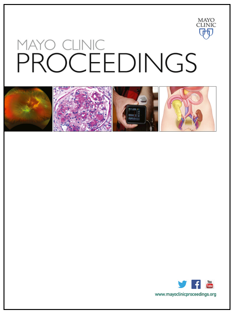


Contact Tracing to Manage COVID19 Spread – Balancing Personal Privacy and Public Health

Suraj Kapa, MD, John Halamka, MD, MS, Ramesh Raskar, PhD

PII: S0025-6196(20)30424-9

DOI: https://doi.org/10.1016/j.mayocp.2020.04.031

Reference: JMCP 2884

To appear in: *Mayo Clinic Proceedings*

Received Date: 14 April 2020

Revised Date: 21 April 2020

Accepted Date: 28 April 2020

Please cite this article as: Kapa S, Halamka J, Raskar R, Contact Tracing to Manage COVID19 Spread – Balancing Personal Privacy and Public Health, *Mayo Clinic Proceedings* (2020), doi: https://doi.org/10.1016/j.mayocp.2020.04.031.

Contact Tracing to Manage COVID19 Spread – Balancing Personal Privacy and Public Health

Suraj Kapa,MD[1]; John Halamka,MD, MS[2]; Ramesh Raskar,PhD[3]

[1]Department of Cardiovascular Medicine,Mayo Clinic,Rochester,MN,USA

[2]President, Mayo Clinic Platform,Mayo Clinic,Rochester,MN,USA

[3]MIT Media Lab,Masschusetts Institute of Technology,Cambridge,MA,USA

Short Title: Contact tracing and personal privacy

Words:1200

Figures:1

Disclosures:SK and RR are working on a contact tracing solution (safepaths.mit.edu). Neither derives nor will derive financial benefit.

Corresponding author:

Suraj Kapa, MD

Associate Professor of Medicine

Mayo Clinic

200 First Street SW

Rochester,MN,55905

Phone:(507)538-6325

Fax:(507)255-2550

E-mail:kapa.suraj@mayo.edu

The COVID19 pandemic has opened up an important conversation regarding how we balance interests of privacy with public health. Examples of contact tracing tools to identify exposed individuals have shown promise in facilitating early responses to minimize disease spread. These tools use geographic location data from SmartPhones and other devices to understand when and where infected individuals may have interacted with others. These tools, however, have come at the expense of personal data privacy, leading to concern over widespread use despite public health benefits.[1]

The decision to deploy contact tracing at a regional, national, or global level needs to take into consideration a balance between individual data privacy and societal benefit. The concept of "data sharing" to facilitate public good is not new – for example, phone-based map applications give real-time traffic data by aggregating user information via global positioning systems (GPS) active on devices. These data libraries are maintained by private companies with permissions enabled by device users. In many cases (eg, with credit cards, smartwatches, and other devices) the user may be unaware of data being stored on the device. However, citizens make decisions to share data in return for various benefits (targeted advertisements, instructions on optimal driving directions, and so on). In the United States, this sharing of data usually occurs with private corporate entities. However, when sharing data for contact tracing, it is expected that the data is aggregated by a central authority such as a state or national government.

In the midst of the COVID19 pandemic, privacy concerns related to sharing historical location data have been raised, even as public health benefits have become obvious.[2] For example, studies in South Korea and Singapore suggest that use of citizens' location data facilitated mitigation without the same level of societal lockdowns used in Europe and the United States. While additional waves of disease are possible, the initial results are promising. However, concerns have been raised regarding acceptability of surrendering privacy over location data to governmental entities. This has raised considerable debate in public and academic spheres regarding how to balance privacy risks against public health benefits.[3,4]

The balance between public good and private data ownership thus comes to a forefront with contact tracing. In using contact tracing, the goal is to understand the movement history of infected individuals to then be able to inform healthy users who may have crossed paths with an infected individual of potential exposure. In an ideal scenario, users who were exposed would then be isolated or tested so as to mitigate further spread. (Figure 1A)

There are two scenarios for contact tracing – one that allows for user identification and one that allows for individual privacy. (Figure 1B) In Scenario 1 (most commonly used to date), a central authority aggregates data and responds to that data via direct interaction with the user (who is identifiable) or via law enforcement. In Scenario 2, the user's data stays encrypted when provided to a central authority.[5] In this iteration, the central authority never knows who the data originated from, but does know if data was from an infected individual or not. Isolation of exposed individuals remains feasible by a user's personal device being able to access aggregated data, recognizing possible intersection with an infected individual, and informing the user of potential exposure. Thus, in Scenario 2, no central authority or law enforcement body is aware of the identities of exposed individuals, but users can still be made aware of potential exposure and respond accordingly (eg, pursuing testing or self-isolating).

Such a privacy-first approach may still be met with skepticism but ideally will allow for implementation of tools to mitigate the current pandemic and potentially future outbreaks. By enabling individuals to understand exposure history, to have full control over their data, and to share their data privately, it may be possible to balance privacy and public health. Ultimately, such data may allow for mitigation of spread by cutting "branches of spread" earlier on.

There are several limitations, though. Societal level benefit is dependent on broad and diverse user adoption. This may occur through legal regulations enforcing use or public addresses to raise awareness and adoption. In many countries where contact tracing is being considered, legal compulsion as a method to raise adoption is being debated. Also, modern perspectives on trust in government may vary, and this may impact perceived importance of personal data privacy. Finally, whether privacy enabled interventions reduce the efficacy of contract tracing due to dependence on private user response rather than direct enforcement by a central authority requires further study.

Conclusion

Given promise of digital solutions to mitigate disease spread, it is critical the science of contact tracing be explored, particularly given their cost-efficiency and scalability. It is feasible to manage privacy and public good by innovating appropriate solutions for how data is aggregated and users are informed of exposures. However, potential benefit to address waves of the current pandemic or future outbreaks can't be under-stated.

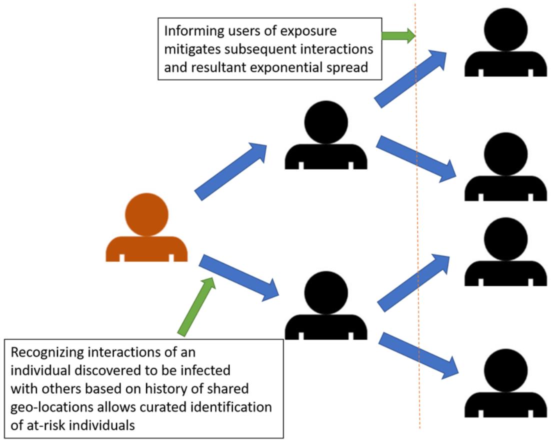
References

1. Servick K. Cellphone tracking could help stem the spread of coronavirus: Is privacy the price? Science [Internet]. 2020. [cited 2020 Mar 22]. Available from: https://www.sciencemag.org/news/2020/03/cellphone-tracking-could-help-stem-spread-coronavirus-privacy-price.
2. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ. Presymptomatic transmission of SARS-CoV-2 – Singapore, January 23-March 16, 2020. MMWR Morb Mortal Wkly. 2020; 69: 411-5.
3. Lee J, Sun J, Wang F, Wang S, Jun CH, Jiang X. Privacy-preserving patient similarity learning in a federated environment: Development and analysis. JMIR Med Inform. 2018; 6: e20.
4. Raskar R, Schunemann I, Barbar R, et al. Apps gone rogue: Maintaining personal privacy in an epidemic. arXiv:2003.08567. [Preprint]. 2020 [cited 2020 Mar 19]; [15 p.]. Available from: https://arxiv.org/abs/2003.08567.
5. Berke A, Bakker M, Vepakomma P, Raskar R, Larson K, Pentland A. Assessing disease exposure risk with location data; A proposal for cryptographic preservation of privacy. arXiv:2003.14412. [Preprint]. 2020 [cited 2020 Apr 8]; [15 p.]. Available from: https://arxiv.org/abs/2003.14412.

Figure 1

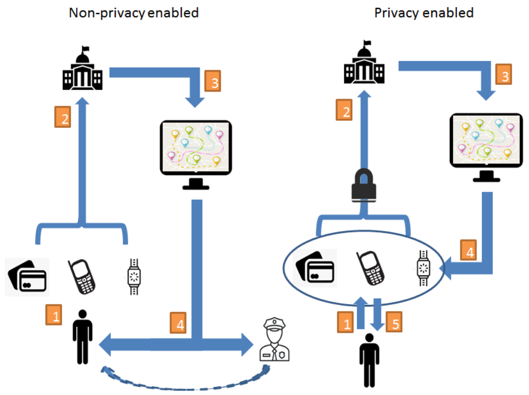
A. Infectious spread occurs exponentially through interactions that occur starting with an infected individual. By recognizing index exposures, it is possible to isolate exposed individuals from the population, mitigating further spread.

B. Two examples of contact tracing are shown. In the left panel, identifiable user data is aggregated on trackable devices (1), sent to a central authority (2), compiled into a centralized data set (3), and because all users are identifiable, the authority approaches users or law enforcement directly to isolate exposed individuals (4). In the right panel, user data exists on personal devices (1) but is fully deidentified, when transmitted to a central authority, keeping the user's identity private (2). The central authority then aggregates data only knowing infectious status of a specific deidentified location history (3) and this data is available for import to the user's devices where personal location history has been stored (4). It is then the user's device that informs the user of potential exposure due to overlap with infected paths (5).



Informing users of exposure mitigates subsequent interactions and resultant exponential spread

Recognizing interactions of an individual discovered to be infected with others based on history of shared geo-locations allows curated identification of at-risk individuals


Non-privacy enabled
Privacy enabled
1
2
3
4
1
2
3
4
5