An evaluation of privacy policies used in digital contact tracing apps for COVID-19

Erin Elizabeth Greig[#], **Reika Grace Shimomura**[#]

Health Humanities Lab, Humanities Research Center, Duke Kunshan University, Kunshan, China

Contributions: (I) Conception and design: Both authors; (II) Administrative support: Both authors; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: Both authors; (V) Data analysis and interpretation: Both authors; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

[#]These authors contributed equally to this work.

Correspondence to: Reika Grace Shimomura. Health Humanities Lab, Humanities Research Center, Duke Kunshan University, Kunshan, China. Email: rs510@duke.edu.

Background: After coronavirus disease 2019 (COVID-19) spread across the globe, quarantine regulations were supplemented by digital contact tracing initiatives in the form of mobile applications and other surveillance. This paper examines digital contact tracing on a global scale using English privacy policy information from mobile digital contact tracing applications (DCTAs) compared to guidelines on digital proximity tracing from the World Health Organization (WHO) published in May 2020.

Methods: Based on the WHO guidelines, six criteria were created to evaluate the ethical development of DCTAs using privacy policies: data deletion, geolocation turned off, time limitation, third parties sharing off, non-commercialization of data, Internet Protocol address (IP) or Unique Device Identifier (UDID) removed. Each criterion was answered by yes, no, or not mentioned to determine compliance with WHO guidelines.

Results: The most respected criterion was the non-commercialization of data, where more than 85% of the applications specified not using the data for commercial purposes. A unique difference was the tracking of geolocation, where 66.7% of applications use Bluetooth while 21.4% rely on geolocation collection. A concern arose from 45.2% of applications not mentioning whether an individual's identifier like IP and UDID would be collected or not. On top of this, privacy policies that satisfy the least respected WHO criterion are more likely to satisfy other WHO criteria, and those that failed the least respected criteria are more likely to fail in other criteria.

Conclusions: The results of this paper suggest that tracking of geolocation is the area that has the largest area for improvement in privacy policy development. To protect the privacy of an individual, following the WHO regulations worldwide is recommended and being transparent on the privacy policies by addressing all information in the criteria is essential. In the case of the COVID-19 pandemic, the regulations encourage the use of QR codes and Bluetooth in digital contact tracing since they minimize geolocation tracking, although sharing of data with third parties is another concern as it is up to the discretion of the developer with no current international body regulation.

Keywords: Coronavirus disease 2019 (COVID-19); digital contact tracing; app; privacy policy; privacy

Received: 06 April 2022; Accepted: 27 June 2022.

doi: 10.21037/jphe-22-27

View this article at: https://dx.doi.org/10.21037/jphe-22-27

Introduction

At the end of January 2020, the World Health Organization (WHO) declared the outbreak of coronavirus disease 2019 (COVID-19) a Public Health Emergency of International Concern (PHEIC) and on the 11th of March 2020 declared the outbreak a pandemic (1). Since then, the virus has spread globally, impacting over 200 confirmed countries (2).

Identifying who needs to be quarantined due to known contact with an infected individual, if manually done, involves direct communication between people if they had known meeting times. However, if they were also in contact with people in public, then knowing the exposures and taking action to inform people is difficult to conduct manually, and this approach is inefficient in preventing the new carrier's spread of the virus.

Contact tracing has been carried out in past pandemics and disease outbreaks, but prior to the technological boom of the recent century, this was done through phone calls and door-to-door check-ins with community health workers (CHWs) (3). Now, with an average of 107 mobile cellular subscriptions per 110 people globally, many countries pursued a digital approach to tracing (4). Contact tracing aids in quarantine implementation with faster close-contact identification (5). The implementation of digital contact tracing applications (DCTAs) can be divided into centralized and decentralized applications, where the definitions are given based on the approaches of storing the data and generating data to notify the users. Decentralized contact-tracing matches contacts and notifies users by downloading the contact database from a server while centralized contact-tracing collects anonymous IDs to a central server as a centralized database and notifies users (6).

While the current literature on the implementation of DCTAs highlights the importance of privacy and ethical concerns as they affect the perception of DCTAs and participation, global comparisons are difficult and are limiting the research (7). Research done by Akinbi *et al.* limits the systematic review on the challenges of DCTAs to East Asia. The primary studies on DCTAs show privacy concerns for 45% of the 61 selected primary studies. User behavior and participation (16%) highlight that around a quarter of the studies want to understand how to increase participation in surveillance. In addition, ethical issues (12%) and lack of trust (10%) were also covered (8). Although recommendations have been made on how to address the ethical and privacy concerns as well as how the governments are using the DCTAs, such as having ethical and legal frameworks, current literature is heavily focused on the perceptions of the users of DCTAs (9). Without knowing what the status is of the DCTA developers or providers, there is no actionable feedback for the DCTA developers to improve the perception of DCTAs from ethics and privacy areas.

DCTA categorizations

In this global age, the need for implementing digital contact tracing and the transforming of previously inefficient methods to the platform of phone applications is relatively new and makes users vulnerable to emerging privacy protection policies. The most common DCTA's communication technologies use WiFi, QR code scanning, Bluetooth, and GPS software (6). QR codes require more participation from the public, as they are responsible for manually scanning posted QR codes to check-in at locations, as happens in Australia for example (10).

Bluetooth and GPS, on the other hand, are more passive for the individual, typically collecting data without input from the user. The drawback, however, is that some software can collect and store personal information, transactions, habits, and personal device information (11). GPS typically shares more information than Bluetooth technology does, by sharing real-time location data, and connecting that information to the identification data, leading the developer of the apps and linked government to possibly violate privacy policies, or come close to privacy infringements compared to that of Bluetooth apps (11). For Bluetooth based apps, reliance is on the proximity between devices to identify the contact without an actual location at that point in time (11). Therefore, it delivers the minimum amount of personal data possible to the developer and the user who receives a contact tracing alert would not know where or who they were in contact with.

Ethical concerns

The question of how to balance the privacy of users and the effectiveness of contact tracing as a tool for prevention arises when examining the country paths to DCTAs. If the rate of use of the tracking system is low, then the contact tracing is not effective, akin to not having the tracking system at all. Even if the public uses contact tracing apps, how can people be sure that information will not be used against them—such as the information tied to their personal identity—

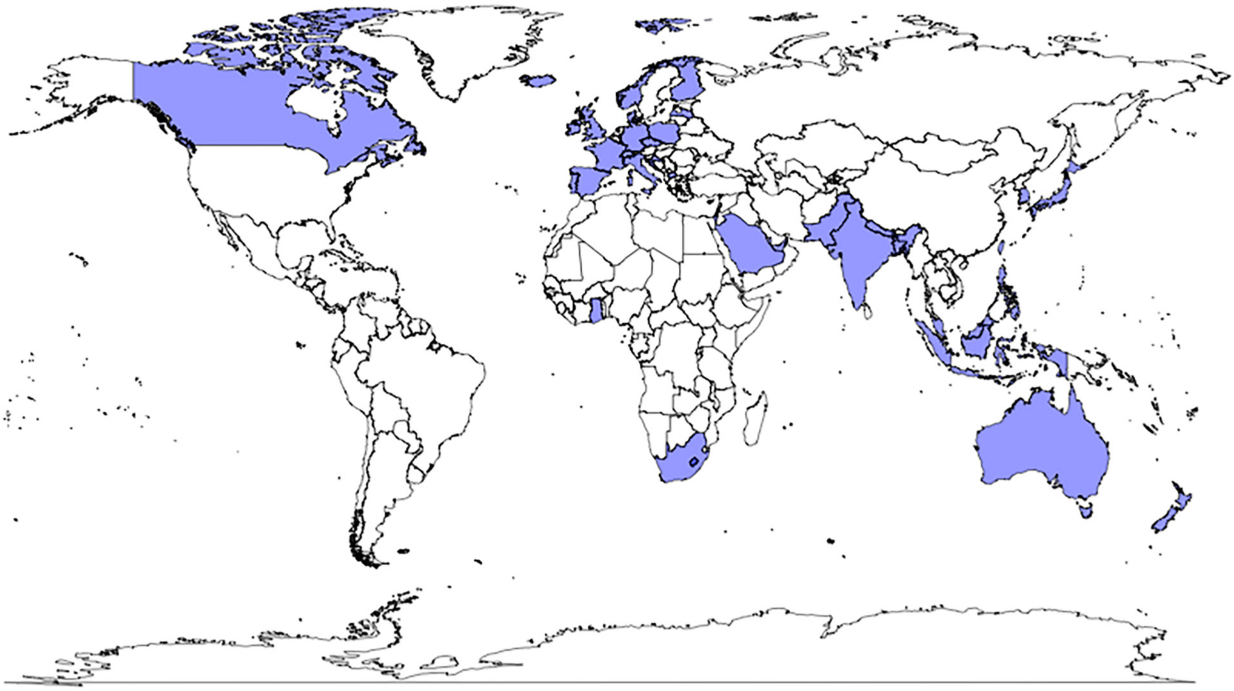

Figure 1 Map of countries with privacy policies used in analysis.

and how can leaks of that information to third parties such as employers be prevented? These concerns tell us the importance of protecting privacy and rights from an ethical perspective to increase the trust of users of contact tracing apps and increasing users to effectively reduce COVID-19 transmission through previously unidentifiable carriers.

One of the ways to facilitate the increased use of DCTAs is to reassure users by improving the readability of privacy policies, by looking into how the privacy policies are worded and phrased (12). However, what the apps' developers are promising to do and what they are not promising to do are underexamined.

WHO guidelines

The WHO developed interim guidance to guide the use of digital proximity tracking technologies for COVID-19 contact tracing, published on May 28th, 2020. This included 17 global principles for DCTAs and their corresponding privacy policies (13). In this paper, privacy policies were examined using the WHO's guidelines on digital proximity tracing to recognize what is and is not mentioned on the DCTA privacy policies across the countries.

Methods

Selection of DCT apps for privacy policy analysis

Based on the database provided by Pagliari (14), countries that either had limited contact tracing or comprehensive contact tracing during a period between May 17th to July 31st, 2021 were listed. There were 155 countries that qualified for these criteria. Countries using DCTAs up to July 31, 2021 were then selected for continuity, which totaled to 64 countries. Only the DCTAs that are endorsed or sponsored by the government (official DCTAs) were selected. For countries that had multiple official DCTAs in different regions of the country, they were all included, such as the UK with three official DCTAs: StopCOVID NI, NHS COVID-19, and Protect Scotland. Lastly, if the privacy policies were not available in English by the app developer or the government's website prior to purchase or download of the app, they were omitted, leaving 42 DCTAs from countries shown in *Figure 1*.

Privacy policies of official DCTAs with other languages were not translated as the precision of the translation cannot be guaranteed without a native speaker and native fluency was not accessible for the research team. If the official

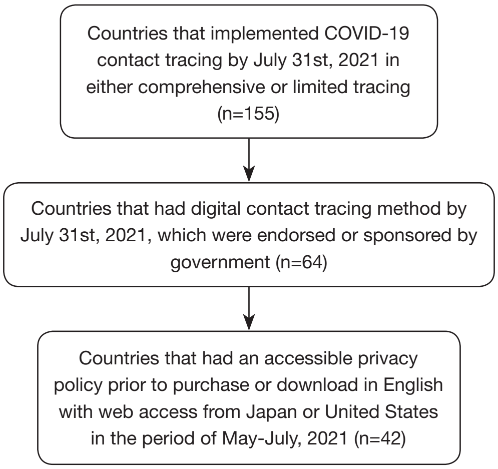
Countries that implemented COVID-19 contact tracing by July 31st, 2021 in either comprehensive or limited tracing (n=155)
Countries that had digital contact tracing method by July 31st, 2021, which were endorsed or sponsored by government (n=64)
Countries that had an accessible privacy policy prior to purchase or download in English with web access from Japan or United States in the period of May-July, 2021 (n=42)

Figure 2 Flow chart of DCTA selection process. DCTA, digital contact tracing application.

DCTAs provided a privacy policy in English on Apple App Store, Android Apps on Google Play, or their official website, then they were included for the study. The selection process of DCTAs is shown in the flow chart in *Figure 2*.

It is important to note that distinctions between the centralized and decentralized approaches, as well as limited and comprehensive tracing methods were not made when evaluating the privacy policies.

Development of analytical framework based on WHO guidelines

To do this evaluation, six overarching categories were created based on the interim guidance of 17 principles from the WHO which are coded as "data deletion", "geolocation turned off", "time limitation", "third parties sharing off", "data commercialization", and "IP or UDID removed" (13). With those categories, each of the 42 privacy policies was evaluated by the explicit, or lack of, comment on these categories. There are some aspects of the WHO's suggested principles that were not incorporated into the six categories such as accuracy of the algorithm, which comes after the use of the DCTAs rather than the category to be mentioned on privacy policies before its use (13).

Results

For the first category, "data deletion", refers to DCTAs allowing users to withdraw from using the DCTA and have the data related to the users deleted, which was inspired by the principle, "voluntariness" (13). The second category, "geolocation turned off" refers to the apps tracking locations using geolocation or GPS. It is preferred to not use the physical location to be privacy-preserving, as other technology such as Bluetooth uses the proximity of the devices. The third category, "time limitation," refers to whether the DCTAs implement a time limitation for the data storage or data retention. Time limitation for the data storage was separated from the app being deleted automatically after a certain time span in these criteria, as the former one refers to whether the older data gets removed while the user is active on the app, while the latter refers to whether the user can withdraw from being an active user at any point in time or whether the app will be deleted after it becomes irrelevant (i.e., COVID-19 is no longer a concern). The fourth category is "third parties sharing off" which indicates that the data is stored and used by the developer and/or specified governmental organizations only. The fifth category is "non-data commercialization" which refers to not using data for commercial or advertising purposes. The sixth category is "IP or UDID removed", which indicates that the DCTAs will not record or store an IP or UDID which can be used to identify an individual.

The most respected criteria were "non-data commercialization" and "data deletion", where more than 85% of the applications specified not using them for commercial purposes and privacy policies explicitly stated that data will be deleted when users withdraw from DCTA as shown in *Table 1*.

Unique differences were shown in "geolocation turned off" and "third parties sharing off" criteria. There were 66.7% of DCTAs that used Bluetooth while 21.4% relied on geolocation collection. Although a few, 11.9% of the DCTAs did not explain how the contact tracing is done, so users would not know what information was used until installing the DCTA. This suggests that tracking geolocation and sharing data with third party providers are the areas in need of most improvement in privacy policy development, as are in lowest compliance with WHO guidelines. A trend in the data showed that many policies did not explicitly state what would happen to user data, with an average of around 14% of the policies not mentioning storage, data transfer, location tracking, and sale of data, and almost half of the policies not mentioning whether the IP address of the individual would be recorded or stored. "IP or UDID removed" shows that not mentioned category was the highest count of 45.2%, to which users would not know whether the individual's identifier was collected by the DCTAs. In some cases where

Table 1 WHO guideline analysis on 42 privacy policies evaluated

Criteria codes	Yes, N (%)	No, N (%)	Not mentioned, N (%)
Data deletion	36 (85.7)	0 (0)	6 (14.3)
Geolocation turned off	28 (66.7)	9 (21.4)	5 (11.9)
Time limitation	35 (83.3)	0 (0)	7 (16.7)
Third parties[†] sharing off	26 (61.9)	9 (21.4)	7 (16.7)
Non-commercialization of data	36 (85.7)	0 (0)	6 (14.3)
IP[‡] or UDID[§] removed	18 (42.9)	5 (11.9)	19 (45.2)

[†], Third parties: anyone outside of a specified government and application developers; [‡], IP: Internet Protocol address; [§], UDID: Unique Device Identifier.

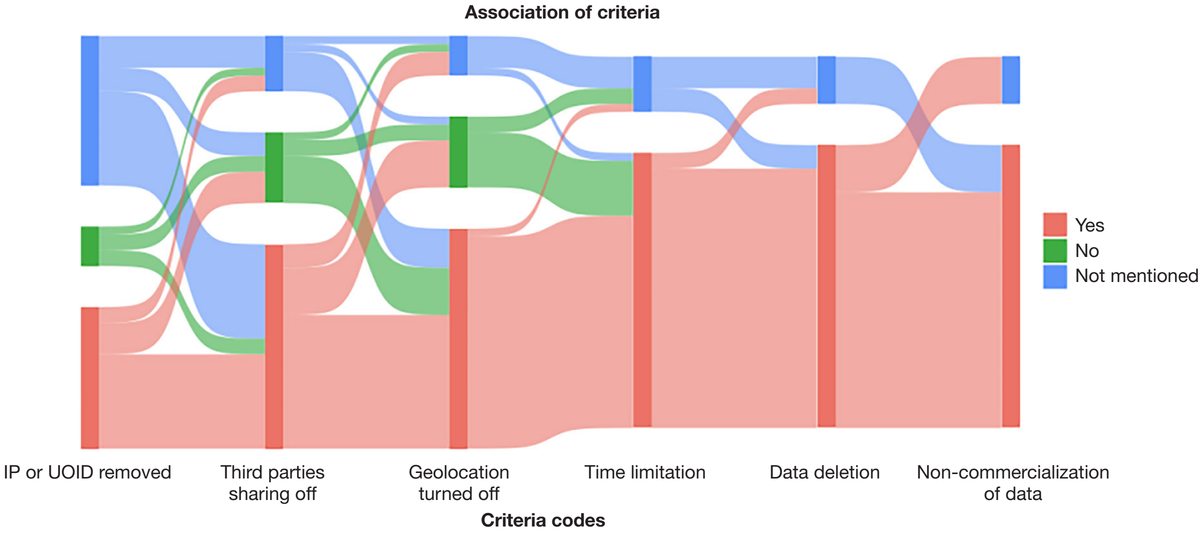

Figure 3 Association of criteria in the order of less respected to more respected criterion. IP, Internet Protocol address; UDID, Unique Device Identifier.

the information needs to be shared with third parties, explicitly stating which information is shared, with whom it is shared, as well as having the option of choosing to share the information or not all need to be available.

Another trend that was observed from the privacy policies was the association between categories. When the criteria were ordered by lower to higher "Yes" percentage and made into a Sankey diagram with the flow of "Yes", "No", and "Not mentioned" answers, *Figure 3* was created. The width of the flows is proportional to the number of DCTAs in this diagram and the direction of the flow changes as the DCTA's "Yes", "No", "Not mentioned" answers change for each criterion.

Privacy policies that explicitly mentioned what were determined as strong privacy protection protocols from the WHO in one category were more likely to meet the criteria for all six categories. DCTAs that respected the criterion that had the biggest split, "IP or UDID removed", were more likely to consistently respect all the other criteria. There is a bottom red portion which consistently meets "Yes" throughout the criteria. Likewise, those policies that did not offer protection of data, especially policies that did not state what would happen to user data ("Not mentioned"), were more likely to fail all six criteria as shown in *Figure 3*. There is more movement of the flows in the top portion of *Figure 3*, which shows the incompleteness of the privacy policies of some DCTAs.

Discussion

The results of this paper suggest that tracking geolocation and sharing data with third party providers are the areas in need of most improvement in privacy policy development. At the time of this crisis, trust in the interventions that policymakers propose is important to implement effectively in a population and adhering to accepted levels of privacy invasion is thus important as well (6). Although none of the apps mentioned that the data collected from apps would be used for commercial purposes on the privacy policy, there were still a few apps that did not mention where the data would go. It is also important to note that what is written on the privacy policy is not necessarily proven to be followed by the function of the app unless the coding of the apps is open to the public on platforms like GitHub. Some of the privacy policies mention that data may be shared with third parties which included the respective governments, Ministry of Health, parties in collaboration for the development of the app, and other healthcare related parties with the purpose to update the number of cases and getting the treatments provided if necessary. However, the privacy policies do not mention the exact procedures of what would happen to a user after being contacted by those parties after being notified as positive or at high risk. Even though users are contributing the data by participating in the DCTAs, if the users are uncertain about what responsibilities are enforced by being in contact with symptomatic users, then it would discourage the users from participating.

As Bluetooth allows for the least privacy invasion, from the ethical standpoint of do not harm, it would be highly recommended to have the DCTAs running on Bluetooth to eliminate the geolocation information and preserve anonymity. However, globally not all mobile phones are Bluetooth compatible, and the "do not harm" principle would be difficult to achieve. It is important to acknowledge that removing all the harms, whether foreseeable or unforeseeable and direct or indirect, of using DCTAs is not possible. Therefore, during an emergency like COVID-19 outbreaks or pandemics, making decisions based on utilitarianism where the benefits outweigh harms is of utmost importance.

What the privacy policies showed is that transparency in the function and usage of DCTAs is one of the fundamental rights that users should have before using the apps. There were a consistent amount of DCTAs that fell into the category of "not mentioned" for each criterion, especially "IP or UDID removed" which shows that users need to agree on the privacy policy that does not cover all the information on how their data would be used. Having access to all the information on what users are agreeing on and understanding the implications of using the DCTAs are crucial to protecting users from exploitation and avoiding having the users in a vulnerable position. This is necessary no matter if DCTAs are required or optional in any country. Who can take the responsibility of monitoring the privacy policies and punishing the violations of not providing appropriate privacy policies is another issue. Platforms that provide or sell applications such as the App Store and Google Play would be a systematic and effective way to intervene and support the users' privacy by requiring the developers to have a complete privacy policy available in multiple languages on global scale, helping with local integration and acceptance of health policies. The definition of a complete privacy policy needs to also be revised regularly as more technologies become available beyond Bluetooth, geolocation, camera, audio, and other data that can be collected on a mobile device.

In addition, to protect the privacy of an individual, consulting the WHO regulations worldwide is recommended, however mandating these regulations is challenging due to differing levels of support of centralization and decentralization. We would be amiss without commenting on the WHO guidelines taking a western-centric perspective implying more developed security structures and less government involvement. We argue that nations should have the ultimate authority over privacy policy regulation, but the WHO guidelines should take de facto importance in the absence of regulations from nation states superseding them. For decentralized contact tracing, depending on the scale of the implementation, alternative options are possible on a case by case basis. For example, on a scale of organizations or institutions, symptom reports without identifying individuals and then being contacted if users need assistance with treatments could be possible. On a state-level or community-level, contact summary aggregation is possible if Bluetooth is available (15). If not regularly monitored but for different occasions such as concerts and other big gathering events, QR codes can be used as a methodology of DCT to inform the audience of any exposures during that high-risk event.

Regardless of what the optimum protection of privacy is and the interest of the government, when considering the implementation of DCTAs to the public, potential ethical implications and analyzing the pros and cons would be highly recommended, so the privacy policy can address all issues and be agreed upon and explicit before using the app.

Conclusions

At minimum, privacy policies should explicitly state whether they will store or record personal data. The ambiguity gleaned from the results in *Figure 3* limit transparency and can create doubt and distrust with the creator of the DCTAs. Learning from the COVID-19 DCTAs is a crucial step for ethical and effective development and implementation of future DCT in pandemics. One future area of research could include examining the accuracy of ethical technology to ensure that it is functioning in the way it claims to be. Another could be analyzing measurable indicators, where information could be collected on download numbers of applications as a proxy for successful encouragement of use. This could also involve more privacy policies from additional countries with the assistance of native speakers to ensure linguistically accurate translations.

DCTAs are multi-faceted and privacy policies are not the only aspect of DCTAs that deserve attention. The aptly named "pingdemic" of the UK suggested that high alert rates reduced engagement with the DCTA which suggests an overt concern with DCTA use as opposed to a covert concern of privacy policies. The sheer number of people told to isolate put manufacturing lines in danger of shutting down and put potential economic crisis in the near future (16). On top of this, DCTAs can cause a false sense of security. If an individual is under the impression that all proximal contacts are using the DCTA, reduced notifications can be seen as positive. However, individuals may not always report symptoms and some close contacts may not be using the apps at all. These are both avenues for future research on efficacy of DCTAs outside of the privacy policy realm. Even with effective privacy policies protecting data, how effective can DCTAs truly be?

This also raises questions of influence and where people get their ideas on safety and efficacy of applications, indicating another future avenue of surveying where people find news on DCTAs and how they evaluate them. This study would focus on the acceptance of emerging DCTA technology.

In the future, it is likely that there will be more big data to make use of in outbreaks—such as the possibility of a vaccine passport. Therefore, not only digital contact tracing apps, but any technologies that may use the individual's health data need to have thorough investigations on the ethical implications and minimize the possibility of identifying individuals with big data, be considered by the technology developers, and regulated by the governments and global policies.

Acknowledgments

We are grateful for advice and comments we received from Prof. Daniel Weissglass and Prof. Meifang Chen. We would like to thank reviewers at this journal for their input on this manuscript, which has resulted in a much stronger paper.
Funding: This work was funded by the Health Humanities Laboratory at Duke Kunshan University, itself an effort of the Humanities Research Center at the same institution.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Mellissa Withers and Mary Schooling) for the series "Global Urban Health: Findings from the 2021 APRU Global Health" published in *Journal of Public Health and Emergency*. The article has undergone external peer review.

Data Sharing Statement: Available at https://jphe.amegroups.com/article/view/10.21037/jphe-22-27/dss

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://jphe.amegroups.com/article/view/10.21037/jphe-22-27/coif). The series "Global Urban Health: Findings from the 2021 APRU Global Health" was commissioned by the editorial office without any funding or sponsorship. This work was supported by the Health Humanities Laboratory from Duke Kunshan University Humanities Research Center for article purchases and conference fees. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Archived: WHO timeline - covid-19 [Internet]. World Health Organization. World Health Organization; 2020 [cited 2022 Mar 2]. Available online: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19
2. Sanyaolu A, Okorie C, Hosein Z, et al. Global Pandemicity of COVID-19: Situation Report as of June 9, 2020. Infect Dis (Auckl) 2021;14:1178633721991260.
3. Mooney G. "A Menace to the Public Health" - Contact Tracing and the Limits of Persuasion. N Engl J Med 2020;383:1806-8.
4. Mobile cellular subscriptions (per 100 people) [Internet]. Data. The World Bank; [cited 2022 Mar 2]. Available online: https://data.worldbank.org/indicator/IT.CEL.SETS.P2
5. Klenk M, Duijf H. Ethics of digital contact tracing and COVID-19: who is (not) free to go? Ethics Inf Technol 2020;23:1-9.
6. Shahroz M, Ahmad F, Younis MS, et al. Covid-19 digital contact tracing applications and techniques: A review post initial deployments. Transportation Engineering 2021;5:2666-91.
7. Budd J, Miller BS, Manning EM, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020;26:1183-92.
8. Akinbi A, Forshaw M, Blinkhorn V. Contact tracing apps for the COVID-19 pandemic: a systematic literature review of challenges and future directions for neo-liberal societies. Health Inf Sci Syst 2021;9:18.
9. Alanoca S, Guetta-Jeanrenaud N, Ferrari I, et al. Digital contact tracing against COVID-19: a governance framework to build trust. International Data Privacy Law 2021;11:3-17.
10. COVIDSafe app [Internet]. Australian Government Department of Health. Australian Government Department of Health; 2021 [cited 2022 Mar 2]. Available online: https://www.health.gov.au/resources/apps-and-tools/covidsafe-app
11. Hale T, Angrist N, Goldszmidt R, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021;5:529-38.
12. Zhang M, Chow A, Smith H. COVID-19 Contact-Tracing Apps: Analysis of the Readability of Privacy Policies. J Med Internet Res 2020;22:e21572.
13. World Health Organization. Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing. Interim guidance. 2020 May 28. [cited 2022 Feb 28]. Available online: https://www.who.int/publications/i/item/WHO-2019-nCoV-Ethics_Contact_tracing_apps-2020.1
14. Pagliari C. The ethics and value of contact tracing apps: International insights and implications for Scotland's COVID-19 response. J Glob Health 2020;10:020103.
15. Wang C, Pujol D, Zhang Y, et al. Poirot: Private Contact Summary Aggregation [Internet]. 34th Conference on Neural Information Processing Systems (NeurIPS 2020). 2020 Dec 6-12; Vancouver, Canada.
16. Rimmer A. Sixty seconds on . . . the pingdemic. BMJ 2021;374:n1822.

doi: 10.21037/jphe-22-27

Cite this article as: Greig EE, Shimomura RG. An evaluation of privacy policies used in digital contact tracing apps for COVID-19. J Public Health Emerg 2022.