Listening to Bluetooth Beacons for Epidemic Risk Mitigation

Gilles Barthe[1] Roberta De Viti[2] Peter Druschel[2] Deepak Garg[2]
Manuel Gomez-Rodriguez[2] Pierfrancesco Ingo[2] Heiner Kremer[3]
Matthew Lentz[4] Lars Lorch[5] Aastha Mehta[2] Bernhard Schölkopf[3]

Abstract

During the ongoing COVID-19 pandemic, there have been burgeoning efforts to develop and deploy smartphone apps to expedite contact tracing and risk notification. Unfortunately, the success of these apps has been limited, partly owing to poor interoperability with manual contact tracing, low adoption rates, and a societally sensitive trade-off between utility and privacy. In this work, we introduce a new privacy-preserving and inclusive system for epidemic risk assessment and notification that aims to address the above limitations. Rather than capturing pairwise encounters between smartphones as done by existing apps, our system captures encounters between inexpensive, zero-maintenance, small devices carried by users, and beacons placed in strategic locations where infection clusters are most likely to originate. Epidemiological simulations using an agent-based model demonstrate several beneficial properties of our system. By achieving bidirectional interoperability with manual contact tracing, our system can help control disease spread already at low adoption. By utilizing the location and environmental information provided by the beacons, our system can provide significantly higher sensitivity and specificity than existing app-based systems. In addition, our simulations also suggest that it is sufficient to deploy beacons in a small fraction of strategic locations for our system to achieve high utility.

Introduction

Containing infectious diseases such as the ongoing COVID-19 pandemic requires effective testing, contact tracing, and isolation of infected individuals (TTI) [1, 2]. Among these, contact tracing is an important tool that can help direct limited test resources to those most likely to be infected, by identifying infected individuals and those who came into close contact with them during their infectious period. Furthermore, contact tracing can provide insight into the circumstances of contagion, which in turn informs the implementation of public health policies and interventions to help contain disease spread. For instance, outbreaks of COVID-19 in meat packing plants in Germany [3] provided insights into conditions that can potentially breed infection hotspots, leading to regulatory action [4].

To expedite contact tracing, a number of digital contact tracing systems have been proposed [5–11]. In most of these systems, individuals install a smartphone application that records instances of physical proximity with devices of other individuals via close-range Bluetooth exchanges, referred to as *encounters*. Once an individual is diagnosed with the disease, they share their recent encounter history with a health authority, used to notify other at-risk individuals who might have come into contact with the diagnosed individual during the individual's period of contagiousness.

Despite strong efforts to deploy these smartphone-based, pairwise encounter-based contact tracing systems, below denoted SPECTs, there is only scarce evidence of their effectiveness [12, 13]. This may be due to poor interoperability with manual contact tracing [12], low adoption rates ($< 25\%$) in countries that do not mandate their use [14], technical limitations imposed by smartphones (e.g., battery life, application

[1]Max Planck Institute for Security and Privacy, Bochum, Germany. [2]Max Planck Institute for Software Systems, Saarbrücken and Kaiserslautern, Germany. [3]Max Planck Institute for Intelligent Systems, Tübingen, Germany. [4]VMware Research, Palo Alto, USA. [5]ETH Zürich, Zürich, Switzerland.

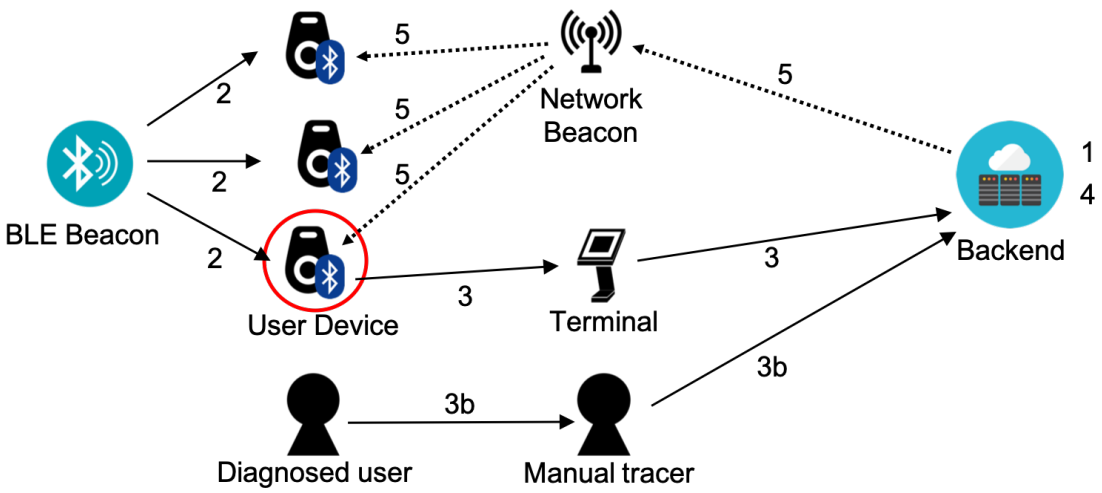


Figure 1: PanCast's architecture. **1.** Beacons and dongles are registered with the backend. **2.** Dongles record encounters with BLE beacons. **3.** Diagnosed users or healthy volunteers may upload their history of encountered beacons to the backend via a terminal. **3b.** Optionally, health workers can manually feed inputs from users into the backend system. **4.** The backend updates the risk database with uploaded encounters. **5.** Risk information is periodically broadcast from the backend to network beacons, which broadcast the information to nearby dongles.

programming interfaces) [15], and a societally sensitive trade-off between utility and privacy. The latter generally precludes the use of location and associated environmental features (e.g., indoor vs. outdoor). This is a major disadvantage, since such measurements could help us predict infection risk, identify spatiotemporal infection clusters, and understand the means of disease transmission via quantitative estimation of transmission parameters, which would lead to better models of epidemic dynamics and the fine-grained effect of contact restrictions.

In the present work, we describe PanCast, a privacy-preserving, secure and inclusive system for epidemic risk assessment and notification that addresses the above limitations. In PanCast, Bluetooth beacons placed in strategic locations continuously broadcast ephemeral IDs. A subset of these beacons, called network beacons, also broadcast risk information associated with times when individuals who tested positive were near specific beacons. Individuals carry inexpensive, zero-maintenance, small electronic dongles in the form of cards or key fobs that listen to these beacon broadcasts passively (i.e., without transmitting anything), store the beacons' ephemeral IDs in local memory, and later compare the stored IDs against risk information broadcast from nearby network beacons. When an individual tests positive for the infectious disease, they disclose (a selected subset of) the list of ephemeral IDs stored in their dongle. The information about which locations were dangerous at which times (which depends on users' visits as well as on location features) then gets included in the broadcast risk messages. A user learns their risk of contagion due to proximity to sites visited by diagnosed individuals but does not learn anything about locations visited by other individuals. Figure 1 shows an overview of PanCast's architecture. Refer to *Materials and Methods* and the PanCast's white paper [16] for more details on PanCast's design, data collection and processing, risk dissemination, risk score calculation, and privacy and security properties.

By design, PanCast enables the participation of technology-challenged, economically disadvantaged, or physically challenged individuals who cannot or do not wish to use smartphones. Moreover, it ensures data minimization in accordance with existing regulations for (manual) contact tracing—a healthy individual can use the system in a purely passive radio mode, individuals who test positive control which information they disclose (much like in a manual contact tracing interview), and this disclosed information is accessible only to individuals at risk and in a privacy-preserving way. Due to the strong level of privacy afforded by the

passive mode, and in contrast to existing digital contact tracing systems, PanCast makes use of location and environmental information provided by the beacons to enhance risk score estimation and identify infection clusters [17–20]. Moreover, PanCast makes use of this information to achieve bidirectional interoperability with manual contact tracing. If an individual who tested positive owns a dongle, they can use the information saved in their dongles to better recall visited locations during a contact tracing interview. If an individual participating in manual contact tracing does not own a dongle, a contact tracer could manually create an entry in the risk database for any of the individual's reported visits to places with installed beacons. By virtue of these properties, PanCast partially mitigates the so-called x^2-adoption problem of SPECTs [10]: If a proportion $x < 1$ of the population has adopted a SPECT, encounters get noticed with probability $x^2 \ll 1$.

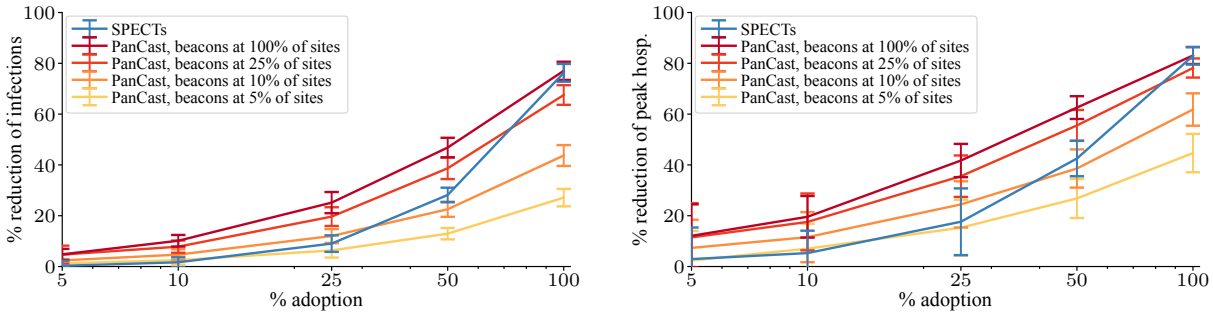
To evaluate the performance of our system compared to existing digital contact tracing systems, we perform epidemiological simulations using an agent-based model [21]. Our results suggest that, by achieving bidirectional interoperability with manual contact tracing, the system helps control viral spread even with low adoption. Moreover, by utilizing location and environmental information, our system provides significantly higher sensitivity and specificity than existing apps. When the variability in environmental transmission rates is high across sites, we find that PanCast is able to achieve up to 33% higher sensitivity (and 32% higher precision) than SPECTs at the same specificity value. Finally, our results also show that deploying beacons in a small fraction of strategic locations already achieves high utility.

Results

We simulate the deployment of our system exemplarily in Tübingen, a town in southwest Germany. In the simulations, we place beacons across five types of sites and identify their real locations using OpenStreetMap [22]: (i) schools, universities and research institutes; (ii) restaurants, cafes and bars; (iii) bus stops; (iv) offices and shops; and (v) supermarkets and convenience stores. To model the exposure of individuals at these sites, we use a recently introduced agent-based epidemiological model where sites are explicitly represented [21]. This model utilizes the spatial distribution of site locations, high-resolution population density data, country-specific information about household structure and region-specific age demographics to simulate the frequency and duration of individual visits to each site. Moreover, the model quantifies the influence that these individual mobility patterns, environmental drivers, as well as testing, contact tracing and isolation of infected individuals (TTI) have on the rate of transmission by each infected individual.

To compare the efficacy of PanCast with generic smartphone-based, pairwise encounter-based contact tracing systems (SPECTs) as described in the introduction, we simulate a range of epidemiological scenarios in which either PanCast or a SPECTs is employed over a six month period. As in the real world, this is always combined with manual contact tracing. All results stated in the following are averages of 100 random roll-outs of the simulation. No other interventional measures (such as e.g., site closures) are implemented during the simulation. Refer to *Material and Methods* for details on the epidemiological simulations and the implementation of the various tracing measures.

PanCast outperforms SPECTs under partial adoption Figure 2 shows the reduction of the number of infections and peak hospitalizations achieved by PanCast and SPECTs with respect to the baseline scenario in which exclusively manual tracing is employed. By circumventing the x^2-adoption problem via interoperation with manual tracing, we observe that PanCast generally scales more favourably with decreasing adoption levels. Figure 2b shows that in the ideal scenario, in which beacons are placed at all sites, an adoption level of roughly 30% with PanCast was sufficient to halve the peak hospitalizations compared to the baseline, whereas the same result would require around 60% adoption with SPECTs. As the ideal scenario might be unrealistic, we investigate how the performance of PanCast decreases when beacons are placed strategically. To this end, we study the placement at a proportion of the sites with highest integrated visit time, defined as the sum of individual visit durations at a given site over a period of one month. For each proportion of sites with beacons, there is an adoption level up to which PanCast turns out to be more effective than SPECTs. With beacons at 25% of the sites PanCast shows a clear advantage over SPECTs



(a) Reduction of infections

(b) Reduction of peak hospitalizations

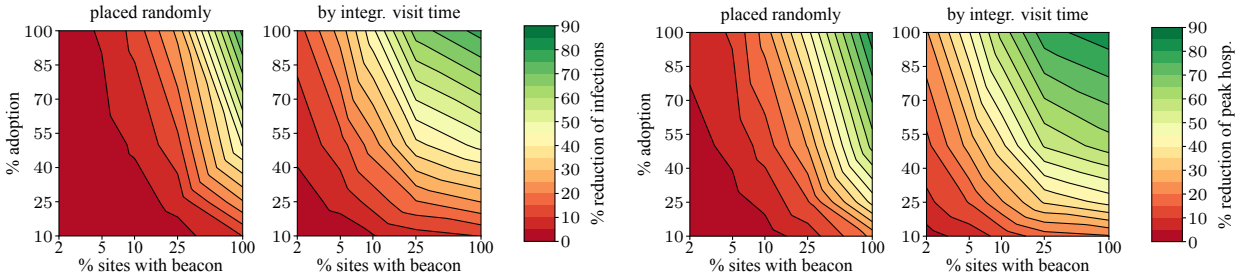
Figure 2: Reduction of overall infections and peak hospitalizations with respect to the baseline scenario in which no digital tracing is employed. The horizontal axis represents the adoption level of the digital tracing system on a log scale. The lines correspond to the mean of 100 roll-outs of the simulation and the error bars represent one standard deviation.

up to adoption levels of 80%. We observe that in order to surpass the efficacy of the SPECT currently employed in Germany with an adoption level of 29% [23], deploying PanCast with beacons placed strategically at >10% of the sites might be sufficient. Finally, equipping only >5% of the sites with beacons may be sufficient to match the performance of SPECTs for adoption levels of up to 15%, as currently observed in most countries [24].

Strategic placement of beacons allows for high utility at low cost In the previous experiment, we have placed beacons strategically at the sites where exposures are more likely to happen. Figure 3 compares this to a random placement strategy. The results show that a random placement requires equipping a large proportion of roughly 50% of the sites with beacons in order to achieve a significant reduction of infections and peak hospitalizations for low and intermediate adoption levels. This disadvantage can be circumvented by placing beacons at the sites with the highest integrated visit times. In this case, we observe that even for proportions of sites with beacons below 25%, PanCast is capable of significantly reducing infections and peak hospitalizations for low and intermediate adoption levels. Finally, note that these estimates are conservative as we rank the sites only according to integrated visit time and do not account for e.g., differences in site specific transmission rates. Any features known to affect infection risks (e.g., indoor vs. outdoor) could be used to further improve beacon placement.

Leveraging environmental information improves sensitivity and specificity As PanCast is aware of the sites at which encounters happen, it can leverage this information to better predict infection risks and thus increase the sensitivity (true positive rate) of individuals flagged as potentially infected. We simulate scenarios in which the transmission rate during an encounter between an infectious and a susceptible individual depends on the specific site type. Since environmental factors have shown to exhibit vastly different transmission rates, such as indoor vs. outdoor, room size, ventilation and air quality [25], by way of example, we increase the transmission rate at social sites by a factor k and decrease the transmission rate at bus stops by the same factor k, for $k \in \{2, 10\}$. While PanCast, using site information, has access to the true transmission rates, we assume for SPECTs to only have access to the average transmission rate. In both systems, we isolate and test individuals whose infection risk is above a certain threshold. By varying the threshold between 0 and 1 we obtain the receiver operating curves shown in Figure 4.

Across the range of specifity values, or true negative rate, PanCast generally shows a relative advantage over SPECTs. For low variability in transmission rates, i.e., when $k = 2$, we find that PanCast affords a marginal increase in sensitivity of up to 4% compared to SPECTs at approximately the same specificity.



(a) Reduction of infections (b) Reduction of peak hospitalizations

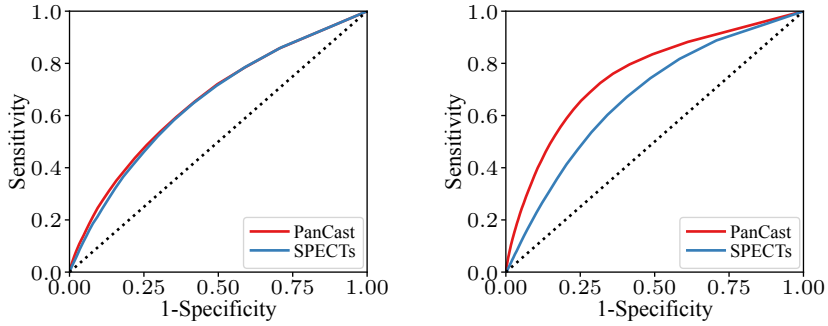
Figure 3: Effects of strategic beacon placement on the reduction of overall infections and peak hospitalizations for PanCast averaged over 100 simulation roll-outs. Performance given is relative to the baseline scenario without any digital tracing. In both (a) and (b), the left plot indicates the expected reduction when beacons are placed randomly at sites, and the right panel when beacons are placed at sites with largest integrated visit times (i.e., overall time spent at the sites).

However, when the variability in transmission rate is higher, i.e., when $k = 10$, PanCast becomes significantly more effective than SPECTs, showing an up to 33% higher sensitivity; see Figure 4. In addition to the relative advantages in terms of sensitivity, we also find that PanCast achieves a 1% ($k = 2$) and 32% ($k = 10$) higher precision, respectively, compared to SPECTs averaged in the specificity value range of 0.6 to 0.8. These results suggest that PanCast can be significantly more effective in terms of sensitivity and precision while keeping the burden on the population (as measured by specificity) the same. The relative advantage of PanCast would further increase as more fine-grained and accurate information characterizing the infection risks of sites becomes available. While such knowledge is presently scarce, a system such as PanCast would allow us to gather it.

Discussion

In the face of an infectious pandemic, effective digital contact tracing solutions should scale with existing (manual) contact tracing, help acquire and leverage environmental information, and be accessible to anyone. In this paper, we introduce PanCast, a secure, privacy-preserving, inclusive, and interoperable digital contact tracing and risk notification system, specifically designed to address several limitations of SPECTs. Our simulations have shown that depending on deployment and adoption levels, PanCast is capable of matching and in some cases strongly exceeding the efficacy of SPECTs under realistic circumstances. At the same time it improves on inclusion and user transparency and provides strong guarantees on security and user privacy. We presently further elaborate on these aspects.

While manual contact tracing has played an important role in epidemic mitigation strategies [12], digital tracing solutions have not been able to prove their efficacy partly due to low adoption levels [12, 13]. As shown by the simulations, bidirectional integration of digital and manual tracing has the potential to yield great benefits for epidemic mitigation. While SPECTs cannot effectively interact with manual tracing, PanCast is designed to allow for a bidirectional interoperation—supporting manual tracing on the one hand and leveraging information obtained from manual tracing on the other hand. This is especially important as the adoption levels of digital systems are relatively low in most countries [24], making the x^2-adoption problem particularly severe. In addition to manual tracing, PanCast can also complement other digital tracing systems by implementing additional BLE-based contact tracing protocols in some of the deployed beacons—both centralized like PEPP-NTK [5] or decentralized like DP3T [9]. By doing so, any individual who tested positive and used these other BLE-based systems could be used to populate the PanCast risk database. In this context, PanCast could also provide relevant location dependent information to augment contact-tracing data from these other BLE-based systems, increasing the precision and specificity of their notifications.



(a) Low variance in transmission rate (b) High variance in transmission rate

Figure 4: Receiver operating characteristic (ROC) curves for PanCast and SPECTs. The variance level corresponds to the dispersion of transmission rates across different site types. Low and high variance is achieved by increasing the transmission rate at social sites by a factor $k = 2$ and $k = 10$, respectively, and decreasing the transmission rate at bus stops correspondingly by the same factor. Averaged over 100 roll-outs of the epidemiological simulations, the line represents the sensitivity (true positive rate) and specificity (true negative rate) achieved across operating thresholds of the system.

Moreover, in principle, PanCast also allows for the implementation of backward contact tracing [26], where not only contacts of the positively tested individual get traced, but also the contacts of their contacts.

In addition to aspects concerning the interoperability with other systems, effective digital tracing should be able to both leverage and help acquire circumstantial information that influences exposure. Firstly, disease transmission depends on environmental factors, such as indoor vs. outdoor, room size, speaking volume, ventilation and air quality [25]. In contrast to SPECTs, which only collect pairwise encounters, PanCast can leverage environmental information to improve infection risk predictions. Our simulations show that by taking into account environmental factors, PanCast achieves a significantly better precision and recall in the realistic scenario in which the transmission rate strongly depends on the environment. Secondly, while SPECTs are limited to contemporaneous transmissions where individuals have been close in both space and time, PanCast allows for capturing non-contemporaneous transmissions, occurring for instance after an infectious individual left a small and poorly ventilated room. Here, we point out that our simulations do not explicitly show this advantage. Currently, no reliable estimates for the true infection risk as a function of space-time distance and environmental parameters exists, and we opted to implement rough estimates in the simulations. However, we do believe that non-contemporaneous infections play a role [27] and it would be negligent to rule out such cases by design. Thirdly, the deployment of PanCast may allow for gathering the data needed to estimate these or other disease parameters, which could subsequently be used to evaluate the system and assist research. This also concerns the overdispersion of infections observed for COVID-19—most infected individuals do not infect anyone, while a few *superspreaders* infect many [17–20]. If digital contact tracing can identify infections events spatially, one may be able to trace entire clusters [26, 28]. As PanCast has access to location information, diagnosed users can help authorities identify potentially relevant superspreading events, which can in turn inform manual tracing or even media broadcasting actions.

Despite PanCast's favorable results in the simulations, the comparisons we make, and conclusions we draw have to be considered in the context of the epidemiological model employed. While we have used a suitable modeling framework, explicitly representing the sites at which transmissions occur, the simulations of both SPECTs and PanCast do not take into account practical considerations such as, e.g., battery status or imperfect beacon coverage within a specific site. The simulations moreover assumed that risk dissemination to the dongles is instantaneous. This assumption is reasonable if individuals' movement is mostly local, which is a good approximation especially if individuals' movement is constrained by government interventions [29]. It would be unrealistic if individuals frequently traveled to far away sites, in which case network beacons

around the world would have to broadcast extensive non-local risk information. However, our detailed protocol for risk dissemination (Section 4 in the PanCast's white paper [16]) ensures that even if individuals travel to far away sites, the time it takes for dongles to receive risk information stays within acceptable limits.

Ultimately, digital tracing can only be effective when easily deployed, widely accessible, and available at low cost. PanCast dongles and beacons require only minimal hardware and can therefore be provided cheaply. Additionally, as dongles require near-zero maintenance and no technical expertise on the part of users, the system can be used by individuals who, for financial or personal reasons, cannot or do not want to use smartphone apps. As these individuals might be over-represented among the elderly and socio-economically disadvantaged, PanCast specifically allows for the inclusion of groups that may be particularly vulnerable to the disease [30]. While a broad installation of PanCast would require the distribution of dongles to a population as well as equipping a significant amount of sites with beacons, the system could also be effectively employed in an incremental and/or local fashion at low costs. Individual institutions such as schools or companies could employ the system on their premises to trace and control epidemic spread among their members. Such local deployment would already provide valuable information on location-dependent transmission parameters, and allow tracking them over time as a virus may change its characteristics in response to evolutionary pressure generated by containment measures. This has the potential to contribute to a much more fine-grained, data-driven understanding of pandemics.

Materials and methods

System components

Figure 5 shows an overview of PanCast's architecture. PanCast comprises two types of Bluetooth Low Energy beacons (BLE-only and BLE+network), personal devices (dongles), terminals, and a backend platform that relays risk notifications and aggregates data for epidemiological analysis. All beacons and dongles are registered and authenticated with the backend, receive a secret key from the backend at the time of registration, and have a coarse-grained timer and a small amount of flash storage. When a user receives a dongle, they receive a list of one-time passwords (OTPs) that are also stored in the dongle. A diagnosed user can use these OTPs to authenticate to the dongle and to control the upload of data. Below, we elaborate on the components of the system and their respective functionalities.

Beacons. PanCast employs two types of beacons: *BLE beacons* are commodity, battery-operated BLE-only beacons and require no network connection to the backend. *Network beacons* also use BLE, but additionally require mains power and a network connection to provide connectivity to backend servers. The beacons serve two purposes:

(i) Every beacon provides a localization point in a specific place (e.g., an office, a bar, a public bus). A beacon periodically broadcasts an ephemeral id, its device id, and a location id. The device id and the location id are cryptographically signed by the backend, while the ephemeral id is generated by the beacon by hashing a beacon-specific secret key, its location id, and an epoch number derived from its local clock. Because the hash includes the epoch number, a fresh ephemeral id is broadcast in every epoch. This localizes every encounter between the beacon and a dongle to a specific epoch. The epoch length is set upfront to a small value, e.g., 15 minutes.

(ii) Network beacons additionally broadcast global risk information received from the backend periodically, using a protocol that balances privacy and efficiency.

Beacons are installed in specific locations, for instance, under the guidance of health authorities or by organizations that choose to place them on their premises. Stationary beacons broadcast a GPS coordinate or a named identifier (e.g., city, zip code). Beacons may also be installed in mobile locations such as trains or buses; in this case they broadcast ids that identify their trajectories. All beacons are registered with the

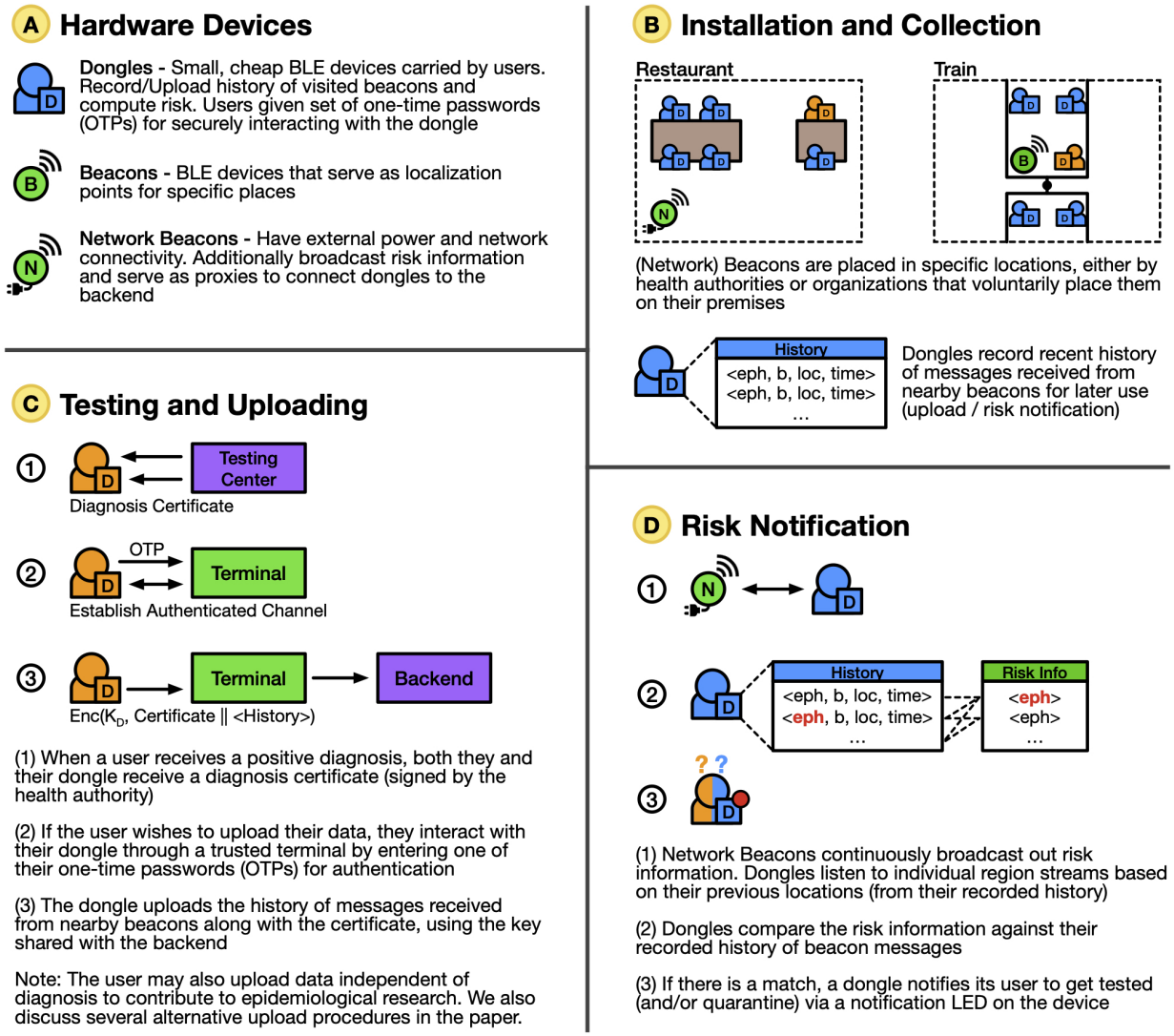


Figure 5: PanCast's hardware devices, installation and collection, testing and uploading, and risk notification.

backend using their id and their location id comprising their stationary coordinate or trajectory, as well as information about their location that may be epidemiologically relevant (e.g., indoor, outdoor, ventilation, air quality, ambient noise level). This information can be used by the backend when computing infection risks or performing epidemiological analyses. Simple, battery-operated BLE beacons can account for the majority of beacons. These can be cheap and easy to install, because they do not require mains power or network connectivity. We expect them to be installed wherever infection transmission is likely to occur (e.g., in places where people congregate). A smaller number of network beacons provide nearby dongles with risk information. To reduce installation costs, we expect them to be installed where power and network connectivity is already available, e.g., next to WiFi base stations.

Dongles. Dongles are small, simple devices that users can attach to a keyring, or wear on the wrist or around the neck. They operate off a coin battery and have a minimal user interface in the form of a LED that indicates risk status and battery condition, and a button to control the LED notification. Dongles

continuously listen for BLE transmissions from nearby beacons. They receive ephemeral ids from both types of beacons and store them along with the timestamps. When in proximity of a network beacon, they additionally receive risk information, compare that information with the device's stored history of ephemeral ids received from beacons, and alert the user in case they were in spatio-temporal proximity of a diagnosed individual under circumstances that suggest a possible transmission. Dongles usually operate entirely passively. They only transmit information when a user chooses to anonymously reveal their information via a user terminal.

Backend service. The backend maintains several databases. First, it maintains a database of registered beacons, their locations/trajectories, and the secret key used by each beacon to generate the unique sequence of ephemeral ids it broadcasts. A second database contains registered users, their dongles, and the cryptographic keys required to authenticate each dongle. A third database, called the risk database, contains the uploaded encounter histories of individuals who recently tested positive. Finally, a fourth database, called the epidemiology database, contains the encounter histories of healthy users who chose to contribute their data for epidemic analytics. The encounter database is available to health authorities for analytics, e.g., to identify hotspots, superspreading events, and to estimate epidemiological parameters. The backend uses this database to transmit risk information to network beacons, which then broadcast it to nearby user dongles.

User terminals. User terminals are provided at locations that issue dongles as well as health care facilities that do testing. Terminals allow users to connect to their dongles over BLE to change their privacy settings, inspect what data is recorded on their dongles, upload data to the backend and, when allowed or required by law, decide what subset of the recorded information they wish to upload when they are tested positive. Users can also use personal computers or smartphones as terminals to perform these tasks in the comfort of their homes, or use the smartphone of a care provider who visits their home.

System functionality

Data upload and risk dissemination Diagnosed individuals may share their encounter data with health authorities to enable dissemination of risk information. Upon receiving a positive test result, users obtain a certificate with which they can verify their infectious state when uploading the information stored in their dongles to the backend via a terminal. Once the backend receives a user's encryption key derived from one of the user's OTPs, it decrypts their dongle's encounter entries and verifies their consistency. If the entries are consistent, the backend adds the encounter entries to the risk database and/or the epidemiology database. The global risk information is broadcast to the users' dongles via network beacons. Dongles use this information to compute a risk score based on the number of matched ephemeral ids and other features of each encounter encoded in the beacon broadcast. If the risk exceeds a certain threshold, the dongle notifies the user via a LED so they can self-isolate and get tested. For the case that users carry their dongles openly visible in public spaces, the LED can be temporarily deactivated in order not to accidentally reveal a potentially incoming risk notification to third parties. Refer to Sections 3.3 and 4 in PanCast's white paper [16] for details on data upload and risk dissemination respectively.

Risk score calculation Whenever a user dongle receives risk information from the backend, it updates the owner's risk score locally (within the dongle). The individual risk score is proportional to the period during which the individual and diagnosed individuals were near the same beacons, as measured by the number of ephemeral ids contained in the risk information matching those stored in the dongle. Each ephemeral id may be weighted differently according to beacon-dependent parameters, such as indoor/outdoor, air quality, ventilation, and ambient noise. These parameters were stored by the dongle when it received the beacon's transmission. How these features are weighted depends on parameters provided by the backend as part of the risk information. The parameters can be determined by the backend using machine learning techniques [21] and reflect the latest scientific knowledge about the disease.

Security and Privacy

From a security point of view, the system's data collection and risk dissemination protocols are immune to many of the attacks that SPECTs are vulnerable to [31, 32]. Moreover, PanCast's simpler devices have a smaller attack surface compared to smartphones. From a privacy point of view, PanCast provides significant improvements over SPECTs. User devices never transmit any information unless their owners choose to do so (e.g., after the owner is tested positive). When an individual is diagnosed, they explicitly consent to the transmission and can select the information they wish to transmit from their device to the health authority. Even the most privacy-conscious users who disclose nothing and never transmit anything from their dongles (i.e., only listen passively) receive risk notifications arising from space-time proximity to other diagnosed users who choose to disclose information after their diagnosis. Moreover, a user learns about potential risks in a visited location only when a diagnosed individual visited the location within a certain time window, but does not otherwise learn the location history of a diagnosed individual. Finally, the risk broadcast provides strong differential privacy guarantees for the number of risk entries and the number of diagnosed individuals contained in a risk broadcast. This ensures that an adversary, even one in possession of offline information about some users, learns nothing about the health status or whereabouts of the remaining users through the system. Refer to Section 6 in PanCast's white paper [16] for a detailed discussion of PanCast's security and privacy characteristics.

Epidemiological Simulations

Epidemiological model To study the effects of PanCast, we simulate different epidemic scenarios under various contact tracing policies using a spatiotemporal epidemic model recently introduced by Lorch et al. [21]. This agent-based model combines points of interests (POIs) of real cities and regions from OpenStreetMap [22] and population density data from Facebook [33] together with assumptions on the mobility behavior to simulate mobility traces of individuals in a given area. The spatial distribution of visits generated by the event-based mobility model follows the gravity model [34]. It assumes that the probability that an individual visits a specific site decreases with the distance between their household and the site. For a detailed overview over the relevant parameters refer to [21]. By way of modeling transmissions explicitly at the sites where they occur, the model is directly applicable to simulating the effects of our proposed system and comparing it with existing SPECTs. The model explicitly captures non-contemporaneous infections, allows for variability in transmission rate at different POI types, and can faithfully incorporate contact-tracing measures at POIs, such as manual contact tracing. Disease dynamics are described by a SEIR compartmental model [35] and are modeled using temporal point processes [36]. This allows the model to explicitly represent events when individuals check in at POIs, get in contact with and infect each other, change epidemiological state, and are affected by or interact with TTI measures.

Model parameters Site-specific and household transmission rates are estimated using Bayesian optimization, as described by Lorch et al. [21], by matching the cumulative number of positively tested individuals in the simulation with real COVID-19 case data over a period of eight weeks early in the pandemic. Following previous studies, we set the proportion of asymptomatic individuals among all infected individuals in the population to 0.4 [37–39] and the relative asymptomatic transmission rate to 0.55 [40]. After an infectious individual leaves a site, we assume that their relative non-contemporaneous transmission rate to others decays with a half-life ten times shorter than that estimated for aerosols under laboratory conditions [27]. We truncate the possibility of non-contemporaneous transmission when the relative transmission rate drops below 10%, which occurs 21 minutes after an individual leaves a site. The remaining disease parameters related to the SEIR progression from having been exposed to recovery are taken from the literature [41–46] and follow the values stated in [21].

Contact tracing Our simulations include manual contact tracing as well as digital contact tracing using either PanCast or SPECTs. Independent of the tracing method, whenever a contact person of a diagnosed individual gets successfully traced, the infection risk is estimated by taking into account the duration of

the contact and possibly environmental factors (only for PanCast). Individuals with infection risk above a certain threshold are quarantined for two weeks, get tested within the next 24 hours, and receive the outcome of the test 48 hours later. We choose the infection risk threshold to correspond to a 15 minute contact with a symptomatic individual in the model, which is in accordance with SPECTs currently employed in Germany, Switzerland, the United Kingdom, France, and Australia [47–51].

All simulations using PanCast or SPECTs also implement manual tracing by assuming that a proportion $p_{\text{reachable}} = 0.5$ of visitors leave their contact details at social, office and education sites, so that they can be reliably contacted, e.g., via phone. Upon receiving a positive test result, we assume that every individual participates in a manual contact-tracing interview independent of their participation in digital tracing. In the tracing interview, we assume a person only remembers a fraction $p_{\text{recall}} = 0.1$ of their visit history of the past 14 days. This accounts for the fact that most people cannot remember every place they have visited and some people might not be willing to share any data or cannot be reached for an interview. Note that our choice of the parameter values is conservative as PanCast directly benefits from larger values of p_{recall} and $p_{\text{reachable}}$ via interoperation with manual tracing as opposed to SPECTs. If an infected person i recalls a visit at which they encountered a person j and individual j is reachable, the contact between i and j is traced successfully and j gets quarantined and tested if the exposure risk exceeds the previously stated 15 minute threshold.

For SPECTs, we assume that a certain proportion p_{digital} of the population has adopted the digital tracing technology. When an individual using SPECTs is tested positive, all contacts of the past 14 days that also use the system get traced. SPECTs and manual tracing do not interoperate but complement each other. Any person may still participate in manual contact tracing, but no contact information can be shared between the two systems as SPECTs do not operate using space-time information.

For PanCast too we assume that the same proportion p_{digital} of the population has adopted the system and is carrying dongles. We place beacons at a proportion p_{beacon} of the sites. This can be done at random or strategically, by taking into account quantities related to the site-specific probability of infections, e.g., by ranking the sites according to their integrated visit time. Whenever a person carrying a dongle gets tested positive, all contacts at sites with beacons who also carry dongles get traced. In addition, the information can be used to trigger manual tracing action at all sites registered by the dongle of the positive-tested individual (i.e., PanCast supports manual tracing). Likewise, when a person who tested positive does not carry a dongle but participates in a manual contact interview and recalls a visit to a site with a beacon, all individuals carrying dongles at this site can be traced (i.e., manual tracing supports PanCast). In our simulations, we assume that individuals receive PanCast's risk broadcast instantaneously. In practice, this assumption may be violated, but PanCast's protocol for risk dissemination (refer to Section 4 in PanCast's white paper [16]) ensures that the time it takes for dongles to receive risk information stays within acceptable limits.

Data Availability

Code and data to reproduce our epidemiological simulations are available at https://github.com/covid19-model/simulator/tree/beacon.

References

[1] Sebastian Contreras et al. The challenges of containing SARS-CoV-2 via test-trace-and-isolate. *Nature Communications*, 12:378, January 2021.

[2] Matthias Linden et al. Überschreitung der Kontaktnachverfolgungskapazität gefährdet die Eindämmung von COVID-19. *Dtsch. Arztebl. International*, 117(46):790–791, 2020.

[3] More than 1,300 workers test positive: Germany fights to control coronavirus spread at meat plant. `https://www.thelocal.de/20200622/more-than-1300-workers-contract-coronavirus-dark-side-of-meat-industry-sparks-row-in-germany`. Accessed on 15 December 2020.

[4] Entwurf eines Gesetzes zur Verbesserung des Vollzugs im Arbeitsschutz. https://dip21.bundestag.de/dip21/btd/19/219/1921978.pdf. Accessed on 15 Dec 2020.

[5] PEPP-PT (20 April 2020) Data Protection and Information Security Architecture. https://github.com/pepp-pt/pepp-pt-documentation/blob/master/10-data-protection/PEPP-PT-data-protection-information-security-architecture-Germany.pdf. Accessed on 15 December 2020.

[6] OpenTrace. https://github.com/opentrace-community/opentrace-android. Accessed on 15 December 2020.

[7] TraceTogether. https://www.tracetogether.gov.sg/. Accessed on 15 December 2020.

[8] ROBERT: ROBust and privacy-presERving proximity Tracing. https://github.com/ROBERT-proximity-tracing/documents/blob/master/ROBERT-specification-EN-v1_0.pdf. Accessed on 15 December 2020.

[9] Decentralized Privacy-Preserving Proximity Tracing. https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf, 2020.

[10] Justin et al. Chan. Pact: Privacy-sensitive protocols and mechanisms for mobile contact tracing. *IEEE Data Engineering Bulletin*, 43(2):15–35, July 2020.

[11] TCN Coalition. https://github.com/TCNCoalition/TCN. Accessed on 15 December 2020.

[12] Isobel Braithwaite, Thomas Callender, Miriam Bullock, and Robert W Aldridge. Automated and partly automated contact tracing: a systematic review to inform the control of covid-19. *The Lancet Digital Health*, 2020.

[13] Early Evidence of Effectiveness of Digital Contact Tracing for SARS-CoV-2 in Switzerland. https://github.com/digitalepidemiologylab/swisscovid_efficacy/blob/master/SwissCovid_efficacy_MS.pdf. Accessed on 15 December 2020.

[14] Global contact tracing app downloads lag behind effective levels. https://qz.com/1880457/global-contact-tracing-app-downloads-lag-behind-effective-levels/. Accessed on 15 December 2020.

[15] Google/Apple Exposure Notification (GAEN) system. Exposure Notifications: Using technology to help public health authorities fight COVID-19. https://www.google.com/intl/en_us/covid19/exposurenotifications/.

[16] Gilles Barthe, Roberta De Viti, Peter Druschel, Deepak Garg, Manuel Gomez-Rodriguez, Pierfrancesco Ingo, Matthew Lentz, Aastha Mehta, and Bernhard Schölkopf. Pancast: Listening to bluetooth beacons for epidemic risk mitigation. *arXiv:2011.08069*, 2020.

[17] Muge Cevik, Julia Marcus, Caroline Buckee, and Tara Smith. SARS-CoV-2 transmission dynamics should inform policy. *SSRN 3692807*, 2020.

[18] Agus Hasan et al. Superspreading in early transmissions of COVID-19 in indonesia. *Scientific Reports*, 10(1):22386, 2020.

[19] Max SY Lau, Bryan Grenfell, Michael Thomas, Michael Bryan, Kristin Nelson, and Ben Lopman. Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. *Proceedings of the National Academy of Sciences*, 117(36):22430–22435, 2020.

[20] Yunjun Zhang, Yuying Li, Lu Wang, Mingyuan Li, and Xiaohua Zhou. Evaluating transmission heterogeneity and super-spreading event of COVID-19 in a metropolis of China. *International Journal of Environmental Research and Public Health*, 2020.

[21] Lars Lorch, William Trouleau, Stratis Tsirtsis, Aron Szanto, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. Quantifying the effects of contact tracing, testing, and containment measures in the presence of infection hotspots. *arXiv:2004.07641*, 2020.

[22] OpenStreetMap. `https://www.openstreetmap.org/`. Accessed on 15 Dec 2020.

[23] RKI. Übersicht zu aktuellen und früherer Zahlen und Fakten zur Corona-Warn-App; 18.12.2020. `https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/WarnApp/Archiv_Kennzahlen/WarnApp_KennzahlenTab.html`. Accessed on 23 December 2020.

[24] Statista. Adoption of government endorsed COVID-19 contact tracing apps in selected countries as of July 2020. `https://www.statista.com/statistics/1134669/share-populations-adopted-covid-contact-tracing-apps-countries/`. Accessed on 15 December 2020.

[25] Baoming Wang, Hui Chen, Yik Lung Chan, and Brian G Oliver. Is there an association between the level of ambient air pollution and COVID-19? *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 319(3):L416–L421, 2020.

[26] William J. Bradshaw, Ethan C. Alley, Jonathan H. Huggins, Alun L. Lloyd, and Kevin M. Esvelt. Bidirectional contact tracing could dramatically improve COVID-19 control. *Nature Communications*, 12(1):232, January 2021.

[27] Neeltje Van Doremalen et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England Journal of Medicine*, 382(16):1564–1567, 2020.

[28] Akira Endo et al. Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks. *Wellcome open research*, 5:239–239, October 2020. Publisher: F1000 Research Limited.

[29] Germany to limit non-essential travel to 15-kilometre radius. `https://www.reuters.com/article/us-health-coronavirus-germany-travel-idINKBN29A1MH?edition-redirect=in`. Accessed on 12 Jan 2020.

[30] Elizabeth Williamson et al. Opensafely: factors associated with covid-19-related hospital death in the linked electronic health records of 17 million adult nhs patients. 05 2020.

[31] Serge Vaudenay. Centralized or Decentralized? The Contact Tracing Dilemma. 2020.

[32] Lars Baumgärtner et al. Mind the gap: Security & privacy risks of contact tracing apps. *arXiv:2006.05914*, 2020.

[33] Facebook Data for Good. `https://dataforgood.fb.com`. Accessed on 15 Dec 2020.

[34] George Kingsley Zipf. *Human behavior and the principle of least effort*. Addison-Wesley Press, Oxford, UK, 1949.

[35] Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.

[36] Abir De, Utkarsh Upadhyay, and Manuel Gomez-Rodriguez. Temporal point processes. Technical report, Technical report, Saarland University, 2019.

[37] Hiroshi Nishiura et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International journal of infectious diseases*, 94:154, 2020.

[38] Enrico Lavezzo et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo. *Nature*, 584(7821):425–429, 2020.

[39] Luca Ferretti et al. Quantifying dynamics of SARS-CoV-2 transmission suggests that epidemic control and avoidance is feasible through instantaneous digital contact tracing. *Science*, 2020.

[40] Ruiyun Li et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490):489–493, 2020.

[41] World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). `https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf`.

[42] Xi He et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 26(5):672–675, May 2020.

[43] Roman Wölfel et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*, 581(7809):465–469, 2020.

[44] Stephen A. Lauer et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, 2020.

[45] Dawei Wang et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA*, 2020.

[46] Natalie M. Linton et al. Incubation period and other epidemiological characteristics of 2019 novel Coronavirus infections with right truncation. *Journal of Clinical Medicine*, 9(2), 2020.

[47] Corona-Warn-App Team. Corona-Warn-App: Documentation. `https://github.com/corona-warn-app/cwa-documentation/`.

[48] Federal Office of Public Health, Switzerland. How the SwissCovid app issues warnings. https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/swisscovid-app-und-contact-tracing.html#-1674020142.

[49] United Kingdom National Health Service. NHS COVID-19 app risk-scoring algorithm. `https://covid19.nhs.uk/risk-scoring-algorithm.html`.

[50] Ministre des Solidarites et de la Sante. TousAntiCovid : responses a vos questions. `https://solidarites-sante.gouv.fr/soins-et-maladies/maladies/maladies-infectieuses/coronavirus/tousanticovid`.

[51] Australian Government. Background to COVIDSafe. `https://covidsafe.gov.au/background.html`.

Acknowledgements

We would like thank Nuria Oliver, Viola Priesemann, Nasim Rahaman, Peter Schwabe, Clara Scheidewind, Michael Meyer-Hermann, and the CIFAR contact tracing working group for helpful feedback.

Author Contributions

G.B., P.D., D.G., M.G.R., A.M. and B.S. designed research; G.B., R.D.V., P.D., D.G., M.G.R., P.I., H.K., M.L., L.L, A.M. and B.S. performed research; L.L. and H.K. analyzed data; and G.B., R.D.V., P.D., D.G., M.G.R., P.I., H.K., M.L., L.L, A.M. and B.S. wrote the paper.

Competing Interests

We declare no competing interests.

Materials and Correspondence

Correspondence should be addressed to Manuel Gomez-Rodriguez (manuelgr@mpi-sws.org) and Bernhard Schölkopf (bs@tuebingen.mpg.de).