opts in could reduce the time required to find important risk factors. For example, contact tracing apps could infer the venue type based on nearby Wi-Fi routers and this information could be coupled with other exposure details and the eventual infection status.

There are many features that can be used to predict risk of infectiousness. Some features that are used or could be used for contact tracing are: duration and proximity of exposure, timing of exposure relative to symptom onset, any symptoms experienced by the contact, timing of exposure relative to test (and CT count of test if PCR or darkness of line if antigen), PPE, type of environment (specifically airflow and activity), disease variant, past vaccinations/infections, higher order infection information in the nearby contact network, and probably many others. Real-world data, numerical modelling, or back-of-the-envelope calculations could be used to determine the influence of each source of data on the resulting risk distribution. Doing so allows us to estimate whether risk density is shifted in a useful region, and to focus limited effort on the data sources that will do the most to improve the efficiency of quarantine recommendations.

5.3 Dynamic Strategy

In Chapter 3, the primary focus was on guiding quarantine decisions at the margin. This was an attempt to make progress in a complex dynamical system with limited control. In an ideal case however, strategy can shift from responsive harm reduction, to more intentional control, borrowing concepts from Optimal Control [12] where possible. Here the aim is to outline a high level strategy to find optimal trajectories in disease prevalence over time.

Several researchers have explored the use of Optimal Control for guiding NPI choices before [146] and during [147, 148] the Covid-19 pandemic. These works have used compartmental disease models with the transmission rate modulated by broad social distancing. They have chosen control strategies to minimize an objective function computed as a weighted sum of infections (or deaths) and the time integral of control cost (which is either a linear or quadratic function of control intensity).

Several simplifying assumptions in previous models prevent them from being able to explore strategies for local disease elimination, and so they may be recommending substantially sub-optimal solutions. The deterministic compartmental models used are unable to reach a state of complete disease elimination (unless starting at 0 prevalence, or control is unrealistically modeled as 100% effective) [149]. By neglecting this absorbing state, these studies miss the opportunity for control to be completely relaxed without an increase in deaths. In addition, they only consider broad social distancing, which has a cost independent of disease prevalence.

As highlighted in Chapter 3, targeted interventions have a cost that is proportional to prevalence, and so can be substantially cheaper than broad social distancing. By neglecting

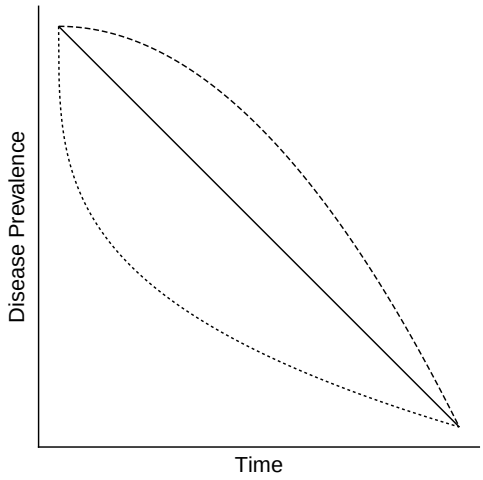


Figure 5.3: Possible trajectories for disease prevalence over time. The cost of each depends on the cost of rate of change and cost of infections.

targeted interventions, previous studies haven't been able to explore optimal control when there are increasing returns to disease mitigation. Finally, previous studies have treated the rate of imported cases as a fixed value. In reality, the importation rate is a function of border policy, and is very important when local disease prevalence is low (as discussed later in Section 5.4).

In this section we will consider an isolated population, and in Section 5.4 we will address the influence of border policy and imported cases on strategy. To compare trajectories, we can approximate the cost of both the derivative (equivalent to growth rate- which depends on NPIs) and case load (which will have a cost depending on health outcomes[3]), then integrate the total cost of the trajectory over time. Numerically optimizing this trajectory cost would yield a solution giving the best dynamic NPI policy (which implicitly depends on available risk distribution via the efficiency of targeted interventions).

To simplify things, we will separate the population into two categories as in Chapter 3: those in quarantine/isolation with complete social distancing, and those in the remaining population with a value for social distancing that can be set between 0 and 1 [4]. The first

[3]And potentially a binary variable representing the psychological/political benefit of complete local elimination

[4]This is a simplification of the more general case where social distancing can be a continuous variable set for each person separately

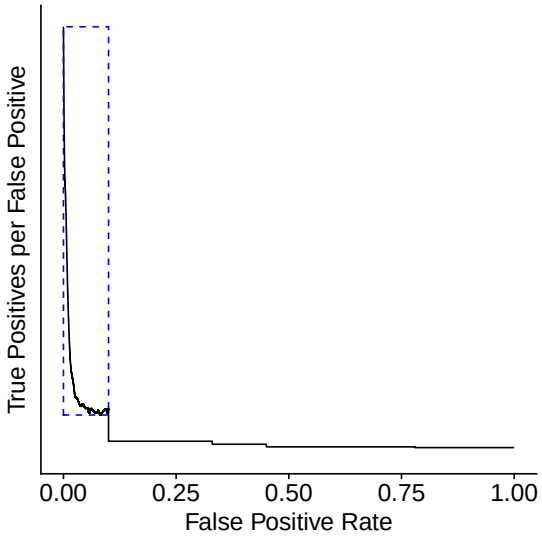


Figure 5.4: Example risk distribution for population with two sources of information: received exposure notifications and rough location. EN subpopulation shown in rectangle and exaggerated for visualization. Stepped pattern due to varying disease prevalence.

step is to develop a process to choose the optimal risk threshold for a specified value of R_t when we have control over Q_i and Q_n. To do this we will use a risk distribution like the one discussed in Section 5.1, but for the whole population. Like most tasks requiring reasoning under uncertainty, building this risk distribution by combining multiple sources of data is challenging to do accurately. However, it is better to make an estimate and be off by an order of magnitude than to make no estimate at all or delay too long and take actions that are off by several orders of magnitude.

As a rough approach, we can combine risk for multiple subpopulations via sorting if they refer to different people, average if multiple assessments for same person using same data, or summation if for the same person using different data. For people in the general population with no information, the remainder of the risk can be assigned uniformly (it is useful to start with an estimate for total number of infections in whole population and ensure resulting risk profile sums to this). As an example, we could take risk estimates for the subpopulation of people who received exposure notifications and assign the remainder of the risk depending on infection rate in the rough location of residence. By virtue of receiving an EN they are likely higher than average risk (even if it is a very low risk EN). So the risk curves are roughly appended like in Figure 4 (where the size of the

subpopulation receiving an EN at a given time is exaggerated for visualization). We could refine this distribution by adding any of the types of information mentioned above, or in the more intermediate regime with things like location or symptom status.

We can use this risk distribution to compute the minimum cost needed to achieve a specified value for Rt using the following algorithm.

Algorithm 1 Quarantine policy: computes minimum social distancing cost of achieving specified R_t (implicitly depends on available information for risk distribution)

$sort(\boldsymbol{r}[])$ # sort population from highest to lowest estimated risk of infectiousness
$Q_i = 0$
$Q_n = 0$
j = 0
while $j < P$ **do**
    $r_{thresh} = computeThreshold(Q_i, Q_n)$
    $R_t' = R_t * (I - Q_i)/I$
    if $r[j] > r_{thresh}$ ***and*** $R_t' > R_{target}$ **then**
        $Q_i = Q_i + r[j]$
        $Q_n = Q_n + 1 - r[j]$
    else
        break
    end
    j++
end
$Cost = (Q_i + Q_n) * f(1) + (P - Q_i - Q_n) * f(1 - R_{target}/R_t * I/(I - Q_i))$

Essentially this works by applying the marginal quarantine decision incrementally from high risk to low risk and adding to the expectation of Q_i and Q_n each time, only stopping when R_{target} is achieved or an additional quarantine is no longer worthwhile.

We can then generate a function giving minimum social distancing cost as a function of R_{target}, which implicitly depends on the available risk distribution. As a (quite strong) simplifying assumption, we can assume that the risk distribution is possible to estimate based on disease prevalence in the population. The validity of this assumption depends on the types of risk information considered- for example it is possible to predict the amount of information available from randomized testing as a function of disease prevalence. Combined, this allows us to compute the minimum cost of achieving a given R_{target} value as a

function of disease prevalence (by first estimating the risk profile from prevalence and then computing the cost of the R_{target} value for the risk profile).

The last piece needed to compute the cost of the time derivative is a mapping from R_t (onward transmissions per case) to the growth rate in time: r_t. By assuming that the mean generation time, T_c, is unchanged by interventions, we can compute this mapping as [150]:

$$r_t = (R_t - 1)/T_c \tag{5.1}$$

We can now express the cost of achieving different rates of change as a function of disease prevalence, allowing us to compute the cost of potential trajectories over time. The cost given by this function only represents social distancing; it should also be summed with a penalty term for health outcomes by integrating the total number of infections over the outbreak. The function may not be differentiable, but it is low dimensional and can be discretized in time, allowing us to use numerical optimization to compute the best trajectory. A high level description of the design problem is given by Equation 5.2, where $f_r()$ is a function giving the cost of achieving a given growth rate (using equation 5.1 and algorithm 1), $I(t)$ denotes infections at time t, and d is the rate of death of infected people [5].

$$\begin{aligned} \min_{r_t} \quad & E[\int f_r(r_t, I(t)) + d * I(t)\, dt] \\ \text{s.t.} \quad & \dot{I}(t) = r_t * I(t) \end{aligned} \tag{5.2}$$

Clearly there are several challenges with implementing an approach like this in the real world: social limitations which will be discussed in section 6.1, as well as more technical challenges with estimation under uncertainty and the importance of stochasticity when disease prevalence is low[6]. However, such a framework is useful for providing intuition about how to use information to target interventions in different stages of an outbreak, and why an approach like the one Vietnam took would be worthwhile. This method can also be used to roughly compare the cost of continued suppression via broad social distancing, vs targeted suppression enabled by a lower case prevalence and good information, vs. temporary elimination via costly but brief suppression.

[5] summing deaths this way was chosen mainly for convenience of notation - other methods may be better

[6] Stochasticity can be partially addressed by sampling the number of new cases at each timestep as a function of both prevalence and r_t and averaging over several of these generated trajectories.

5.4 Border Policy

Prior to the Covid-19 pandemic, travel restrictions were thought to be a relatively ineffective disease control strategy - only serving to briefly delay the inevitable outbreak [151]. A simple model by Tomba et al estimates that with a 99% reduction in importations, a local epidemic would be delayed by 3 weeks [152]. However, during the Covid-19 pandemic, several countries demonstrated the effect radical border policies (including testing, quarantine, and a reduced numbers of travelers) could have on the rate of introductions [153]. The frequency of introductions also depends on prevalence in neighboring countries, which is simplest to model as a fixed quantity, but there is some work on economic approaches for countries to negotiate disease strategies [154]. When local disease prevalence is high, imported cases aren't significant [155]. However importations are a crucial part of the epidemic control problem when local prevalence is low.

There is a cost to increased border stringency – both in dollars from reduced tourism or trade, and as a reduction in wellbeing from inability to travel or see friends/family. Conversely, each introduction has an expected cost of suppression and an expected cost of health outcomes. Border stringency can reduce the rate of introduction, so the policy should be chosen as a trade-off between the cost of reduced movement and the cost of more new cases.

This can be made slightly more formal by estimating the rate of introductions as a function of policy (using a per-person estimate accounting for mistakes, false negatives, and long incubation periods), then estimating the cost of each policy, and combining the two to get the cost as a function of the fraction of importations prevented. The cost per imported case depends on the broader control strategy. For regions aiming to maintain local elimination, the cost of a new introduction could be estimated using a stochastic model of the implemented suppression strategy (like the branching model used by Plank et al [156], but also including the use of targeted interventions). This would yield an estimate of the expected cost per introduction in terms of dollars, quarantine days, and negative health outcomes, which could be roughly converted to the same units [7]. Then we can find the optimal border policy (for the local elimination scenario) by choosing the frequency of introductions that minimizes the sum of the border stringency and outbreak suppression functions.

For a more general-purpose analysis of border policy, the decision can be incorporated into the previously discussed Optimal Control problem by modifying Equation 5.2 to include imported infections (to keep things simple, exported infections are not considered

[7] This is again an order of magnitude estimate, not an exact calculation

here[8]). In Equation 5.3, C gives the counterfactual rate of imported infections[9] (with no border control), b_t gives the fraction reduction in importations at time t, and $f_b()$ computes the cost of achieving that border policy.

$$\begin{aligned} \min_{r_t} \quad & E[\int f_r(r_t, I(t)) + f_b(b_t) + d * I(t)\, dt] \\ \text{s.t.} \quad & \dot{I}(t) = r_t * I(t) + b_t * C \end{aligned} \tag{5.3}$$

Even without solving this optimization problem, we can see how border policy relates to local control measures. For any policy $\{r_t, b_t\}$ at time t, we can infinitesimally shift the policy to $\{r_t + dr, b_t - dr * I(t)/C\}$ without changing $\dot{I}(t)$ (assuming $C \neq 0$). The change in cost associated with this policy shift is $dt * (f_r'(r) * dr - f_b'(b) * dr * I(t)/C)$. If $f_r'(r) > f_b'(b) * I/C$ this means we can achieve the same outcome at lower cost by loosening local control measures and strengthening border control (and the opposite if $f_r'(r) < f_b'(b) * I/C$). Therefore, having $f_r'(r) = f_b'(b) * I/C$ is a necessary (but not sufficient) condition for a given policy to be optimal.

As with Chapter 3, these cost functions may be difficult to evaluate exactly, but in some scenarios even order of magnitude estimates could reveal whether there are substantially more efficient policy choices available.

5.5 Design Problem

In sections 5.3 and 5.4 we were interested in making the best decisions given fixed capabilities. Alternatively, these models can reveal which capabilities are the most valuable to enhance - either by increasing the reliability of existing methods (e.g. refinement of border testing timing) or by developing new strategies (e.g. digital contact tracing). Effective changes can be interpreted as reducing the functions $f_r()$ or $f_b()$ (achieving the same amount of disease containment at a lower cost). The most valuable capabilities to enhance are ones that reduce the cost in areas of the parameter space, $\{r_t, b_t, I(t)\}$, that a country expects to spend a lot of time in.

For regions with a very high expected cost of an imported case, it is worthwhile to put significant effort into ensuring border policies are functioning as intended. Analyses

[8]For most scenarios, exported infections would not be significant unless the 'border' is chosen in a way that caused a significant fraction of the infected population to leave

[9]The counterfactual importation rate could depend on time

like in Chapter 4 could be done to estimate the number of breakthrough cases depending on test timing and quarantine duration. Additionally, unexpected transmissions could be investigated thoroughly to ensure the same mistake is not repeated [157, 145]. An approach from cyber security, penetration testing, could be used to validate airborne precautions using a technology like [158] that releases detectable (and safe) airborne particles. Using this approach, points of failure like limos with shared air [157] could be identified at lower cost by sending someone through the quarantine process with a device that released these particles, and later swabbing people and surfaces to test for unintended contamination.

Interconnected regions may not be able to bring the frequency of introductions close to 0 (or the cost function may be very steep below the level where state-sanctioned travel has been forbidden). Reducing the cost of local suppression (when $I(t)$ is relatively low) as much as possible can allow for low disease prevalence to be maintained even with frequent re-introductions. It is worthwhile to explore the use of technology to improve this capability.

5.6 An Idealized (Tech-based) Solution

With risk aware strategy, high cooperation, and modifications of existing technology, infectious diseases like COVID-19 could be contained at a very low expected cost per person. In this section we will explore an idealized response to a novel infectious disease (or COVID-19 / Influenza) before addressing real-world limitations in Chapter 6. Borrowing from algorithm design in computer science, we would like our solution to have a cost that is roughly proportional to the number of infected individuals, not the total population, so that with prevalence $<< 1$, the cost per person (in the total population) is negligible. For a completely targeted approach (without any broad social distancing) to be able to reduce cases ($R_t < 1$), the number of transmissions per case that occur outside an identified subpopulation (in time to prevent additional transmissions) must be less than 1. There are two components to this: detecting cases and finding a suitable subpopulation. We will aim to develop a strategy that can reliably identify such a subpopulation for diseases transmitted by air or fomite, with some robustness to delays and unidentified cases.

Since this is an idealized case, we will assume 100% participation and compliance with recommendations. This would include usage of custom devices for people without smartphones (similar to Singapore's tokens [46]). The proposed system functions similarly to the one in Chapter 2, with 3 major modifications (and the minor modification of longer storage of local records to allow backward contact tracing for diseases with an incubation period greater than 2 weeks).

Wi-Fi router logging

As proposed by Prasad and Kotz [25], devices also record the MAC addresses of nearby Wi-Fi routers and store this locally along with signal strength and time period. People who test positive upload their record of Wi-Fi logs, which is stored privately by public health. Wi-Fi data (or the hash of Wi-Fi data) for locations deemed to have transmission risk[10] are broadcast to the notification server along with a time period. To check for an exposure, all devices download this data from the notification server and compare it to their logs (in addition to the existing Bluetooth-based protocol).

2nd (or higher) degree notifications

Higher order contact tracing (contacts of contacts of cases) in Vietnam [159] and backward contact tracing in Japan [160] were both effective at controlling the spread of Covid-19. Modeling work shows that risk estimation from higher order tracing better prioritizes scarce resources [161], and that including backwards tracing is more efficient than only forward tracing for disease spread with superspreading dynamics [162]. Bestvina et al proposed using Monte Carlo simulation of disease spread within an app-based network to estimate infection risk [163], which was implemented by Aarogya Setu [164], but was much more challenging to use with decentralized protocols[11]. Exposure Notification provides a report type for 'recursive' notifications [58], but (to the best of my knowledge) it was not used. In theory, this recursive notification would allow people who were notified of direct exposure to then notify their contacts. However, this process (as currently implemented) would require a new permission number from a public health authority in order to send this recursive notification.

A lower friction implementation could gate the triggering of recursive notification based on knowledge of data originally broadcast over BLE by the index case and not shared elsewhere (knowledge of this data is evidence of being near the index case recently, which prevents system misuse). Furthermore, a setting could be used to enable automatic recursive notification upon receiving a primary notification.

[10] with the potential for some logs to be redacted by the user

[11] due to quantity of information that would need to be transmitted and privacy implications of methods to reduce that data size

Exposure sharing for limited contact network assembly

Information about notified contacts and details of the exposure could be valuable for manual contact tracing and to better understand transmission routes. This capability is fairly simple to include in a decentralized protocol (without enabling broad surveillance). To do this, notified devices upload the details of the exposure (duration, proximity, time, and key of device matched with) to a private server operated by public health along with their name and contact info. The local contact chain can be reconstructed if the index case shared their identity with public health (which could be linked to the key they used to notify their contacts).

Analysis of system benefits

The system cannot be used for mass surveillance, as data is stored locally on devices. When exposure data is transmitted to the public health server the user is notified (and if someone did not want this to happen – following a protest, for example - they could delete the data on their app). Additionally, legislation could be introduced to prohibit use of the system for non-health related purposes.

This system has several benefits over the original version. Essentially, it allows rapid assembly of the 2nd degree contact network around an index case in a way that decision makers could use to assess risk, assist with testing and quarantine, and learn about transmission dynamics. Pre-built software could be used to combine this contact network information with any other identified networks, reported symptoms, and test results (including genomic sequencing). Second degree contact tracing could prevent transmissions even if the first-degree contacts are not notified in time. Additionally, second degree contact tracing is equivalent to backward contact tracing when using a wide enough time range. If fomite transmission is occurring (or transmission via particles staying in the air), notifications could be sent to people who visited a location after a case using Wi-Fi data. Appendix C explores how a system like this could be used in an idealized response to a novel epidemic.

Chapter 6

Discussion

The focus of this work so far has been on improving tradeoffs between disease suppression and the cost to achieve it. This has been achieved by developing technology and strategies that take advantage of the heterogeneity of infection status (over time and within a population). Specifically, we have developed a technology for sharing exposure information with the people who need to act on it, a framework for using risk of infectiousness to make decisions about quarantine that minimize overall social distancing, and methods to tailor quarantine and testing policy based on inferred distributions about the disease being controlled. We've also expanded on this approach by proposing strategies that make better use of data on infection risk within an idealized population.

In this chapter, Sections 6.1 and 6.2 discuss several practical limitations observed when applying targeted NPIs in the real world. Sections 6.3 and 6.4 investigate whether alternative approaches could perform better, and Section 6.5 examines more general limitations. Finally, Section 6.6 discusses directions for future work and potential strategies to make real-world progress on this complex but important problem.

6.1 Observations of Disease Dynamics

Unfortunately, low-cost elimination of COVID-19 using targeted NPIs was not the norm in many countries. Instead, we mainly saw oscillations of disease prevalence at a higher level for over a year before vaccines became widely available[1]. In many of these countries (such as Canada, which I am most familiar with), seroprevalence in the population remained too

[1]Oscillations also continued after vaccines became available, but that analysis is more complicated.

low for infections to have been significantly mediated via immunity[2], suggesting that behaviour modification was the dominant mechanism that prevented most people from being infected [165]. This oscillation is exactly what we would predict if targeted NPIs alone were insufficient and fear-motivated broad social distancing was also needed to suppress transmissions. In this model, people increase social distancing when there are more cases, either independently or as instructed by public health. This model is supported by observations of variants almost fully displacing each other (without reaching relevant levels of immunity); as cases rise from a new more competitive variant, social distancing increases to the level required to keep disease prevalence steady. This increases social distancing to a level where the prior variant is rapidly eliminated- demonstrating that elimination of the prior variant was possible the entire time, but only occurred when prompted by rising cases of the new variant. This fear-based control mechanism has the opposite property of targeted interventions- reduced disease prevalence makes control harder rather than easier.

Disease control mainly occurring through responsive and broad social distancing results in persistent costs being imposed on a large fraction of the population for the duration of the pandemic. We were fortunate that vaccines for COVID-19 were developed as quickly as they were because this is an extremely costly quasi-equilibrium that seems difficult to escape in either direction. Another implication of this dynamic is that it makes the longer-term influence of NPIs very difficult to measure - with technical or strategic improvements in suppression, cases would drop, then people would relax social distancing, causing cases to return close to their prior level[3]. The upside of this (as we argued in Chapter 3) is that improvements in targeted NPIs can indirectly reduce the burden of broad social distancing.

6.2 Failure Modes in Exposure Notification

This section will examine the possible impact of the EN system during the COVID-19 pandemic. For a number of reasons, evaluation of the effect of targeted NPIs (and NPIs in general) has been difficult. However, there are some exceptions, like the study evaluating the impact of manual contact tracing in the UK based on a quasi-randomized excel error in September 2020 [166]. They found that cases with timely manual contact tracing resulted in 63% fewer infections 6 weeks in the future than cases without. Wymant et al. analyzed the impact of the NHS app and found a reduction in cases between 100,000 and 900,000

[2]Since R0 was considerably above 1, unmitigated, simple SIR models predict the disease would have infected roughly $1 - 1/R0$ fraction of the population.

[3]This problem is made even more difficult by new variants, changing behaviour, and a shifting immunity landscape

during a time in which there were 1.9 million total cases [59]. The UK had one of the best EN deployments (along with Germany), so we can interpret the results from Wymant et al. as a rough upper bound on the benefit EN had in other regions during the Covid pandemic.

Quantification of NPI effectiveness is challenging because of how much the number of cases in the future depends on disease dynamics over the time period[4]. To make predictions about the impact of targeted NPIs, it would be useful to know the fraction reduction in R_t caused by those targeted NPIs. Defining this fraction as x, then if $R_0(1-x) < 1$, the targeted NPIs would be able to independently suppress transmissions. Both manual contact tracing and EN prevented a large number of infections in the UK, but not enough to achieve the type of disease control based on efficient local elimination discussed in Section 5.3 (clearly, because the disease was not locally eliminated in the UK).

To estimate the benefit of EN in other regions and look for areas of improvement, we can return to a step-by-step analysis of the notification process (similar to what we began in Section 2.2, Ferretti et al's adoption analysis [28], and Tomas Pueyo's analysis [167]). In the Covid Watch White Paper (Chapter 2), we identified the 3 factors (actually 4 separate steps) needed to notify someone of an exposure event:

1. The index case needs to receive a positive test result
2. The exposure needs to have been properly measured and judged to be high risk
3. Both the index case and the contact need to have downloaded the app prior to the exposure (which is actually 2 steps)

To directly prevent counterfactual transmissions, the index case not only has to receive a positive test result (point 1), but they also need to receive it before their contact has become infectious. Many locations had PCR wait times of more than a week during periods of high demand. Additionally, there are two more steps that we originally did not anticipate limiting performance that have proven challenging:

4. The index case must choose to upload their positive test result and receive the authentication to do so
5. After receiving a notification, the contact must choose to (and be able to) quarantine or get tested so that additional transmissions do not occur

[4]For example, if an NPI reduced R_t from 1.1 to 1.0, then over 6 generation periods the removal of that NPI would appear to cause 77% more infections: $(1.1^6 - 1)/1 = 0.77$

Combined, there are at least six points of failure in the process. Each of these steps rarely had more than a 50% success rate in the general population (although there may have been subpopulations that managed to perform better). In a reasonably functioning deployment of an EN app we could expect a success rate for each of these steps between 20% and 80%. Within this range, the worst-case result is a 0.0064% (0.2^6) reduction in R_t and the best case is a 26% (0.8^6) reduction in R_t.

Clustering of app usage could at most[5] reduce the adoption dependency from quadratic to linear, which would be equivalent to removing a point of failure- changing the bound to 0.2^5–$0.8^5(0.032\% - 33\%)$. Also, indirect effects could conceivably increase the efficacy, e.g., increased awareness about symptoms, receiving a notification and quarantining before they would have otherwise been infected, or receiving a notification and quarantining after being infected by a different index case. However, it would be surprising if any of these corrections shifted the resulting bound considerably.

As we argued in Chapter 3, even a 1% reduction in Rt is very valuable because it substitutes for much more costly broad social distancing. However, a limited resource to consider is decision-makers' attention. Use of a system like EN is worthwhile if it can realistically achieve a success rate that is worth decision makers' time (compared to other strategies they would otherwise be deploying or optimizing). The value provided to decision makers is improved if the EN system can not only be used to prevent transmissions, but also to learn about disease dynamics in real time.

This challenge with robustness is not unique to EN, but it is especially pronounced because of how many points of failure there are. As another example, manual contact tracing requires four steps equivalent to the ones listed above (excluding the two download steps) and success rates will vary depending on available capacity. There are technical and social changes that could improve the success rate for many of these steps, but some of these changes come with difficult tradeoffs.

6.3 Improving the Succes Rate of Exposure Notification

There are several modifications that could be made to the EN system to increase the success rate of specific steps. Overall, aligning incentives can encourage higher success rates – for example earlier testing and therefore access to treatments like Paxlovid for app users that

[5] Although the dynamics of this mixed-population scenario are not straightforward

receive a notification. The factor for the fraction of the population using the EN system occurs twice in the calculation for success rate, so it is the most important to have close to 1. There are many ways to increase adoption, but perhaps the most effective (short of mandating it) would be to make participation opt-out. As an example of the influence this can have, registration as an organ donor was between 86% and 100% in the 7 European countries where it was opt-out, and between 4% and 27% in the 4 European countries where it was opt-in [168]. Similar to organ donation, the choice to participate in EN is pro-social, unlikely to cause harm, but thinking through the pros and cons is more effort than many are willing to invest. As another example, Apple's 'Find My' system operates similarly to EN using encrypted Bluetooth messages and it is enabled by default on Apple devices [169].

In many EN deployments, the notification process has been limited by the availability and delay time of PCR testing. Allowing self-reported rapid antigen test results or symptoms [33] could reduce this bottleneck. The challenge with allowing self-reports is that it could enable intentional disruption by falsely reporting a positive test and notifying others. The potential damage of false reports is somewhat limited if testing rather than quarantine is recommended for those receiving a notification. Schemes that make submitting a false report more challenging could also reduce this risk, e.g. requiring a photo of the positive antigen test.

The fraction of people who choose to upload a verified positive result could be increased by encouraging pre-authorization of key upload while at the testing site. Germany's Corona-Warn-App [89] allows people to scan a QR code when tested, so that they can be notified in the app as soon as test results are available. This allows app users to choose to notify others after receiving a positive test result without the friction caused by requirements to enter an authentication code.

The fraction of infected contacts who are assessed to be high enough risk to recommend behaviour modification could be improved in three ways: 1) A lower risk threshold if community prevalence is low, as we recommend in Chapter 3 2) More information on risk (e.g. higher resolution of timing of symptoms of the index case, etc.[6]) 3) Better algorithms to assess risk based on available information (As discussed in Section 5.1)

Finally, the fraction of people who reduce further transmissions could be improved by making recommendations less costly to follow (e.g. shortening or replacing quarantine with testing [126]), or incentivizing compliance (e.g. Finland which payed people to quarantine

[6]Apple and Google currently restrict data on infectiousness to be either "none", "standard", or "high", but more resolution could enable better risk predictions and learning about transmission timing

[170]). By making one or more of these modifications, the fraction of onward transmissions prevented could be significantly improved.

6.4 Comparison to Alternative Protocols

Many other smartphone-based systems were proposed in response to the COVID-19 pandemic [45, 171, 172]. Some of these systems were widely deployed, while many others did not get the opportunity for sufficient testing[7]. Rather than reviewing all proposals, the focus of this section is to analyze whether alternative designs could improve on the failure modes discussed in Section 6.2 (and associated tradeoffs).

Singapore's TraceTogether system [46] (and a few similar designs like PEP-PT [47]) function similarly to EN, with the key difference that public health has access to the notified contacts. This may have improved compliance with (quarantine or testing) recommendations and if contact tracing capacity was available could have improved risk assessment (by combining app data with other data). My proposal as mentioned in Section 5.6 was to enhance EN with this capability by having notified devices send exposure information back to public health. This design has the important benefit over TraceTogether that it has less potential for additional tracking – TraceTogether's system could be used for passive surveillance of all citizens because the government has a mapping between the secret key being used to derive Bluetooth broadcasts and the device owner's name. With this mapping, they could use devices placed around an area to log Bluetooth data, and then use this data to assemble quite precise location logs of everyone. EN augmented with exposure sharing doesn't provide this capability because no useful data leaves the device except for when sending a notification or after receiving one – events that are rare and where the user is explicitly informed of the information being shared. The stated public preference seems to be for less invasive technology (especially anything involving the government), so while the cryptographic difference between these options may appear subtle, I believe a potential reduction in usage is not worth the risk of using TraceTogether over EN or EN augmented with exposure sharing.

Several systems were proposed that are capable of less direct risk assessment (Aarogya Setu [164], MILA [173], NOVID [174], and others), which by considering risk in a wider social network have the potential to reduce transmission with less dependence on correctly tracking a single exposure with an infectious individual. As discussed in Section 5.2, the value of a new source of information or a better predictor depends on whether it is likely

[7] In some cases, likely due to app store restrictions

to change behavior in a useful way (compared to previous information). Practically, this depends on the predictive value of infection (and potentially symptom) information in a larger contact network around an app user, and also the app user's willingness and ability to act on this information. All app-based systems will depend on at least 2 of the 6 critical steps discussed in section 6.2: the notified user must have downloaded the app and successfully modify their behavior based on information from the app so that additional transmissions are reduced. The dependence of these designs on the other 4 steps discussed (infected person has app, is tested, uploads test result, and the contact was correctly measured) is somewhat obfuscated. Clearly some information about contact events and infection status must be entered into the system in order to make useful risk assessments, but the reduction in R_t will depend on what level of risk users are willing to act on and how the disease moves through a specific contact network. In theory, even quite low accuracy risk assessments can be useful if the infection rate in the general population is low, but the trend so far in most countries has been to only modify behaviour more than others if at a significantly higher risk of infection (10%+). NOVID has the interesting property that users could be informed of their risk before being infected (if transmission is primarily occurring in a somewhat persistent contact network), which better aligns incentives for self-protection and could increase the fraction that modify their behaviour.

A challenge with these designs is that they require more trust from users in the entity controlling central servers – something that has even been lacking for the very private EN system. Data obfuscation can help somewhat, but extensive contact network information is still collected in one place for both Aarogya Setu and NOVID designs. MILA's COVI system attempts to avoid this vulnerability by using a Mix Network [173] to pass information between users, but this is quite complex to implement at scale, and if all of the message passing servers were compromised then the contact network could be assembled based on the flow of messages.

Overall, these systems could be more robust than EN to missing specific steps in the process, but whether this results in greater disease suppression depends on information content and willingness to modify behaviour at low risk. These systems have fewer built-in guarantees preventing passive surveillance, so my preference for multiple degree contact tracing is a system like the one described in Section 5.6, which shares no information about typical users, and exactly the information needed by contact tracers for those at higher risk.

Over the course of the COVID-19 pandemic it became clear that venues like indoor bars and restaurants were a frequent location for disease transmission. This likely occurred because of poor airflow, large numbers of people, and no masks (which can't be worn while eating or drinking). Many countries chose a compromise that allowed these venues

to stay open by improving contact tracing for these locations. This was done by either signing in (using pen and paper) or more quickly by using an app to scan a QR code. Several QR-based venue check-in systems were deployed, including in China, New Zealand [175], Germany and the UK. All of these systems function by getting users to scan a QR code when entering a venue – recording where and when they visited. However, some differ on where data is stored (like Bluetooth based tracing, this can be either centralized or decentralized, or some combination of the two), and how automated the notification process is (which is important for being able to handle a large number of infections). Different decision rules for notification were used in different places, but people could be notified if they had been at a venue at a time that was later marked as high risk. The efficacy of this type of system again depends on usage, testing, and behaviour modification. For locations where transmission risk is better predicted by presence in the same airspace, predictive accuracy could be better than using Bluetooth signal strength. However, the efficacy is limited by the fraction of total transmissions that occur within these venues, the willingness to notify people of appropriate risk (as discussed in Chapter 3), and the speed and scalability (if cases rise) of the process if there is a manual component to decision making. A major strength of this type of system is that there seems to be more willingness to either mandate or strongly encourage it[8] - which can significantly increase the adoption rate. For locations unwilling or unable to get people to scan QR codes, a wireless version of this system can either be implemented with specialized beacons[9] or using existing Wi-Fi routers as discussed in Section 5.6. Overall, a venue check-in system could significantly reduce disease transmission stemming from venues – especially when it is automated and set to notify people at relatively low risk (and of course people have to be willing to modify their behaviour after receiving a notification).

Most of the systems described in this section could likely independently suppress a disease like COVID-19 if used as they were designed. However, the real-world challenges are significant. Public trust in government and technology doesn't seem to be high enough to strongly support a high-tech solution - especially one that relies on aggressive behaviour modification based on unintuitive risk calculations. QR based check in systems are an important addition to EN because there is greater public support for usage and so many transmissions occur in these venues. Higher order (3+) network information seems potentially useful for risk assessment but depends on transmission dynamics and people's willingness to modify their behaviour. The main benefits of the centralized systems (on 1-2 degree contact separation) can be incorporated into EN with exposure sharing, which

[8]Potential reasons: ability to enforce usage, more limited scope, getting something in exchange, better understanding of use, attending venues is more optional

[9]Such as the patent pending system we designed [176] to notify EN users of venue-based exposures

I favour as it limits broad surveillance capabilities.

6.5 Retrospective

The initial motivations for this work still remain valid – infectious disease spread is highly heterogeneous so careful risk assessment can be used to build much more efficient methods for suppression. An ongoing challenge is to design interventions that can reliably collect and act on this information in the real world. Countries that can consistently do this can achieve disease elimination much more efficiently - as an example see Vietnam's 3rd degree tracing and year with almost no local transmission of COVID-19 [80].

For multiple reasons, tech-based targeted interventions (and also low-tech broad interventions) have not convinced a fraction of the population. Even people that agree with these approaches in principle may not choose to act on them in practice. E.g., if their phone requested that they quarantine for two weeks with little justification, and this would require that they cancel plans and not get paid over this time. The benefit to the entire population may be well worth this cost, but this is not intuitive, and they receive very little of this benefit themselves. On top of this, there may be insufficient testing to enable these interventions, or incomplete understanding of how to upload a test result or quarantine properly.

It seems quite unlikely that the current version of EN (with the current public attitude towards it) can independently suppress a high R_0 (> 2) disease. Both the fixed and variable costs of deployment are quite low compared to social distancing, so it would still be worthwhile to use in places that are actively social distancing and can expect a non-negligible reduction in Rt. Any of the modifications discussed in Section 5.6 could be incorporated to increase the success rate. Real-time information about transmission dynamics can be very valuable to decision makers and this would be possible to collect in an EN app even if only a small fraction of the population were using it. Small modifications to the system could be made to allow people to opt-in to share useful information or for privacy-preserving statistics to be generated.

A more powerful system such as the one described in Section 5.6 could reduce the dependence on infected people testing and uploading results in time by using both proximity and venue-based tracing for backward and second-degree notification. This system could also increase predictive accuracy and support for quarantine by integrating with the manual contact tracing system (which really should be supporting people through quarantine like in Finland [170]). However, this system is clearly still highly dependent on public participation in several steps. In the best case, a targeted system must have at least one point of

failure: ability to quarantine or isolate if needed to prevent a counterfactual transmission. This places an underlying limit on what can be achieved by potential interventions.

Incomplete buy-in is problematic because in a homogeneous (immune-naive) population, at the minimum $1 - 1/R_0$ fraction of the population must modify (or be ready to modify) their behaviour[10] in some way to suppress transmission ($R_t < 1$). The math gets worse if the population is heterogeneous – subpopulations with a higher concentration of people with unmodified behaviour can sustain an epidemic. In addition, behaviour modification can only be partially effective. So, if those who modify their behaviour have a reduced chance, $b \in (0, 1)$ of becoming infected and transmitting to others, then at least $(1 - 1/R_0)/(1 - b)$ people must modify their behaviour[11]. For a moderately infectious disease and optimistic value of the behaviour modification reduction, b, this implies that a very large fraction of the population must participate in order to achieve disease control (e.g., with $R_0 = 3, b = 0.2$, 83% participation is needed). The problem is made slightly easier if infection grants long-lasting immunity, but it is best to not rely on a sizeable fraction of the population getting infected.

The mechanism of effect of targeted NPIs is successful isolation of infectious individuals. A corollary of the previous argument about participation is that highly effective isolation of individuals known to be infectious is a necessary (but not sufficient) condition for elimination of highly infectious diseases with targeted NPIs. Except for in extremely pro-social populations, this means that at a minimum[12] successful elimination requires firm expectations about the behaviour of individuals that test positive. As an example, Australia in July 2020 found that 1/3 of infected individuals were not home isolating when checked [177]. Australia responded with hefty fines for breaking isolation and was able to effectively suppress transmission for almost a full year after this (September 2020 to August 2021) [177, 178, 179].

It is possible that the moderate participation observed during the COVID-19 pandemic was because for some people the perceived risk of infection with COVID-19 was not worth the perceived cost of avoiding it, and that for a more deadly infectious disease participation would be higher. However, this seems like too large of a gamble to make given the stakes. Even if it were true, there is significant benefit to developing coordinated strategies that lower the average cost of suppression such that even the more hesitant find it worthwhile to (locally) eliminate relatively less deadly diseases.

[10] Or be willing to modify their behaviour if needed, as in targeted interventions

[11] $Rt = 1 = f * b * R0 + (1 - f) * R0 - > f = (R0 - 1)/(R0 - b * R0) = (1 - 1/R0)/(1 - b)$

[12] On top of the non-trivial challenge of finding infectious individuals in the first place

6.6 Potential Design Strategy for NPIs

Except for in a few countries, at the time of this writing (July 2022) there seems to be limited social will to deploy NPIs against COVID-19 (and with vaccines and treatment available many less targeted NPIs are likely not worthwhile as long-term solutions). However, COVID-19 is likely not the last airborne pandemic, so improving the effectiveness of NPIs in advance could be enormously valuable. An advantage we have now over early 2020 is time – time to selectively develop nontrivial technology to make the response more effective.

If possible, it would be excellent to build solutions that use information close to optimally to minimize the cost of suppression (as discussed in Section 5.1). But the cost of mass infection or mass social distancing is very large, so even highly suboptimal (but reliable) methods could provide substantial benefit over the status-quo (for many but not all countries). Recognizing that insufficient public support and fragility to multiple points of failure have been major weaknesses of deployed NPIs, how can we design better NPIs?

Support by a supermajority of the population greatly increases the chances of a solution being successfully implemented in a democracy [180]. Our design strategy for digital contact tracing correctly placed significant importance on public acceptability. However, the design failed to include socially acceptable mechanisms for ensuring robustness. Public opinion seemed to support optional participation in the EN system, but making participation (app download, test upload, quarantine) mandatory was not broadly judged worthwhile (at least for COVID-19 – maybe more extreme strategies would be supported for a more deadly disease). Optional interventions (EN) and interventions primarily for self-protection (mask wearing) seem to achieve greater public support and do prevent transmissions. But these types of strategies don't reliably lead to disease elimination for highly infectious diseases – rational self-serving actors will decrease protective measures when personal risk decreases due to low infection rates [181], and some (potentially 'irrational') actors may not take protective measures at any risk.

Ideally, we can design NPIs that are both robust at any infection rate and socially acceptable (at least when faced with a deadly pandemic). Intuitively, policy backed by social expectations about the behaviour of others will be more successful. Social acceptability may be the most difficult part of the problem, so one design strategy is to start with a search for a robust social contract around control of communicable infectious diseases[13]

[13]Starting with infectious diseases that are actively being controlled. Trade-offs about behaviour around endemic diseases can be more subtle (for example, the benefit of effective isolation depends on the fraction of the population that is susceptible).

and then work backwards to build technology, infrastructure, and policy to support that 'contract'. There is precedent for building social expectations around safety- a close analogy is the standard use of drivers licenses to demonstrate proof of receiving safety training. The danger in shared spaces is high (deadly car crashes), so we expect others to put in a reasonable amount of effort to improve safety (driving training) and are willing to enforce this if necessary (legal requirement for driver's license). As a result, over 97% of drivers in the US have a driver's license[14] [182] – a level of participation that combined with good policy could suppress even a highly infectious disease.

Appendix D proposes a solution based on frequent mass-testing with the design goals of socially acceptable robustness and fast deployment in response to a novel pandemic. The system is intended to be more extensive and more strongly enforce participation than previous mass-testing deployments in democratic countries[15], while being more pleasant and offering more privacy than China's design[16]. All NPIs have some tradeoff between health, cost, and civil liberties. While a system like the one in Appendix D may seem excessive to some against COVID-19, a vast majority would likely agree that it would be worthwhile at some degree of danger. It is prudent to plan the response (and triggering conditions for that response) in advance of a deadly pandemic, especially if doing so provides more time to develop enabling technologies.

Wherever the influence on liberties is judged to be acceptable, better targeting interventions based on risk analysis can substantially reduce the cost of suppression. As an example, the number of tests needed for the previous mass-testing example could be substantially reduced by adjusting the testing frequency depending on the probability of there being cases in a certain location. The probability of a local case could be estimated using wastewater measurements and prevalence estimates of connected regions, so that in low probability regions testing frequency could be reduced to something like once a month[17] instead of once a day.

Even if not mandated, encouraging optional participation in targeted interventions like manual or digital contact tracing can be highly cost effective[18]. The success rate of these interventions can be increased by reducing friction for important steps like app download

[14] Although, unfortunately the unlicensed 3% are responsible for 20% of accidents

[15] For COVID-19, these have typically been limited to a particular school or workforce, or for a brief period of time, or are dependent on self-attestations about home test results

[16] At the time of writing (July 2022), China is using systems that are tied to identity, and often requires people to line up for less-than-pleasant nasopharyngeal swabs

[17] Reducing frequency instead of completely stopping testing will keep infrastructure active, people familiar with system, and provide another method of disease surveillance instead of completely relying on wastewater.

[18] Preferably as a lower cost addition to a robust base strategy

and sharing test results, as well as providing support and financial incentives for disruptive pro-social choices like testing and quarantine.

In a disaster scenario, both information and decision-making capacity are limited. In this context, digital contact tracing apps are more useful if they require little oversight and can provide the information needed to inform broader policy choices (e.g. when are people infectious and how long should quarantine be?). With even a small fraction of the population opting in to share more information with public health, a more powerful system like the one in Section 5.6 could provide valuable real-time information. As an example, the transmission risk associated with different types of venues could be estimated by classifying venue type based on nearby Wi-Fi routers[19]. This could be done using exposures between pairs of users, or to look for patterns correlated with infection by an unknown source (e.g. does frequent use of public transit correlate with higher risk of infection from an unknown source?). Having access to a real-time snapshot with clues about transmission dynamics would allow decision makers to focus on policy that provides the most value.

6.7 Summary

In this work we developed systems for decentralized digital contact notification and tracing that can efficiently suppress infectious disease (in the idealized high-participation case) without enabling mass surveillance. We built a framework to estimate the value of targeted interventions as a substitute for broad social distancing and used it to show that the optimal quarantine risk threshold varies significantly depending on disease prevalence. We showed how to use the joint distribution for infectiousness and test sensitivity to minimize the duration of quarantine without increasing transmissions by incorporating well-timed PCR tests. These approaches reduce the cost of NPIs by performing better risk analyses using better data. From a more general perspective, we can quantify the value of information for targeted interventions using the change in 'risk surplus' it generates and pose dynamic NPI policy decisions as an optimal control problem. Targeted interventions like EN prevented a significant number of infections during the COVID-19 pandemic [59] but failed to reach the level of robustness needed to independently achieve local disease elimination. A promising direction for future work is to develop low-cost NPIs that are both socially acceptable and robust. The goal is extraordinarily worthwhile: being able to deploy effective NPIs offers protection from future pandemics, and if made targeted enough (via either improved

[19]Potentially just based on the name of nearby networks or using more sophisticated lookup tables like the ones big tech companies have

technology or risk analysis) could enable the elimination of infectious diseases that we accept today.

References

[1] S. V. Arx, I. Becker-Mayer, D. Blank, J. Colligan, R. Fenwick, M. Hittle, M. Ingle, O. Nash, V. Nguyen, J. Petrie, J. Schwaber, Z. Szabo, A. Veeraghanta, M. Voloshin, T. White, and H. Xue, "Slowing the Spread of Infectious Diseases Using Crowdsourced Data," *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 2020.

[2] J. Petrie and J. Masel, "The economic value of quarantine is higher at lower case prevalence, with quarantine justified at lower risk of infection," *Journal of the Royal Society Interface*, vol. 18, no. 182, p. 20210459, 2021.

[3] "High contagiousness and rapid spread of Severe Acute Respiratory Syndrome Coronavirus 2," https://stacks.cdc.gov/view/cdc/90194, 2020.

[4] M. Zhang, J. Xiao, A. Deng, Y. Zhang, Y. Zhuang, T. Hu, J. Li, H. Tu, B. Li, Y. Zhou, and others, "Transmission dynamics of an outbreak of the COVID-19 Delta variant B. 1.617. 2—Guangdong Province, China, May–June 2021," *China CDC weekly*, vol. 3, no. 27, p. 584, 2021.

[5] S. I. Mallah, O. K. Ghorab, S. Al-Salmi, O. S. Abdellatif, T. Tharmaratnam, M. A. Iskandar, J. A. N. Sefen, P. Sidhu, B. Atallah, R. El-Lababidi, and M. Al-Qahtani, "COVID-19: breaking down a global health crisis," *Annals of Clinical Microbiology and Antimicrobials 2021 20:1*, vol. 20, no. 1, pp. 1–36, 5 2021. [Online]. Available: https://ann-clinmicrob.biomedcentral.com/articles/10.1186/s12941-021-00438-7

[6] M. E. Halloran, N. M. Ferguson, S. Eubank, I. M. Longini, D. A. Cummings, B. Lewis, S. Xu, C. Fraser, A. Vullikanti, T. C. Germann, D. Wagener, R. Beckman, K. Kadau, C. Barrett, C. A. Macken, D. S. Burke, and P. Cooley, "Modeling targeted layered containment of an influenza pandemic in the United States," *Proceedings of the National Academy of Sciences of the United States*

of America, vol. 105, no. 12, pp. 4639–4644, 3 2008. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0706849105

[7] L. O. Peng, S. Lim, and K. C. Suok, "Use of quarantine in the control of SARS in Singapore," *American Journal of Infection Control*, vol. 33, no. 5, pp. 252–257, 6 2005.

[8] K. T. Eames and M. J. Keeling, "Contact tracing and disease control," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. 1533, pp. 2565–2571, 12 2003. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rspb.2003.2554

[9] A. Daron, V. Chernozhukov, I. Werning, and M. Whinston, "A Multi-Risk SIR Model with Optimally Targeted Lockdown," *NBER Working Paper Series*, vol. 27102, 2020.

[10] "Lockdown Effectiveness: Much More Than You Wanted To Know." [Online]. Available: https://astralcodexten.substack.com/p/lockdown-effectiveness-much-more

[11] "On the dilemma between lives and economy - by Jan Kulveit." [Online]. Available: https://boundedlyrational.substack.com/p/on-the-dilemma-between-lives-and

[12] L. C. Evans, "An Introduction to Mathematical Optimal Control Theory Version 0.2."

[13] 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, 4 2004. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0307506101

[14] N. M. Ferguson, D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke, "Strategies for mitigating an influenza pandemic," *Nature 2006 442:7101*, vol. 442, no. 7101, pp. 448–452, 4 2006. [Online]. Available: https://www.nature.com/articles/nature04795

[15] I. o. Medicine, "Modeling Community Containment for Pandemic Influenza: A Letter Report," *Modeling Community Containment for Pandemic Influenza: A Letter Report*, pp. 1–36, 12 2006.

[16] T. C. Germann, K. Kadau, I. M. Longini, and C. A. Macken, "Mitigation strategies for pandemic influenza in the United States," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 15, pp. 5935–5940, 4 2006. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0601266103

[17] R. J. Hatchett, C. E. Mecher, and M. Lipsitch, "Public health interventions and epidemic intensity during the 1918 influenza pandemic," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 18, pp. 7582–7587, 5 2007. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0610941104

[18] H. Markel, H. B. Lipman, J. A. Navarro, A. Sloan, J. R. Michalsen, A. M. Stern, and M. S. Cetron, "Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic," *JAMA*, vol. 298, no. 6, pp. 644–654, 8 2007. [Online]. Available: https://jamanetwork.com/journals/jama/fullarticle/208354

[19] D. Klinkenberg, C. Fraser, and H. Heesterbeek, "The Effectiveness of Contact Tracing in Emerging Epidemics," *PLOS ONE*, vol. 1, no. 1, p. e12, 12 2006. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000012

[20] S. Saurabh and S. Prateek, "Role of contact tracing in containing the 2014 Ebola outbreak: a review," *African Health Sciences*, vol. 17, no. 1, pp. 225–236, 5 2017. [Online]. Available: https://www.ajol.info/index.php/ahs/article/view/156380

[21] M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones, "A high-resolution human contact network for infectious disease transmission," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 51, pp. 22 020–22 025, 12 2010. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1009094108

[22] E. Yoneki, "FluPhone study: Virtual disease spread using haggle," *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, pp. 65–66, 2011. [Online]. Available: http://www.cl.cam.ac.uk/research/srg/netos/fluphone2/.

[23] K. Farrahi, R. Emonet, and M. Cebrian, "Epidemic Contact Tracing via Communication Traces," *PLOS ONE*, vol. 9, no. 5, p. e95133, 5 2014. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0095133

[24] K. A. Nguyen, Z. Luo, and C. Watkins, "On the Feasibility of Using Two Mobile Phones and WLAN Signal to Detect Co-Location of Two Users for Epidemic Prediction," pp. 63–78, 2015. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-11879-6_5

[25] A. Prasad and D. Kotz, "ENACT: Encounter-based architecture for contact tracing," *WPA 2017 - Proceedings of the 4th International Workshop on Physical Analytics, co-located with MobiSys 2017*, pp. 37–42, 6 2017. [Online]. Available: http://dx.doi.org/10.1145/3092305.3092310

[26] A. Prasad, "Privacy-preserving controls for sharing mHealth data," *Dartmouth College Ph.D Dissertations*, 5 2016. [Online]. Available: https://digitalcommons.dartmouth.edu/dissertations/49

[27] 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, S. Flasche, B. J. Quilty, N. Davies, Y. Liu, S. Clifford, P. Klepac, M. Jit, C. Diamond, H. Gibbs, K. van Zandvoort, S. Funk, and R. M. Eggo, "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," *The Lancet Global Health*, vol. 8, no. 4, 2020.

[28] 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*, vol. 368, no. 6491, 2020.

[29] A. Di Luzio, A. Mei, and J. Stefa, "Mind your probes: De-anonymization of large crowds through smartphone WiFi probe requests," in *Proceedings - IEEE INFOCOM*, vol. 2016-July, 2016.

[30] J. Freudiger, "Short: How talkative is your mobile device? An experimental study of Wi-Fi probe requests," in *Proceedings of the 8th ACM Conference on Security and Privacy in Wireless and Mobile Networks, WiSec 2015*, 2015.

[31] M. Génois, C. L. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, and A. Barrat, "Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers," *Network Science*, vol. 3, no. 3, 2015.

[32] T. White, "COVID-19 Risk Assessment App Idea for Vetting and Discussion - EA Forum." [Online]. Available: https://forum.effectivealtruism.org/posts/8chk6DHZXctGHtNoz/covid-19-risk-assessment-app-idea-for-vetting-and-discussion

[33] "CoEpi." [Online]. Available: https://github.com/Co-Epi

[34] S. Von Arx, I. Becker-Mayer, D. Blank, J. Colligan, R. Fenwick, M. Hittle, M. Ingle, O. Nash, V. Nguyen, J. Petrie, J. Schwaber, Z. Szabo, A. Veeraghanta, M. Voloshin, T. White, and H. Xue, "Slowing the spread

of infectious diseases using crowdsourced data," https://blog.covidwatch.org/en/covid-watch-whitepaper-using-crowdsourced-data-to-slow-virus-spread, 2020, accessed 25 April 2021.

[35] "Inside China's All-Out War on the Coronavirus - The New York Times." [Online]. Available: https://www.nytimes.com/2020/03/04/health/coronavirus-china-aylward.html

[36] D. Normile, "Coronavirus cases have dropped sharply in South Korea. What's the secret to its success?" *Science*, 3 2020.

[37] N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. C. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Gt Walker, C. Walters, P. Winskill, C. Whittaker, C. A. Donnelly, S. Riley, and A. C. Ghani, "Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand." [Online]. Available: https://doi.org/10.25561/77482.

[38] "covid-19_instant_tracing/Manuscript - Modelling instantaneous digital contact tracing.pdf at master · BDI-pathogens/covid-19_instant_tracing · GitHub." [Online]. Available: https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Manuscript%20-%20Modelling%20instantaneous%20digital%20contact%20tracing.pdf

[39] "Smartphone subscriptions worldwide — Statista." [Online]. Available: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

[40] "Smartphone penetration in the US (share of population) 2010-2021 — Statista." [Online]. Available: https://www.statista.com/statistics/201183/forecast-of-smartphone-penetration-in-the-us/

[41] "TCNCoalition/TCN: Specification and reference implementation of the TCN Protocol for decentralized, privacy-preserving contact tracing." [Online]. Available: https://github.com/TCNCoalition/TCN

[42] "GitHub - covidwatchorg/covidwatch-ios-tcn: Covid Watch iOS TCN app." [Online]. Available: https://github.com/covidwatchorg/covidwatch-ios-tcn

[43] "GitHub - covidwatchorg/covidwatch-android-tcn: Covid Watch Android TCN app." [Online]. Available: https://github.com/covidwatchorg/covidwatch-android-tcn

[44] "COVID Watch: Bluetooth Contact Tracing Functionality - YouTube." [Online]. Available: https://www.youtube.com/watch?v=4aQ90emoAqc

[45] N. Ahmed, R. A. Michelin, W. Xue, S. Ruj, R. Malaney, S. S. Kanhere, A. Seneviratne, W. Hu, H. Janicke, and S. K. Jha, "A Survey of COVID-19 Contact Tracing Apps," *IEEE Access*, vol. 8, pp. 134 577–134 601, 2020.

[46] "TraceTogether." [Online]. Available: https://www.tracetogether.gov.sg/

[47] "pepp-pt/pepp-pt-documentation: Documentation for Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT)." [Online]. Available: https://github.com/pepp-pt/pepp-pt-documentation

[48] "documents/DP3T White Paper.pdf at master · DP-3T/documents." [Online]. Available: https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

[49] J. Chan, D. Foster, S. Gollakota, E. Horvitz, J. Jaeger, S. Kakade, T. Kohno, J. Langford, J. Larson, P. Sharma, S. Singanamalla, J. Sunshine, and S. Tessaro, "PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing," 4 2020. [Online]. Available: https://arxiv.org/abs/2004.03544v4

[50] "ito-org/STRICT: STRICT [simply trace infections] is a protocol and concept of how to anonymously track infections without tracking people." [Online]. Available: https://github.com/ito-org/STRICT

[51] L. Loiseau, V. Bellet, T. S. Bento, E. Teissonniere, M. Benoliel, G. Kinsman, and P. Milne, "Whisper Tracing Version 3 - An Open and Privacy First Protocol for Contact Tracing." [Online]. Available: https://nodle.docsend.com/view/nis3dac

[52] "NVD - CVE-2020-12856." [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2020-12856

[53] "TCN_Cost.ipynb - Colaboratory." [Online]. Available: https://colab.research.google.com/drive/1x_KadfmNXiAEaraGbGW7qWipx5KuKywk#scrollTo=EzfyxzveXPWT

[54] R. L. Rivest, J. Callas, R. Canetti, K. Esvelt, D. K. Gillmor, A. Civil, L. Union, Y. T. Kalai, A. Lysyanskaya, A. Norige, R. Raskar, A. Shamir, E. Shen, I. Soibelman, M. Specter, V. Teague, T. Cybersecurity, A. Trachtenberg, M. Varia, M. Viera, and D. Weitzner, "The PACT protocol specification," *The PACT protocol specification*, vol. 1, 2020.

[55] "Exposure Notification — Apple Developer Documentation." [Online]. Available: https://developer.apple.com/documentation/exposurenotification

[56] "Exposure Notifications: Helping fight COVID-19 - Google." [Online]. Available: https://www.google.com/covid19/exposurenotifications/

[57] "Why Apple and Google's Virus Alert Apps Had Limited Success - The New York Times." [Online]. Available: https://www.nytimes.com/2021/05/27/business/apple-google-virus-tracing-app.html

[58] "Publicly-available Exposure Notifications apps — Google API for Exposure Notifications — Google Developers." [Online]. Available: https://developers.google.com/android/exposure-notifications/apps

[59] C. Wymant, L. Ferretti, D. Tsallis, M. Charalambides, L. Abeler-Dörner, D. Bonsall, R. Hinch, M. Kendall, L. Milsom, M. Ayres, C. Holmes, M. Briers, and C. Fraser, "The epidemiological impact of the NHS COVID-19 app," *Nature*, vol. 594, no. 7863, 2021.

[60] A. Brodeur, D. Gray, A. Islam, and S. J. Bhuiyan, "A literature review of the economics of COVID-19," Institute of Labor Economics (IZA), IZA Discussion Papers 13411, 2020. [Online]. Available: https://EconPapers.repec.org/RePEc:iza:izadps:dp13411

[61] O. Yakusheva, E. van den Broek-Altenburg, G. Brekke, and A. Atherly, "The cure is not worse than the disease-a humanitarian perspective," *Available at SSRN 3638575*, 2020.

[62] Z. A. Bethune and A. Korinek, "Covid-19 infection externalities: Trading off lives vs. livelihoods," National Bureau of Economic Research, Tech. Rep. 27009, 2020.

[63] H. d'Albis and E. Augeraud-Véron, "Optimal prevention and elimination of infectious diseases," *Journal of Mathematical Economics*, vol. 93, p. 102487, 2021.

[64] D. Acemoglu, V. Chernozhukov, I. Werning, and M. D. Whinston, "Optimal targeted lockdowns in a multi-group sir model," *NBER Working Paper*, vol. 27102, 2020.

[65] Y. Bar-On, T. Baron, O. Cornfeld, R. Milo, E. Yashiv *et al.*, "Caveats for economists: Epidemiology-based modelling of covid19 and model misspecifications," CEPR Discussion Papers, Tech. Rep., 2020.

[66] M. N. Lurie, J. Silva, R. R. Yorlets, J. Tao, and P. A. Chan, "Coronavirus disease 2019 epidemic doubling time in the United States before and during stay-at-home restrictions," *The Journal of Infectious Diseases*, vol. 222, no. 10, pp. 1601–1606, 08 2020. [Online]. Available: https://doi.org/10.1093/infdis/jiaa491

[67] Y. Gu, "Covid-19 projections using machine learning," https://covid19-projections.com/, 2020.

[68] J. S. Weitz, S. W. Park, C. Eksin, and J. Dushoff, "Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations," *Proceedings of the National Academy of Sciences*, vol. 117, no. 51, pp. 32 764–32 771, 2020.

[69] A. V. Tkachenko, S. Maslov, A. Elbanna, G. N. Wong, Z. J. Weiner, and N. Goldenfeld, "Time-dependent heterogeneity leads to transient suppression of the covid-19 epidemic, not herd immunity," *Proceedings of the National Academy of Sciences*, vol. 118, no. 17, 2021. [Online]. Available: https://www.pnas.org/content/118/17/e2015972118

[70] M. S. Eichenbaum, S. Rebelo, and M. Trabandt, "The macroeconomics of testing and quarantining," National Bureau of Economic Research, Tech. Rep. 27104, 2020.

[71] A. M. Wilson, N. Aviles, J. I. Petrie, P. I. Beamer, Z. Szabo, M. Xie, J. McIllece, Y. Chen, Y.-J. Son, S. Halai, T. White, K. C. Ernst, and J. Masel, "Quantifying sars-cov-2 infection risk within the Google/Apple exposure notification framework to inform quarantine recommendations," 2020, https://www.medrxiv.org/content/early/2020/09/17/2020.07.17.20156539.

[72] V. A. Karatayev, M. Anand, and C. T. Bauch, "Local lockdowns outperform global lockdown on the far side of the covid-19 epidemic curve," *Proceedings of the National Academy of Sciences*, vol. 117, no. 39, pp. 24 575–24 580, 2020. [Online]. Available: https://www.pnas.org/content/117/39/24575

[73] R. M. Anderson and R. M. May, *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[74] C. R. Wells, J. P. Townsend, A. Pandey, G. Krieger, B. Singer, R. H. McDonald, S. M. Moghadas, and A. P. Galvani, "Optimal covid-19 quarantine and testing strategies," *medRxiv*, 2020, https://www.medrxiv.org/content/early/2020/11/08/2020.10.27.20211631.

[75] "Pandemic navigator," https://pandemicnavigator.oliverwyman.com/forecast?mode=country&region=United%20States&panel=mortality, 2020.

[76] Y. J. Park, Y. J. Choe, O. Park, S. Y. Park, Y.-M. Kim, J. Kim, S. Kweon, Y. Woo, J. Gwack, S. S. Kim *et al.*, "Contact tracing during coronavirus disease outbreak, south korea, 2020," *Emerging infectious diseases*, vol. 26, no. 10, pp. 2465–2468, 2020.

[77] R. Laxminarayan, B. Wahl, S. R. Dudala, K. Gopal, C. Mohan B, S. Neelima, K. S. Jawahar Reddy, J. Radhakrishnan, and J. A. Lewnard, "Epidemiology and transmission dynamics of COVID-19 in two Indian states," *Science*, vol. 370, no. 6517, pp. 691–697, 2020. [Online]. Available: https://science.sciencemag.org/content/370/6517/691

[78] W. Chen, N. Zhang, J. Wei, H.-L. Yen, and Y. Li, "Short-range airborne route dominates exposure of respiratory infection during close contact," *Building and Environment*, vol. 176, p. 106859, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360132320302183

[79] "3.5 million people who arrived in canada since march exempted from quarantine requirement," https://globalnews.ca/news/7416775/canada-coronavirus-border-quarantine/, 2020.

[80] "Emerging COVID-19 success story: Vietnam's commitment to containment," https://ourworldindata.org/covid-exemplar-vietnam, 2020, accessed: 2021-03-30.

[81] N. Wetsman, "FDA authorizes Abbott's fast $5 COVID-19 test," https://www.theverge.com/2020/8/26/21403432/fda-authorizes-binaxnow-covid-19-test-abbott-cheap-fast, Aug 2020.

[82] Z. Du, A. Pandey, Y. Bai, M. C. Fitzpatrick, M. Chinazzi, A. P. y Piontti, M. Lachmann, A. Vespignani, B. J. Cowling, A. P. Galvani *et al.*, "Comparative cost-effectiveness of SARS-CoV-2 testing strategies in the USA: a modelling study," *The Lancet Public Health*, vol. 6, no. 3, pp. e184–e191, 2021.

[83] "Contact tracing in the context of covid-19," https://www.who.int/publications/i/item/contact-tracing-in-the-context-of-covid-19, 2021, accessed: 2021-05-11.

[84] "Investigating a covid-19 case," https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/investigating-covid-19-case.html, 2021, accessed: 2021-05-11.

[85] L. Ferretti, A. Ledda, C. Wymant, L. Zhao, V. Ledda, L. Abeler-Dorner, M. Kendall, A. Nurtay, H.-Y. Cheng, T.-C. Ng, H.-H. Lin, R. Hinch, J. Masel, A. M. Kilpatrick, and C. Fraser, "The timing of covid-19 transmission," https://www.medrxiv.org/content/early/2020/09/16/2020.09.04.20188516.full.pdf.

[86] P. Ashcroft, S. Lehtinen, D. C. Angst, N. Low, and S. Bonhoeffer, "Quantifying the impact of quarantine duration on covid-19 transmission," *Elife*, vol. 10, p. e63704, 2021.

[87] X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan *et al.*, "Author correction: Temporal dynamics in viral shedding and transmissibility of covid-19," *Nature medicine*, vol. 26, no. 9, pp. 1491–1493, 2020.

[88] M. Briers, M. Charalambides, and C. Holmes, "Risk scoring calculation for the current NHSx contact tracing app," https://arxiv.org/abs/2005.11057, May 2020.

[89] Corona-Warn-App, "How does the corona-warn-app identify an increased risk?" https://github.com/corona-warn-app/cwa-documentation/blob/master/cwa-risk-assessment.md, 2020.

[90] T. C. Bulfone, M. Malekinejad, G. W. Rutherford, and N. Razani, "Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review," *The Journal of infectious diseases*, vol. 223, no. 4, pp. 550–561, 2021.

[91] R. C. Dorf and R. H. Bishop, "Modern control systems," 1998.

[92] "Arizona ended covid restrictions because hospitals have bed space, health chief says," https://tucson.com/news/local/arizona-ended-covid-restrictions-because-hospitals-have-bed-space-health-chief-says/article_c35f294e-4cd8-5ca3-b6c9-abc61b15ec4b.html, 2020, accessed: 2021-03-30.

[93] "Herd immunity: will the UK's coronavirus strategy work?" https://www.theguardian.com/world/2020/mar/13/herd-immunity-will-the-uks-coronavirus-strategy-work, 2020, accessed: 2021-05-13.

[94] S. Otto, "The tipping point of contact tracing," https://covidmodelling.pwias.ubc.ca/blog/, 2020.

[95] "CDC: Viral Testing Tool." [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/testing/viral-testing-tool.html

[96] B. Adamson, R. Sikka, A. L. Wyllie, and P. Premsrirut, "Discordant SARS-CoV-2 PCR and Rapid Antigen Test Results When Infectious: A December 2021 Occupational Case Series," *medRxiv*, 2022. [Online]. Available: https://www.medrxiv.org/content/early/2022/01/05/2022.01.04.22268770

[97] C. Del Vecchio, G. Brancaccio, A. R. Brazzale, E. Lavezzo, F. Onelia, E. Franchin, L. Manuto, F. Bianca, V. Cianci, A. Cattelan, S. Toppo, and A. Crisanti, "Emergence of N antigen SARS-CoV-2 genetic variants escaping detection of antigenic tests," *medRxiv*, 2021.

[98] C. N. Haas, "On the quarantine period for Ebola virus," *PLoS Currents*, vol. 6, 2014.

[99] B. Ottino-Loffler, J. G. Scott, and S. H. Strogatz, "Evolutionary dynamics of incubation periods," *Elife*, vol. 6, p. e30212, 2017.

[100] S. K. Brooks, R. K. Webster, L. E. Smith, L. Woodland, S. Wessely, N. Greenberg, and G. J. Rubin, "The psychological impact of quarantine and how to reduce it: rapid review of the evidence," *The Lancet*, 2020.

[101] F. A. Soud, M. M. Cortese, A. T. Curns, P. J. Edelson, R. H. Bitsko, H. T. Jordan, A. S. Huang, J. M. Villalon-Gomez, and G. H. Dayan, "Isolation compliance among university students during a mumps outbreak, Kansas 2006," *Epidemiology & Infection*, vol. 137, no. 1, pp. 30–37, 2009.

[102] L. E. Smith, H. W. W. Potts, R. Amlot, N. T. Fear, S. Michie, and J. Rubin, "Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study)," *medRxiv*, 2020.

[103] A. Steens, B. F. de Blasio, L. Veneti, A. Gimma, W. J. Edmunds, K. Van Zandvoort, C. I. Jarvis, F. Forland, and B. Robberstad, "Poor self-reported adherence to COVID-19-related quarantine/isolation requests, Norway, April to July 2020," *Eurosurveillance*, vol. 25, no. 37, p. 2001607, 2020.

[104] "Options to Reduce Quarantine for Contacts of Persons with SARS-CoV-2 Infection Using Symptom Monitoring and Diagnostic Testing." [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-options-to-reduce-quarantine.html

[105] W. H. Organization and others, "Considerations for quarantine of contacts of COVID-19 cases: interim guidance, 25 June 2021," World Health Organization, Tech. Rep., 2021.

[106] "Options to reduce quarantine for contacts of persons with sars-cov-2 infection using symptom monitoring and diagnostic testing," accessed 11 January 2022. [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-options-to-reduce-quarantine.html

[107] C. R. Wells, J. P. Townsend, A. Pandey, S. M. Moghadas, G. Krieger, B. Singer, R. H. McDonald, M. C. Fitzpatrick, and A. P. Galvani, "Optimal covid-19 quarantine and testing strategies," *Nature Communications*, vol. 12, no. 1, pp. 1–9, 2021.

[108] A. Steens, B. F. de Blasio, L. Veneti, A. Gimma, W. J. Edmunds, K. Van Zandvoort, C. I. Jarvis, F. Forland, and B. Robberstad, "Poor self-reported adherence to covid-19-related quarantine/isolation requests, norway, april to july 2020," *Eurosurveillance*, vol. 25, no. 37, p. 2001607, 2020.

[109] L. E. Smith, H. W. Potts, R. Amlot, N. T. Fear, S. Michie, and J. Rubin, "Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the uk (the covid-19 rapid survey of adherence to interventions and responses [corsair] study)," *medRxiv*, 2020.

[110] M. A. Johansson, H. Wolford, P. Paul, P. S. Diaz, T.-H. Chen, C. M. Brown, M. S. Cetron, and F. Alvarado-Ramy, "Reducing travel-related sars-cov-2 transmission with layered mitigation measures: Symptom monitoring, quarantine, and testing," *medRxiv*, 2020.

[111] S. Clifford, B. J. Quilty, T. W. Russell, Y. Liu, Y.-W. D. Chan, C. A. Pearson, R. M. Eggo, A. Endo, S. Flasche, W. J. Edmunds *et al.*, "Strategies to reduce the risk of sars-cov-2 re-introduction from international travellers," *medRxiv*, 2020.

[112] L. M. Kucirka, S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler, "Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based sars-cov-2 tests by time since exposure," *Annals of Internal Medicine*, 2020.

[113] K. Danis, O. Epaulard, T. Bénet, A. Gaymard, S. Campoy, E. Bothelo-Nevers, M. Bouscambert-Duchamp, G. Spaccaferri, F. Ader, A. Mailles *et al.*, "Cluster of coronavirus disease 2019 (covid-19) in the french alps, 2020," *Clinical Infectious Diseases*, 2020.

[114] B. J. Quilty, S. Clifford, J. Hellewell, T. W. Russell, A. J. Kucharski, S. Flasche, W. J. Edmunds, K. E. Atkins, A. M. Foss, N. R. Waterlow *et al.*, "Quarantine and testing strategies in contact tracing for sars-cov-2: a modelling study," *The Lancet Public Health*, vol. 6, no. 3, pp. e175–e183, 2021.

[115] L. Ferretti, C. Wymant, A. Nurtay, L. Zhao, R. Hinch, D. Bonsall, M. Kendall, J. Masel, J. Bell, S. Hopkins *et al.*, "Modelling the effectiveness and social costs of daily lateral flow antigen tests versus quarantine in preventing onward transmission of covid-19 from traced contacts," *medRxiv*, 2021.

[116] C. Atherstone, M. L. Peterson, M. Malone, M. A. Honein, A. MacNeil, C. S. O'Neal, S. Paul, K. G. Harmon, K. Goerl, C. R. Wolfe *et al.*, "Time from start of quarantine to sars-cov-2 positive test among quarantined college and university athletes—17 states, june–october 2020," *Morbidity and Mortality Weekly Report*, vol. 70, no. 1, p. 7, 2021.

[117] J. Hellewell, T. W. Russell, R. Beale, G. Kelly, C. Houlihan, E. Nastouli, A. J. Kucharski, S. Investigators, F. study team, C. C.-. Consortium *et al.*, "Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections," *medRxiv*, 2020.

[118] A. M. Wilson, N. Aviles, J. I. Petrie, P. I. Beamer, Z. Szabo, M. Xie, J. McIllece, Y. Chen, Y.-J. Son, S. Halai *et al.*, "Quantifying sars-cov-2 infection risk within the google/apple exposure notification framework to inform quarantine recommendations," *Risk Analysis*, 2021.

[119] "Cdc: Test for current infection." [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/testing/diagnostic-testing.html

[120] M. Zhang, J. Xiao, A. Deng, Y. Zhang, Y. Zhuang, T. Hu, J. Li, H. Tu, B. Li, Y. Zhou *et al.*, "Transmission dynamics of an outbreak of the covid-19 delta variant b. 1.617. 2—guangdong province, china, may–june 2021," *China CDC weekly*, vol. 3, no. 27, p. 584, 2021.

[121] "Uk government: Tests required before travel to uk and nigeria added to red list." [Online]. Available: https://www.gov.uk/government/news/tests-required-before-travel-to-uk-and-nigeria-added-to-red-list

[122] C. R. Wells, J. P. Townsend, A. Pandey, S. M. Moghadas, G. Krieger, B. Singer, R. H. McDonald, M. C. Fitzpatrick, and A. P. Galvani, "Optimal COVID-19 quarantine and testing strategies," *Nature Communications*, vol. 12, no. 1, pp. 1–9, 2021.

[123] B. J. Quilty, S. Clifford, J. Hellewell, T. W. Russell, A. J. Kucharski, S. Flasche, W. J. Edmunds, K. E. Atkins, A. M. Foss, N. R. Waterlow, and others, "Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study," *The Lancet Public Health*, vol. 6, no. 3, p. e175–e183, 2021.

[124] M. A. Johansson, H. Wolford, P. Paul, P. S. Diaz, T.-H. Chen, C. M. Brown, M. S. Cetron, and F. Alvarado-Ramy, "Reducing travel-related SARS-CoV-2 transmission with layered mitigation measures: Symptom monitoring, quarantine, and testing," *medRxiv*, 2020.

[125] L. M. Kucirka, S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler, "Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure," *Annals of Internal Medicine*, vol. 173, no. 4, 2020.

[126] L. Ferretti, C. Wymant, A. Nurtay, L. Zhao, R. Hinch, D. Bonsall, M. Kendall, J. Masel, J. Bell, S. Hopkins, A. M. Kilpatrick, T. Peto, L. Abeler-Dörner, and C. Fraser, "Modelling the effectiveness and social costs of daily lateral flow antigen tests versus quarantine in preventing onward transmission of COVID-19 from traced contacts," *medRxiv*, p. 2021.08.06.21261725, 8 2021. [Online]. Available: https://www.medrxiv.org/content/10.1101/2021.08.06.21261725v1https://www.medrxiv.org/content/10.1101/2021.08.06.21261725v1.abstract

[127] M. U. G. Kraemer, O. G. Pybus, C. Fraser, S. Cauchemez, A. Rambaut, and B. J. Cowling, "Monitoring key epidemiological parameters of SARS-CoV-2 transmission," *Nature medicine*, vol. 27, no. 11, pp. 1854–1855, 2021.

[128] J. Petrie and J. Masel, "The economic value of quarantine is higher at lower case prevalence, with quarantine justified at lower risk of infection," *Journal of the Royal Society Interface*, vol. 18, no. 182, p. 20210459, 2021.

[129] A. M. Wilson, N. Aviles, J. I. Petrie, P. I. Beamer, Z. Szabo, M. Xie, J. McIllece, Y. Chen, Y. J. Son, S. Halai, T. White, K. C. Ernst, and J. Masel, "Quantifying

SARS-CoV-2 Infection Risk Within the Google/Apple Exposure Notification Framework to Inform Quarantine Recommendations," *Risk Analysis*, vol. 42, no. 1, 2022.

[130] P. Singer and J. Masel, "How (not) to fight COVID 19. Project Syndicate," 2020.

[131] C. McAloon, Collins, K. Hunt, A. Barber, A. W. Byrne, F. Butler, M. Casey, J. Griffin, E. Lane, D. McEvoy, and others, "Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research," *BMJ open*, vol. 10, no. 8, p. e039652, 2020.

[132] "How to Determine a Close Contact for COVID-19 — CDC." [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/determine-close-contacts.html

[133] P. Health Ontario, "Risk Assessment Approach for COVID-19 Contact Tracing."

[134] "SwissCovid App – Staffnet — ETH Zurich." [Online]. Available: https://ethz.ch/staffnet/en/news-and-events/solidarity/pilot-swiss-covid-app.html

[135] L. Luo, D. Liu, X. Liao, X. Wu, Q. Jing, J. Zheng, F. Liu, S. Yang, H. Bi, Z. Li, J. Liu, W. Song, W. Zhu, Z. Wang, X. Zhang, Q. Huang, P. Chen, H. Liu, X. Cheng, M. Cai, P. Yang, X. Yang, Z. Han, J. Tang, Y. Ma, and C. Mao, "Contact settings and risk for transmission in 3410 close contacts of patients with COVID-19 in Guangzhou, China a prospective cohort study," *Annals of Internal Medicine*, vol. 173, no. 11, 2020.

[136] S. Muthaiyah, T. O. K. Zaw, K. S. M. Anbananthen, B. Park, and M. J. Kim, "Data Driven Models for Contact Tracing Prediction: A Systematic Review of COVID-19," *Emerging Science Journal*, vol. 7, pp. 17–28, 9 2022.

[137] H. Yoshikawa, "Can naive Bayes classifier predict infection in a close contact of COVID-19? A comparative test for predictability of the predictive model and healthcare workers in Japan," *Journal of Infection and Chemotherapy*, vol. 28, no. 6, pp. 774–779, 6 2022.

[138] Z. H. Hoo, J. Candlish, and D. Teare, "What is an ROC curve?" *Emergency Medicine Journal*, vol. 34, no. 6, pp. 357–359, 6 2017. [Online]. Available: https://emj.bmj.com/content/34/6/357https://emj.bmj.com/content/34/6/357.abstract

[139] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning Data Mining, Inference, and Prediction (12th printing)*, 2017.

[140] J. Bennetto, "Should ROC curve be concave? - Quora." [Online]. Available: https://www.quora.com/Should-ROC-curve-be-concave-when-looking-from-downside

[141] R. H. Frank, *Principles of microeconomics*. , 2004.

[142] M. DASH and H. LIU, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, no. 1-4, pp. 131–156, 1 1997. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1088467X97000085

[143] J. Tang, S. Alelyani, and H. Liu, "Feature Selection for Classification: A Review."

[144] O. Mohr, J. Hermes, S. B. Schink, M. Askar, D. Menucci, C. Swaan, U. Goetsch, P. Monk, T. Eckmanns, G. Poggensee, and G. Krause, "Development of a risk assessment tool for contact tracing people after contact with infectious patients while travelling by bus or other public ground transport: a Delphi consensus approach," *BMJ Open*, vol. 3, no. 10, p. e002939, 10 2013. [Online]. Available: https://bmjopen.bmj.com/content/3/10/e002939https://bmjopen.bmj.com/content/3/10/e002939.abstract

[145] A. Fox-Lewis, F. Williamson, J. Harrower, X. Ren, G. J. Sonder, A. McNeill, J. de Ligt, and J. L. Geoghegan, "Airborne Transmission of SARS-CoV-2 Delta Variant within Tightly Monitored Isolation Facility, New Zealand (Aotearoa)," *Emerging Infectious Diseases*, vol. 28, no. 3, 2022.

[146] F. Lin, K. Muthuraman, and M. Lawley, "An optimal control theory approach to non-pharmaceutical interventions," *BMC Infectious Diseases*, vol. 10, no. 1, pp. 1–13, 2 2010. [Online]. Available: https://link.springer.com/articles/10.1186/1471-2334-10-32https://link.springer.com/article/10.1186/1471-2334-10-32

[147] R. Djidjou-Demasse, Y. Michalakis, M. Choisy, M. T. Sofonea, and S. Alizon, "Optimal COVID-19 epidemic control until vaccine deployment." [Online]. Available: https://doi.org/10.1101/2020.04.02.20049189

[148] T. A. Perkins and G. España, "Optimal Control of the COVID-19 Pandemic with Non-pharmaceutical Interventions," *Bulletin of Mathematical Biology*, vol. 82, no. 9, pp. 1–24, 9 2020. [Online]. Available: https://link.springer.com/article/10.1007/s11538-020-00795-y

[149] L. J. Allen and A. M. Burgin, "Comparison of deterministic and stochastic SIS and SIR models in discrete time," *Mathematical Biosciences*, vol. 163, no. 1, pp. 1–33, 1 2000.

[150] J. Wallinga and M. Lipsitch, "How generation intervals shape the relationship between growth rates and reproductive numbers," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1609, 2007.

[151] T. D. Hollingsworth, N. M. Ferguson, and R. M. Anderson, "Will travel restrictions control the international spread of pandemic influenza?" *Nature Medicine 2006 12:5*, vol. 12, no. 5, pp. 497–499, 5 2006. [Online]. Available: https://www.nature.com/articles/nm0506-497

[152] G. Scalia Tomba and J. Wallinga, "A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease," *Mathematical Biosciences*, vol. 214, no. 1-2, pp. 70–72, 7 2008.

[153] J. Summers, H.-Y. Cheng, H.-H. Lin, L. T. Barnard, A. Kvalsvig, N. Wilson, and M. G. Baker, "Potential lessons from the Taiwan and New Zealand health responses to the COVID-19 pandemic," *The Lancet Regional Health - Western Pacific*, vol. 4, p. 100044, 11 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2666606520300444

[154] A. Caparrós and M. Finus, "The Corona-Pandemic: A Game-Theoretic Perspective on Regional and Global Governance," *Environmental and Resource Economics*, vol. 76, no. 4, 2020.

[155] S. Cauchemez and C. T. Kiem, "Managing COVID-19 importation risks in a heterogeneous world," *The Lancet Public Health*, vol. 6, no. 9, pp. e626–e627, 9 2021.

[156] M. J. Plank, "Minimising the use of costly control measures in an epidemic elimination strategy: A simple mathematical model," *Mathematical Biosciences*, vol. 351, p. 108885, 9 2022.

[157] "NSW toughens Covid rules for airport transport workers as police seek advice on Sydney limo driver — Australian police and policing — The Guardian." [Online]. Available: https://www.theguardian.com/australia-news/2021/jun/25/sydney-limo-driver-police-seek-urgent-legal-advice-if-bondi-limousine-breached-covid-orders

[158] "DHS to Conduct Airflow Studies in NYC to Help Urban Areas Better Plan for Airborne Threats - Global Biodefense." [Online]. Available: https://globalbiodefense.com/2021/10/14/dhs-to-conduct-airflow-studies-in-nyc-to-help-urban-areas-better-plan-for-airborne-threats/

[159] T. L. Dao, T. D. Nguyen, and V. T. Hoang, "Controlling the COVID-19 pandemic: Useful lessons from Vietnam," *Travel Medicine and Infectious Disease*, vol. 37, p. 101822, 9 2020. [Online]. Available: /pmc/articles/PMC7347475/https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7347475/

[160] Y. Furuse, N. Tsuchiya, R. Miyahara, I. Yasuda, E. Sando, Y. K. Ko, T. Imamura, K. Morimoto, T. Imamura, Y. Shobugawa, S. Nagata, A. Tokumoto, K. Jindai, M. Suzuki, and H. Oshitani, "COVID-19 case-clusters and transmission chains in the communities in Japan," *Journal of Infection*, vol. 84, no. 2, pp. 248–288, 2 2022. [Online]. Available: http://www.journalofinfection.com/article/S0163445321003996/fulltexthttp://www.journalofinfection.com/article/S0163445321003996/abstracthttps://www.journalofinfection.com/article/S0163-4453(21)00399-6/abstract

[161] A. Baker, I. Biazzo, A. Braunstein, G. Catania, L. Dall'Asta, A. Ingrosso, F. Krzakala, F. Mazza, M. Mézard, A. P. Muntoni, M. Refinetti, S. S. Mannelli, and L. Zdeborová, "Epidemic mitigation by statistical inference from contact tracing data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 32, 8 2021.

[162] W. J. Bradshaw, E. C. Alley, J. H. Huggins, A. L. Lloyd, and K. M. Esvelt, "Bidirectional contact tracing could dramatically improve COVID-19 control," *Nature Communications 2021 12:1*, vol. 12, no. 1, pp. 1–9, 1 2021. [Online]. Available: https://www.nature.com/articles/s41467-020-20325-7

[163] "Viratrace." [Online]. Available: https://www.viratrace.org/#/

[164] "Aarogya Setu Mobile App — MyGov.in." [Online]. Available: https://www.mygov.in/aarogya-setu-app/

[165] "Covid-19: My Current Model — Don't Worry About the Vase." [Online]. Available: https://thezvi.wordpress.com/2020/05/31/covid-19-my-current-model/

[166] T. Fetzer and T. Graeber, "Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 33, 2021.

[167] "Coronavirus: How to Do Testing and Contact Tracing — by Tomas Pueyo — Medium." [Online]. Available: https://tomaspueyo.medium.com/coronavirus-how-to-do-testing-and-contact-tracing-bde85b64072e

[168] E. J. Johnson and D. Goldstein, "Do Defaults Save Lives?" *Science*, vol. 302, no. 5649, pp. 1338–1339, 11 2003. [Online]. Available: https://www.science.org/doi/10.1126/science.1091721

[169] "Find My App: Everything to Know - MacRumors." [Online]. Available: https://www.macrumors.com/guide/find-my/

[170] "Sickness allowance on account of an infectious disease - kela.fi." [Online]. Available: https://www.kela.fi/web/en/sickness-allowances-infectious-disease

[171] A. Anglemyer, "Digital contact tracing technologies in epiDemics: A rapid review," *Saudi Medical Journal*, vol. 41, no. 9, p. 1028, 9 2020. [Online]. Available: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013699/fullhttps://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD013699/abstract

[172] A. B. Dar, A. H. Lone, S. Zahoor, A. A. Khan, and R. Naaz, "Applicability of mobile contact tracing in fighting pandemic (COVID-19): Issues, challenges and solutions," *Computer Science Review*, vol. 38, p. 100307, 11 2020.

[173] "MilaCOVI - Mila." [Online]. Available: https://mila.quebec/en/project/covi/

[174] P.-S. Loh, "Flipping the Perspective in Contact Tracing," 10 2020. [Online]. Available: https://arxiv.org/abs/2010.03806v2

[175] "NZ COVID Tracer app — Ministry of Health NZ." [Online]. Available: https://www.health.govt.nz/covid-19-novel-coronavirus/covid-19-resources-and-tools/nz-covid-tracer-app

[176] "WO2022133171A1 - Beacon-based exposure notification systems and methods - Google Patents." [Online]. Available: https://patents.google.com/patent/WO2022133171A1/en?inventor=petrie&assignee=wehealth&oq=wehealth+petrie

[177] "Australian state to impose hefty fines to compel COVID-19 isolation — Reuters." [Online]. Available: https://www.reuters.com/article/us-health-coronavirus-australia-idUSKCN25005D

[178] "COVID-19 Data Explorer - Our World in Data." [Online]. Available: https://ourworldindata.org/explorers/coronavirus-data-explorer?time=earliest..2021-08-08&facet=none&Metric=Confirmed+cases&Interval=7-day+rolling+average&Relative+to+Population=true&Color+by+test+positivity=false&country=~AUS

[179] "Coronavirus (COVID-19) at a glance – 19 November 2020 — Australian Government Department of Health and Aged Care." [Online]. Available: https://www.health.gov.au/resources/publications/coronavirus-covid-19-at-a-glance-19-november-2020

[180] J. M. Buchanan, J. M. Buchanan, and G. Tullock, *The calculus of consent: Logical foundations of constitutional democracy*. University of Michigan press, 1965, vol. 100.

[181] P. C. Jentsch, M. Anand, and C. T. Bauch, "Prioritising COVID-19 vaccination in changing social and epidemiological landscapes: a mathematical modelling study," *The Lancet Infectious Diseases*, 2021.

[182] "USClaims — Unlicensed Drivers Responsible for 20% of Accidents." [Online]. Available: https://usclaims.com/educational-resources/non-licensed-drivers-responsible-for-20-percent-of-all-auto-accidents/

[183] A. Endo, S. Abbott, A. J. Kucharski, S. Funk *et al.*, "Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China," *Wellcome Open Research*, vol. 5, no. 67, p. 67, 2020.

[184] P. Zhu and X. Tan, "Is compulsory home quarantine less effective than centralized quarantine in controlling the COVID-19 outbreak? Evidence from Hong Kong," *Sustainable Cities and Society*, vol. 74, p. 103222, 11 2021.

[185] "China's Shanghai announces two new rounds of mass COVID testing — Reuters." [Online]. Available: https://www.reuters.com/world/china/chinas-shanghai-carry-out-some-covid-testing-residents-say-2022-07-05/

[186] "Coronavirus Spread At University Of Illinois Leads To Student Lockdown : Coronavirus Updates : NPR." [Online]. Available: https://www.npr.org/sections/coronavirus-live-updates/2020/09/03/909137658/university-with-model-testing-regime-doubles-down-on-discipline-amid-case-spike

[187] E. Mahase, "Covid-19: Mass testing in Slovakia may have helped cut infections," *BMJ*, vol. 371, p. m4761, 12 2020. [Online]. Available: https://www.bmj.com/content/371/bmj.m4761https://www.bmj.com/content/371/bmj.m4761.abstract

[188] A. Elbanna and N. Goldenfeld, "Frequency of surveillance testing necessary to reduce transmission of SARS-CoV-2," 2021.

[189] "Swish, Gargle, Repeat: UArizona Researcher Explores Mouth Rinse Test as Alternative to COVID-19 Nasal Swab — University of Arizona News." [Online]. Available: https://news.arizona.edu/story/swish-gargle-repeat-uarizona-researcher-explores-mouth-rinse-test-alternative-covid-19-nasal

[190] D. M. Goldfarb, P. Tilley, G. N. Al-Rawahi, J. A. Srigley, G. Ford, H. Pedersen, A. Pabbi, S. Hannam-Clark, M. Charles, M. Dittrick, V. J. Gadkar, J. M. Pernica, and L. M. Hoanga, "Self-collected saline gargle samples as an alternative to health care worker-collected nasopharyngeal swabs for COVID-19 diagnosis in outpatients," *Journal of Clinical Microbiology*, vol. 59, no. 4, 4 2021. [Online]. Available: https://journals.asm.org/doi/10.1128/JCM.02427-20

[191] "Wargame - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Wargame

[192] R. M. Anderson and R. M. May, *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

APPENDICES

Appendix A

Supplement to *The economic value of quarantine is higher at lower case prevalence, with quarantine justified at lower risk of infection*

A.1 Threshold Simplification

Substituting Equation 3.4 into Equation 3.6, we obtain:

$$r_{thresh} = \frac{f(\max\{0, 1 - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i}\}) - f(1)}{(P - Q_i - Q_n - 1) * [f(\max\{0, 1 - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i-1}\}) - f(\max\{0, 1 - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i}\})]} \tag{A.1}$$

Setting $f(x) = x$ and assuming $1 - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i-1} > 0$ yields:

$$r_{thresh} = \frac{1 - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i} - 1}{1 - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i-1} - 1 + \frac{R_{target}}{R_t} * \frac{I}{I-Q_i}} * \frac{1}{P - Q_i - Q_n - 1} \tag{A.2}$$

Simplifying the numerator and denominator:

$$r_{thresh} = \frac{-\frac{R_{target}}{R_t} * \frac{I}{I-Q_i}}{\frac{R_{target}}{R_t} * \frac{I}{I-Q_i} - \frac{R_{target}}{R_t} * \frac{I}{I-Q_i-1}} * \frac{1}{P - Q_i - Q_n - 1} \tag{A.3}$$

Factoring out $\frac{R_{target}}{R_t}$:

$$r_{thresh} = \frac{\frac{R_{target}}{R_t}(-\frac{I}{I-Q_i})}{\frac{R_{target}}{R_t}(\frac{I}{I-Q_i} - \frac{I}{I-Q_i-1})} * \frac{1}{P-Q_i-Q_n-1} \tag{A.4}$$

Multiplying the numerator and denominator by $(I-Q_i-1)(I-Q_i)$:

$$r_{thresh} = \frac{(I-Q_i-1)(I-Q_i)(-\frac{I}{I-Q_i})}{(I-Q_i-1)(I-Q_i)(\frac{I}{I-Q_i} - \frac{I}{I-Q_i-1})} * \frac{1}{P-Q_i-Q_n-1} \tag{A.5}$$

Multiplying through and simplifying:

$$r_{thresh} = \frac{-(I-Q_i-1)}{(I-Q_i-1)-(I-Q_i)} * \frac{1}{P-Q_i-Q_n-1} \tag{A.6}$$

$$r_{thresh} = \frac{-(I-Q_i-1)}{-1} * \frac{1}{P-Q_i-Q_n-1} \tag{A.7}$$

$$r_{thresh} = \frac{I-Q_i-1}{P-Q_i-Q_n-1} \tag{A.8}$$

A.2 Fluctuating R'_t

In Model Section 2.1 from the main text, we assume that D instantaneously adapts to achieve R_{target} given prevailing conditions and current quarantine policies. This is motivated by the observation of relatively stable R'_t averaged over longer timescales. In this section, we show that Equation 2.1 (from the main text), and hence subsequent findings, are still a good approximation even when including stochastic effects, time delays, and uncertainty about system state.

To make this argument, we assume that for a certain region and time frame, the virus is neither eliminated nor creates herd immunity, and that control measures are relatively stable. Over N serial intervals between times t_0 and t_N, the initial and final case prevalence are related by $I(t_N) = I(t_0)\prod_{n=0}^{N-1} R'_t(t_n)$. Equivalently, the geometric mean of R'_t over this time period can be expressed using the initial and final case prevalence, $(\prod_n (R'_t(t_n)))^{1/N} = (\frac{I(t_N)}{I(t_0)})^{1/N}$. For most regions, case prevalence doesn't vary by more than a couple orders of

magnitude on the scale of months, which yields a (geometric) mean R'_t value quite close to 1. As an example, with $I(t_N)/I(t_1) = 100$, and $N = 36$ (6 months with a serial interval of 5 days), $100^{1/36} = 1.14$. Approximating the geometric mean of the effective reproductive number as 1, and expanding to include distancing and quarantine terms:

$$1 \approx \Big[\prod_{n=0}^{N-1}(R'_t(t_n))\Big]^{1/N} = \Big[\prod_{n=0}^{N-1}(R_t(t_n)(1-D(t_n))\frac{I(t_n)-Q_i(t_n)}{I(t_n)})\Big]^{1/N} \tag{A.9}$$

Taking the log of both sides and then multiplying by $1/N$:

$$\frac{1}{N}\log(1) \approx \frac{1}{N}\sum_{n=0}^{N-1}[\log(R_t(t_n)) + \log(1-D(t_n)) + \log(\frac{I(t_n)-Q_i(t_n)}{I(t_n)})] \tag{A.10}$$

$\frac{1}{N}\sum_{n=0}^{N-1}$ is the definition of the average over time:

$$0 \approx E[\log(R_t(t))] + E[\log(1-D(t))] + E[\log(\frac{I(t)-Q_i(t)}{I(t)})] \tag{A.11}$$

Assuming control measures don't change drastically over time, we replace the mean of the log with the log of the mean:

$$0 \approx \log(E[R_t(t)]) + \log(E[1-D(t)]) + \log(E[\frac{I(t)-Q_i(t)}{I(t)}]) \tag{A.12}$$

Taking the exponential of each side:

$$1 \approx E[R_t(t)] * E[1-D(t)] * E[\frac{I(t)-Q_i(t)}{I(t)}] \tag{A.13}$$

This is equivalent to Equation 2.1 (from the main text) with R'_t set to 1, showing that by averaging over time, the requirements for precise knowledge and control of system state can be relaxed.

A.3 Imported Cases and Stochasticity

For regions where imported cases make up a significant fraction of total cases, policy-makers should consider the modified dynamics when choosing R_{target}. Imported cases act as a forcing term, with the number of cases at a future timepoint being the sum of locally obtained infections and imported infections. The mechanics of choosing a risk threshold do not change, but it is important to highlight that R_{target} only determines a portion of future cases. A complete solution must also consider the effect of border policies.

When the total number of cases is small, e.g. in a region with excellent local control for which rare imported cases are the primary concern, discrete and stochastic effects become important.

Quantitative models of tiny case count scenarios are challenging, because of stochasticity in the length of time needed to return to elimination, combined with curvature in the harms of shutdown as a function of its duration. This is further exacerbated by superspreader dynamics [183]. Modelling the number of secondary cases with a negative binomial distribution with an overdispersion parameter 0.1 and mean 2.5 yields a variance of 65. In locations not aiming for local elimination, stochasticity averages out over time, and expected values can be used. When local elimination is being attempted, it would be useful to make use of discrete, stochastic models when choosing R_{target}, e.g. to calculate the probability distribution of time to elimination and the expected damage from shutdowns of different severity. With the choice of R_{target} set, the previous arguments mostly still hold (although some consideration may be given to the difference in variance in transmissions based on broad vs. targeted distancing).

If local elimination has already been achieved, then a still more complicated case is to consider the risk of an outbreak from an imported case, the time delay and growth in the outbreak before it is noticed, and the subsequent shutdowns needed to re-achieve local elimination and thus relax general social distancing. It is likely that the considerable costs of such an outbreak warrant strict quarantine for incoming individuals. A robust testing and contact tracing system capable of quickly containing any new outbreak is also key.

A.4 Cost Function Dependence

Here we compare the risk threshold r_1 derived from the linear cost function $f_1(x) = x$ to the risk threshold r_2 derived from a non-linear cost function $f_2(x)$. As discussed in Section 3.3.1, disparate considerations all support the assumption that a realistic choice for $f_2(x)$

will be strictly increasing and concave up. We also assume that $f_2(x)$ is differentiable. For both of these derivations, we define the distancing setpoint as $D_0 = max(0, 1 - \frac{R_{target}}{R_t}\frac{I}{I-Q_i})$, the required distancing if one more infectious person were isolated as $D_- = max(0, 1 - \frac{R_{target}}{R_t}\frac{I}{I-Q_i-1})$, and the total quarantined or isolated population as $Q = Q_i + Q_n$. By expanding Equation 2.6 from the main text, we obtain the two risk thresholds (with r_1 corresponding to f_1 and r_2 corresponding to f_2):

$$r_1 = \frac{f_1(1) - f_1(D_0)}{f_1(D_0) - f_1(D_-)} \frac{1}{P - Q - 1} \tag{A.14}$$

$$r_2 = \frac{f_2(1) - f_2(D_0)}{f_2(D_0) - f_2(D_-)} \frac{1}{P - Q - 1} \tag{A.15}$$

Substituting $f_1(x) = x$, and taking the ratio of the two thresholds gives:

$$\frac{r_2}{r_1} = \frac{\frac{f_2(1)-f_2(D_0)}{1-D_0}}{\frac{f_2(D_0)-f_2(D_-)}{D_0-D_-}}. \tag{A.16}$$

We use the fact that $\frac{f_2(1)-f_2(D_0)}{1-D_0} \leq \max_{x\in[0,1]} f_2'(x)$ and $\frac{f_2(D_0)-f_2(D_-)}{D_0-D_-} \geq \min_{x\in[0,1]} f_2'(x)$ to obtain a bound on the ratio of the two thresholds:

$$\frac{r_2}{r_1} \leq \frac{\max_{x\in[0,1]} f_2'(x)}{\min_{x\in[0,1]} f_2'(x)}. \tag{A.17}$$

Because $f_2(x)$ is concave up (indicating that additional isolation is more costly), we have $\frac{f_2(1)-f_2(D_0)}{1-D_0} \geq \frac{f_2(D_0)-f_2(D_-)}{D_0-D_-}$, and therefore $r_2 \geq r_1$.

In order to approximate the risk threshold, we substitute Equation 2.7 from the main text into Equation A.16, and obtain:

$$r_2 = \frac{\frac{f_2(1)-f_2(D_0)}{1-D_0}}{\frac{f_2(D_0)-f_2(D_-)}{D_0-D_-}} \frac{I - Q_i - 1}{P - Q_i - Q_n - 1}. \tag{A.18}$$

For differentiable $f_2(x)$ and relatively large $I - Q_i$, $D_0 - D_-$ is small, and $\frac{f_2(D_0)-f_2(D_-)}{D_0-D_-}$ can be approximated as $f_2'(D_0)$. Let $r_{pop} = \frac{I-Q_i}{P-Q_i-Q_n}$ be the risk of the unquarantined population. Using the approximation, $r_{pop} \approx \frac{I-Q_i-1}{P-Q_i-Q_n-1}$ and referring to r_2 as r_{thresh}, we obtain Equation 3.8.

Appendix B

Supplement to *Realistic consideration of quarantine compliance suggests earlier PCR testing and shorter quarantine following SARS-CoV-2 exposure*

B.1 Dependence on Quarantine and Isolation Effectiveness

Here we show that the optimal choice of quarantine length and test date(s) depend only on the ratio C_Q of quarantine compliance to isolation compliance, and not the individual values of these two compliances.

A policy π produces a vector $\boldsymbol{\phi}(\pi)$, whose four elements sum to the expected number of transmissions that would occur given no compliance with either quarantine or isolation. Those expected transmissions are divided among the four elements according to whether they occur: [during isolation, during quarantine after a negative test result has been received, during quarantine at other times, while not recommended for quarantine or isolation]. Because policies differ in the timing of testing and quarantine, they differ in the distribution of total expected transmissions (in the absence of compliance with either isolation or quarantine, but with testing on the recommended date) among these four

elements.

The reduction in transmissions due to scaling from quarantine and isolation is given by:

$$J(\pi, C_I, C_Q, C_N) = \boldsymbol{\phi}(\pi) \cdot \begin{bmatrix} C_I \\ (C_Q * C_I) * (1 - C_N) \\ C_Q * C_I \\ 0 \end{bmatrix}$$

Lemma: for any two policies, π_a and π_b and $\alpha > 0$, if $J(\pi_a, C_I, C_Q, C_N) \geq J(\pi_b, C_I, C_Q, C_N)$, then $J(\pi_a, \alpha C_I, C_Q, C_N) \geq J(\pi_b, \alpha C_I, C_Q, C_N)$.

Expanding

$$J(\pi_a, C_I, C_Q, C_N) \geq J(\pi_b, C_I, C_Q, C_N),$$

we have

$$\boldsymbol{\phi}(\pi_a) \cdot \begin{bmatrix} C_I \\ (C_Q * C_I) * (1 - C_N) \\ C_Q * C_I \\ 0 \end{bmatrix} \geq \boldsymbol{\phi}(\pi_b) \cdot \begin{bmatrix} C_I \\ (C_Q * C_I) * (1 - C_N) \\ C_Q * C_I \\ 0 \end{bmatrix}.$$

Multiplying both sides by α yields

$$\boldsymbol{\phi}(\pi_a) \cdot \begin{bmatrix} \alpha C_I \\ (C_Q \alpha C_I) * (1 - C_N) \\ C_Q \alpha C_I \\ 0 \end{bmatrix} \geq \boldsymbol{\phi}(\pi_b) \cdot \begin{bmatrix} \alpha C_I \\ (C_Q \alpha C_I) * (1 - C_N) \\ C_Q \alpha C_I \\ 0 \end{bmatrix},$$

and thus

$$J(\pi_a, \alpha C_I, C_Q, C_N) \geq J(\pi_b, \alpha C_I, C_Q, C_N).$$

Therefore if C_I is required to be a positive number, the ordering of policies in terms of expected transmissions is independent of C_I. Since the order of policies ranked by expected transmissions and quarantine days is constant for a fixed ratio of isolation and quarantine compliances, the recommended policy will also be constant.

Appendix C

An Idealized Outbreak Response

Early in an outbreak of a novel pathogen little may be known about transmission modes, timing of infectiousness, timing and type of symptoms, and timing of test sensitivity (and test capacity may be quite limited). It would be sensible to use fairly uninformative priors on these distributions based on known pathogens (with more weighting for similar pathogens if known). If the decision is made to suppress transmission, then an approach like the one outlined in Section 5.3 can be used (by attempting to choose the optimal trade-off between broad distancing, targeted quarantine, and border restrictions while lowering disease prevalence along an 'optimal' trajectory). Total case numbers will likely be quite low at this point- leading to a low quarantine risk threshold (Chapter 3). With relatively uninformative risk calculations from contact information and a low risk threshold, an optimal strategy would likely recommend quarantine for a large fraction of the detected second order contact network- even quite brief encounters (as Vietnam implemented in early 2020 [159]). This contact network could be detected almost instantaneously using the system described in Section 5.6 (with decentralized proximity notification, recursive notification, and exposure sharing).

Quarantine is highly disruptive for individuals, but extremely valuable (in expectation) for the broader population, so people asked to quarantine could be compensated at a level roughly aimed to make it worthwhile for them, but not so high that it incentivizes intentional exposure to the virus. Unenforced home quarantine has had limited success, so either dedicated facilities or more strictly enforced home quarantine would be preferable [184]. Transportation to facilities could be done using dedicated vehicles with proper fomite and airborne precautions (completely separate compartments, not just masks [157]). People in quarantine and isolation could frequently have swabs collected and stored even

if testing is not available – this data would be valuable for estimating test sensitivity and infectiousness distributions. Any observed symptoms would also be recorded.

As more information became available the distributions used for risk assessments could be refined using Bayesian updates. With greater predictive power the risk distribution will concentrate – leading to fewer people who are judged high enough risk to warrant quarantine. Unexpected infections could be thoroughly investigated to prevent repeated mistakes, e.g. New Zealand's analysis of transmission in quarantine facility [145]. Much of this investigation cannot be automated because there are many possible unmodelled transmission routes. However, data collected on these unexpected cases can aid in the investigation, including recent locations and the local contact network (using the system from Section 5.6), along with genomic sequencing (if useful). If the mode of unexpected transmissions can be determined, a comparison could be made between the cost of population-wide actions to reduce it[1] vs. collecting more information to enable accurate future predictions of risk and expanded quarantine.

If a significant fraction of new cases do not present because they are asymptomatic or paucisymptomatic, suppression becomes more challenging, but is still possible. The notification system (described in Section 5.6) has a fair amount of robustness to this due to backward and second (or greater) degree contact tracing, meaning that we can tolerate a missing link in the transmission chain. However, it would be best not to rely on noticeable symptoms at all to achieve disease control. The above approach can be augmented with broad surveillance testing. A less disruptive version could use frequent wastewater testing at a spatial resolution where one case could be reliably detected- then prompting local testing.

For a broad range of infectious diseases, rapid detection of 2nd degree contact networks and aggressive quarantine policy would be sufficient for low-cost local elimination with no broad social distancing needed. As risk assessment improves fewer contacts would be estimated as high risk and testing could replace quarantine, minimizing the expected cost to suppress a novel introduction of the disease.

[1] In some cases this could be quite difficult- for example transmission via air ducts in large residential buildings

Appendix D

A Potential Design for Mass Testing

Several countries, workplaces and schools implemented some form of mass testing during the COVID-19 pandemic [185, 186, 187]. The goal of this section is to design a system for mass testing that would be feasible in a broader range of countries and would be able to robustly contain a novel infectious disease. We will consider modifications to existing approaches and technological advancements that could conceivably enable this within a few years. Specifically, I claim that innovation in automation and logistics could massively scale testing availability, prioritization of pleasantness and accessibility could increase public support, and that with careful design there is an acceptable balance between robustness and civil liberties.

Mass testing deployments in western countries have been limited in time or the fraction of the population participating (e.g. Slovakia testing the whole population twice, but not on an ongoing basis [187]). The expense of massively scaling testing was likely a contributing factor preventing more widespread deployment. The limited scope of deployment limited the possible success because people testing regularly could still be infected by members of the broader community, and cases could rebound after time-limited interventions. With a targeted deployment of testing, even a limited quantity of tests could be sufficient for suppression. However, this is more complicated to deploy properly, and the public may not be comfortable sharing information needed for highly targeted testing[1].

Another option is to innovate in automation of PCR testing[2] and logistics to drastically reduce the cost. Much more ambitious deployments would be possible if the cost of a test were around $5 instead of $100 (testing every person daily at $100 per test is close to the

[1] Also it could seem unfair for only specific subcommunities to need to get tested

[2] Or other types of testing that could be used in the early stages of a pandemic

GDP of wealthy countries). PCR tests are highly sensitive, so it seems unlikely for someone to test negative while actively infectious with an arbitrary airborne pathogen. If this is true, then for some testing frequency and test delay (and implicitly isolation effectiveness), a mass testing policy would be successful. For the original COVID-19 variant, Elbanna and Goldenfeld estimated that testing twice a week would be sufficient to bring R_t below 1 (assuming perfect isolation of those who test positive) [188].

Mass testing seems to be an important component of China's "dynamic Covid-zero" policy (as of July 2022). The independent effect of mass testing is difficult to estimate because they are using it in combination with a few other strategies like venue based tracing and localized lockdowns. The combined approach seems to be effective at containing outbreaks, but it is difficult to determine what the cost of this is, as the main source of information is Chinese media – which is politically incentivized to tell a positive story. Western media sources have predominately portrayed the approach as excessive, but this framing could also be politically motivated. In either case, now that vaccines and treatments are available, public support for any kind of Covid-zero strategy is likely very limited outside of a few countries [3]. However, support for a mass-testing approach may be greater when faced with a novel deadly pandemic. There are a few modifications that could improve the likelihood of achieving this support: reduced cost (as previously discussed), less authoritarian-seeming rules (to be discussed shortly) and increasing the pleasantness of the testing experience. To improve the testing experience, gargle [189, 190] or saliva tests could be used instead of nasal swabs [4] and time spent waiting at or traveling to testing sites could be reduced.

A more difficult part of the design is deciding on tradeoffs between robustness to non-compliance and authoritarian capabilities (or the perception of them). The most straightforward approach is to reduce the number of points of failure (with 2 steps instead of 6, the success rate doesn't need to be as high). Ideally, any remaining critical steps will have an intuitive justification; if necessary, it is much easier to enforce participation for things that the public believes are necessary and important. In a mass testing deployment, there are two critical steps that depend on the population: 1) frequent testing and 2) effective isolation if positive. The honor system may be too optimistic, but both steps have not-too-excessive methods of enforcement. Frequent testing could be encouraged by requiring proof of a recent negative test to enter public spaces. Isolation at home can be subject to random check-ins with fines to discourage non-compliance.

[3] And even if there was support, border policy with countries not aiming for elimination would be a significant challenge

[4] It seems like nasal and gargle tests have comparable sensitivity for COVID-19 [190]. Even if more pleasant testing was 10-20% less sensitive, this could be still be worthwhile if it improves public support.

Another area to be careful about is the validity of test results – whether tests are being taken properly and by the right people. For example, self-serving people who suspect they could be infected have an incentive to not test properly. A robust system will preferably have the ability to detect this failure mode and adjust if it became a significant problem. For this reason, designs where the sample collection process can be observed are preferable. Having testing records securely stored and associated with people's identity makes understanding the disease easier for public health departments and researchers. However, in some countries requiring this may decrease public support and disincentivize participation among some populations (undocumented immigrants, libertarians, etc.). A potential solution is to offer anonymous testing but heavily incentivize[5] a version that associates test results with a person's health record.

Based on these considerations, the following is a potential implementation:

1. Sample collection sites easily available – unstaffed, in style of self-checkout kiosks[6]
2. Gargle or saliva-based sampling, with detachable QR code to get test results later
3. Scalable service like Uber or Amazon delivery frequently picks up samples[7]
4. Highly automated[8] labs perform PCR tests
5. Results posted to public database linked to QR codes (ideally within 8 hours of collection)
6. QR code demonstrating a recent negative test permits entry into public spaces
7. People who test positive are required by law to effectively isolate – free transportation, financial support, food, and accommodation provided if needed

The likelihood of success for a mass-testing intervention (as with most NPIs) increases with the ability to detect, diagnose, and adapt to failure modes. In this case, some important areas to monitor are isolation effectiveness, testing compliance, and the frequency of transmissions occurring before testing positive. This proposal is based on several partially

[5]For example, allowing entry into public spaces immediately after testing instead of waiting for test results

[6]Potentially with a cheap camera to discourage sabotage or intentionally testing incorrectly

[7]With option of testing at home for those with accessibility issues, symptoms, or who wish to pay for it

[8]Potentially more innovation needed to make this scalable and cheap

validated assumptions, which ideally can be strengthened with further analysis, numerical modeling, wargaming [191], and real-world experiments. While preliminary, it offers some optimism that innovation can enable an acceptable compromise between civil liberties and robust disease control.

Glossary

Non-Pharmaceutical Intervention (NPI): Health interventions that do not depend on medication (e.g. contact tracing, social distancing).

$\boldsymbol{R_0}$: The "basic reproduction number" [192], which gives the expected number of cases infected by one initial case in a susceptible population.

$\boldsymbol{R_t}$: The "effective reproduction number" [192], which gives the expected number of cases infected by one case at time t. This value depends on immunity in the population and health interventions.

SIR model: a compartmental model of infectious disease spread with three compartments: susceptible, infectious, and recovered [192].

Heatmap: A plot using colour to show how a quantity varies over space. E.g. concentration of infectious individuals varying by location.

ROC curve: A "Receiver operating characteristic" curve shows how the performance of a classifier varies with a changing decision threshold [138]. It plots the true positive rate (correctly predicted true positives divided by total positives) against the false positive rate (incorrectly predicted false positives divided by total negatives).

Pareto frontier: The set of feasible points that are optimal for at least one choice of objective function in a multi-objective optimization problem. The Pareto frontier can be expanded by relaxing limiting constraints.

Index case: In contact tracing, the index case is the first person within a cluster who is identified as being infected with a communicable disease.

Secondary attack rate: The probability of infection among contacts exposed to an infected individual.