A research agenda for digital proximity tracing apps

von Wyl Viktor[ab], Bonhoeffer Sebastian[c], Bugnion Edouard[d], Puhan Milo Alan[a], Salathé Marcel[d], Stadler Theresa[d], Troncoso Carmela[d], Vayena Effy[e], Low Nicola[f]

[a] Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland
[b] Institute for Implementation Science in Health Care, University of Zurich, Switzerland
[c] Institute of Integrative Biology, Eidgenössische Technische Hochschule Zürich (ETHZ), Switzerland
[d] École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[e] Institute of Translational Medicine, Eidgenössische Technische Hochschule Zürich (ETHZ), Switzerland
[f] Institute of Social and Preventive Medicine, University of Bern, Switzerland

Introduction

Digital proximity tracing apps are a health technology designed to speed up the tracing of contacts of people found to be infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Apps have been released in several countries, including Switzerland, as additional measures to combat the SARS-CoV-2 pandemic [1]. Developers started work on the apps early in the pandemic, when the numbers of cases of coronavirus disease 2019 (COVID-19) were increasing quickly and classic, interview-based, contact tracing services became overwhelmed [2]. A mathematical study, first published as a preprint in March 2020, found that, because SARS-CoV-2 can be transmitted before the symptoms are apparent, classic, interview-based contact tracing is too slow to control transmission [3].

Digital proximity tracing is based on practical hardware technologies (e.g., Bluetooth low energy), but has undergone only limited real-life testing [1]. The effects of digital proximity tracing, either alone or as an adjunct to classic contact tracing, have not yet been shown. Effectiveness will depend as much on factors such as public confidence and trust in protection of privacy, which affect uptake and adherence with recommended actions, as on technical performance [4]. Research programmes will need to accompany the release of proximity tracing apps to assess how well they work, whether they produce the desired effects, and how technological, *acceptability* and *usability* aspects can be further optimised. In this article, we summarise how digital proximity tracing and classic contact tracing should work, we discuss methods for their evaluation, and present an agenda for research to assess the benefits, harms and costs of this new technology during an epidemic. To this end, we use the frameworks of health technology assessment [5] and comparative effectiveness studies in clinical research [6]. Key words printed in *italics* are defined in appendix 1.

How should a proximity tracing app work?

Proximity tracing apps are designed to trigger a notification on a smartphone when their user has been exposed to one or more people with confirmed SARS-CoV-2 infection (index cases) who are also using the app. In brief, digital proximity tracing enables estimation of the exposure of an app user to other users using the same app, via low energy Bluetooth signals. The app sends and receives ephemeral (regularly changing), pseudo-random identification numbers (IDs). The SwissCovid app uses decentralised digital privacy-preserving proximity tracing (D3-PT) protocols that store only contact data on the phone, thus protecting the privacy of the user [7, 8]. Signal strength and cumulative duration of all observed contact events using this protocol are stored locally on the phone for 14 days. If app users receive a positive result for active SARS-CoV-2 infection, by reverse transcriptase polymerase chain reaction (RT-PCR) testing, they can upload the ephemeral IDs that were transmitted during the contagious period, e.g., starting 2 days before the date of symptom onset, from their phone onto a central server.

All phones using the proximity tracing app periodically connect to the central server and compare their locally stored list of recorded ephemeral IDs with the list obtained from the server. The smartphone will determine whether the cumulative time and distance derived from identified matches between ephemeral IDs marked as contagious and locally stored, observed IDs, indicate sufficient exposure to an infected person, for example, proximity of <1.5 m for a total of 15 minutes. If this is the case, this person (a contact) will receive an exposure notification through the proximity tracing app, and information on how to proceed, e.g., to call a telephone information line. If deemed necessary, the contact will be advised to go into quarantine and, if infected, they might not have transmitted SARS-CoV-2 onwards. The mechanism of proximity tracing, its connection to classical contact tracing and the epidemiological rationale have been discussed elsewhere [8]. Although exact

Correspondence:
Prof. Viktor von Wyl, PhD,, Digital and Mobile Health Group, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Hirschengraben 84, CH-8001 Zurich, Viktor.vonwyl[at]uzh.ch

procedures will vary across countries, the overall principle is expected to be the same.

How should a proximity tracing app work with classic contact tracing?

Classic, interview-based contact tracing is one part of the test, trace, isolate, quarantine (TTIQ) package of interventions to help control SARS-CoV-2 transmission [9]. Classic contact tracing begins when a person with SARS-CoV-2 infection (an index case) receives the diagnosis. Testing for SARS-CoV-2 is usually prompted by symptoms. Because presymptomatic transmission is possible 1 to 3 days before symptom onset [10], the index case might already have transmitted infection to others when they are tested. It may then take several days before the index case receives the result, has a contact tracing interview, and contacts are notified and begin quarantine. Proximity tracing apps are intended to accelerate the last step of contact notification; the app alerts potentially exposed users as soon as the index case receives his/her positive test result and activates the upload of his/her ephemeral IDs onto the central server. Proximity tracing app users who receive an alert can immediately contact an adviser on a telephone hotline and start quarantine earlier. The app can also extend the reach of classic contact tracing because it sends alerts to people who might not be known to an index case. Furthermore, proximity tracing will still function in a situation when classic contact tracing gets overwhelmed by high numbers of new infections.

Proximity tracing app evaluation should be scaled up immediately

Large-scale assessments of proximity tracing apps in real-world populations are still lacking and will likely have to wait until the apps have been rolled out. This is because the release of lockdown measures in many countries has put pressure on health authori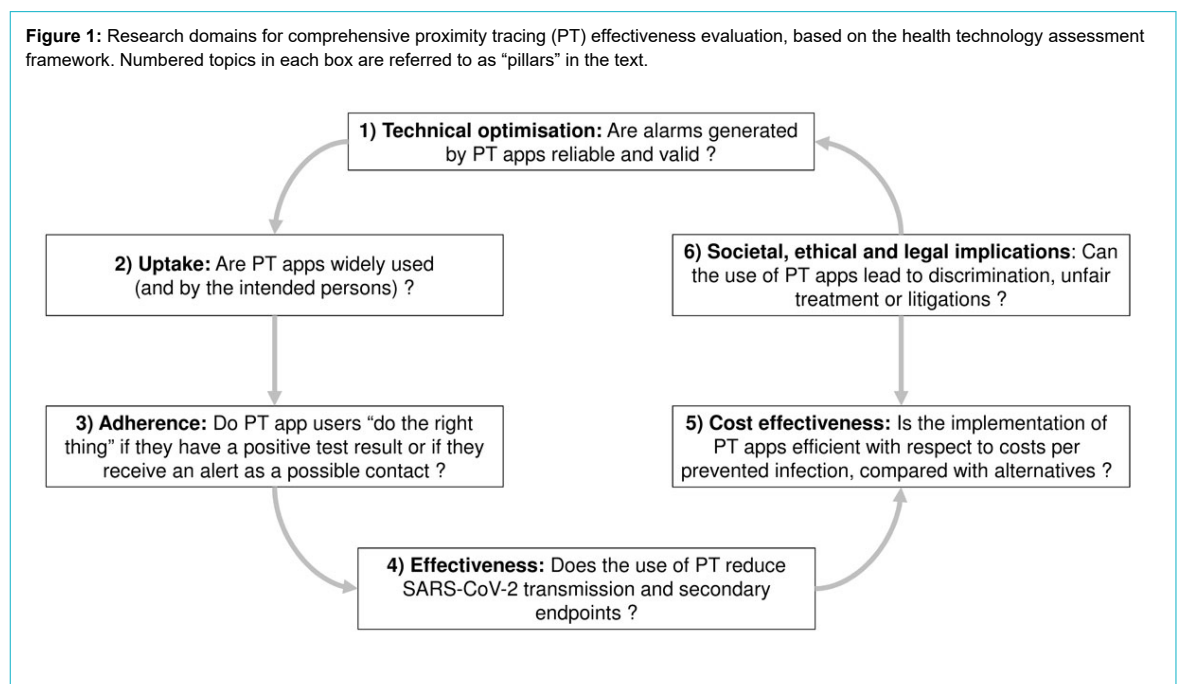ties to deploy apps as quickly as possible to support other pandemic containment measures, with limited time for extensive testing and optimisation.

A research programme can be implemented more easily when transmission is still under control. As of 1 July 2020, the number of new cases of COVID-19 is still relatively low in several countries that have introduced proximity tracing apps, but is increasing. Whereas low, stable numbers of new cases are usually manageable for classic contact tracing, clusters of infection and super-spreading event outbreaks can quickly overwhelm capacity. If classic contact tracing is suspended, however, the collection of important information for proximity tracing app evaluation would be impaired.

Proposal for a pragmatic proximity tracing effectiveness research agenda

Proximity tracing apps are a health technology. The framework of health technology assessment addresses different dimensions and properties of proximity tracing apps, including technical properties, uptake, adherence, safety, *efficacy* or *effectiveness*, economic attributes or impacts and social, legal, ethical and political impacts (fig. 1) [5]. Specific *study questions* for each aspect require appropriate study designs and outcome measures, using a range of quantitative and qualitative health services research methods. Given the differences between the dimensions of health technology assessment, research teams should be interdisciplinary, drawing expertise from diverse fields such as epidemiology, virology, statistics, mathematical modelling, health economics, psychology, and other social sciences.

We propose a research agenda that includes six interdependent pillars, which should result in iterative improvements in proximity tracing apps. For each pillar, we outline the rationale and main research questions and methods (fig. 1 and table 1).

Figure 1: Research domains for comprehensive proximity tracing (PT) effectiveness evaluation, based on the health technology assessment framework. Numbered topics in each box are referred to as "pillars" in the text.

Technical optimisation

Initial studies of proximity tracing apps should focus on technical improvement (fig. 1), such as improving the exposure estimation algorithm based on reliable low energy Bluetooth measurements. Most researchers agree that the definition of close contact according to distance and duration, e.g., <1.5 m for 15 minutes or more, should be scrutinised from technical and epidemiological perspectives. Studies should assess the reliability of exposure measurement as *sensitivity* (the ability of proximity tracing to correctly identify persons with real risk exposures), and *specificity* (the ability to correctly identify persons without risk exposures). The fractions of true positive and false positive risk exposure alarms will be of particular interest. *False positive* notifications denote triggered warnings when there was objectively no risk for transmission of SARS-CoV-2. They can occur for different reasons. For example, false positive notifications can be triggered by measurement inaccuracy and exposure risk misclassifications, that is, the actual distance and duration of contact exposure to index cases were insufficient for transmission, but the app inaccurately identified a potential exposure risk. Other false positive warnings can occur if barriers (such as plexiglass walls) or protective measures (protective gear such as masks worn by index case and contact) effectively prevented transmission, but were not recognisable by the proximity tracing app.

Epidemiological studies are needed to determine the percentage of persons alerted by the proximity tracing app who eventually develop symptoms or test positive for SARS-CoV-2. This will be a biological outcome measure of the "quality" of warnings (in terms of proportions of true positives and false positives) and could provide in-formation for further fine-tuning of exposure definitions. Studies of social and ethical aspects are also needed (pillar 6). From the perspective of preventing SARS-CoV-2 transmission, it might be desirable to err on the side of sending more persons into quarantine, rather than to miss a potential exposure. However, putting a large number of contacts into quarantine is probably unacceptable from a social and economic viewpoint because quarantine is disruptive to personal lives and the economic costs due to income and productivity losses can be substantial. The findings of these studies should feed back into, and inform, an iterative process of technical calibration to find an optimal balance between the different evaluation dimensions (the different pillars in fig. 1).

Uptake

Automated systems record the number of downloads of the app. By 4 July 2020, five days after its release, >1,000,000 people were already actively using the SwissCovid app [11]. In addition to the number of users, as outlined in figure 1, pillar 2, researchers should examine users' social and demographic characteristics in random samples of the general population in repeated cross-sectional surveys. The demographic characteristics of proximity tracing app users could then be compared with the characteristics of people newly diagnosed with COVID-19 in the same time period. If the characteristics of these groups differ, there may be a need to find ways to increase proximity tracing app uptake in certain population groups.

Acceptability, adherence, usage

Research about whether and how people use the proximity tracing app has analogies with the assessment of medica-

Table 1: Pillars of health technology assessment for proximity tracing (PT) apps, research questions, study designs and possible outcome measures.

Pillar	Research question	Study designs	Possible outcomes
1. Technical optimisation	How reliable is the detection of signal strength by Bluetooth low energy algorithm?	Laboratory measurement	Refined Bluetooth signal attenuation
	How accurately does the PT app define a close contact?	Test accuracy study	Sensitivity, specificity % of alerted users who are infected
2. Uptake	What is the uptake of the PT app? How does PT app uptake differ in the whole population, amongst those with smartphones, and amongst those with up to date smartphones?	Downloads Smartphone usage statistics Cross-sectional surveys	% population downloaded % survey respondents reporting use
3. Acceptability, adherence	Why do people download a PT app, or not? Are users adequately informed about the privacy protections and potential risks associated with the app use? How do users and non-users assess issues of trust and privacy? Were there adverse experiences?	Cross-sectional surveys Qualitative studies	Reasons for use or non-use Perspectives and views of app users or non-users
	Do people follow the PT app instructions when alerted?	Contact tracing data App use data Cross-sectional surveys	% notified persons who call information service
4. Efficacy, effectiveness, safety	Does the PT app identify more contacts or identify contacts more quickly than classic contact tracing?	Randomised controlled trial Cohort study Before-and-after study	Number of contacts per index case Time from index case diagnosis to notification or quarantine
	Does the PT app reduce onward transmission of SARS-CoV-2?	Randomised controlled trial Cohort study	Attack rates
	What is the country-level effectiveness of PT apps? What is their contribution to transmission prevention compared with other measures?	Mathematical modelling study	Reduction in incidence Averted cases
5. Economic impact	Is the app cost-effective?	Economic impact studies	Direct and indirect costs of apps Cost per contact identified Cost per prevented case
6. Social, legal and ethical aspects	Does PT app use or non-use have unwanted ethical, societal consequences? Do users have the right to challenge notification alerts? What is the oversight mechanism of the PT app and how transparently does it operate?	Qualitative research methods Legal opinion	Discrimination Stigmatisation Surveillance attitudes Public accountability

tion adherence ("drugs don't work in people who don't take them") [12]. Findings from such studies will help to elucidate the uptake of proximity tracing apps in the *target population* and to assess the strength of public support and trust in the apps. The decentralised protocols of the Swiss-Covid app mean that very limited data are available about usage of the app. Additional studies (fig. 1, pillar 3) will therefore be needed to examine whether app users are following recommendations, that is, whether they call the infoline in the event of an exposure notification or whether they enter positive test results into the app to trigger notifications. A cross-sectional study could collect these data by inviting proximity tracing app users at random to complete an anonymised survey on their motivations for app usage and response to exposure notifications. Qualitative research can explore issues such as reasons why people do not use proximity tracing apps, or concerns about privacy. International, comparative studies could examine attitudes and behaviours in different places that have introduced centralised or decentralised proximity tracing apps.

Efficacy and effectiveness

The overall goal of proximity tracing apps is to break chains of SARS-CoV-2 transmission by leading to swifter identification of close contacts and implementation of testing and quarantine than with classic contact tracing alone. Proximity tracing can also potentially reach persons who are not socially connected (or known by name) to the index case and would therefore be missed by classic contact tracing. Pillar 4 (fig. 1) is at the heart of effectiveness research.

The design of studies and the outcomes measured depend on the stage during an epidemic at which a proximity tracing app is introduced and the availability of data for measuring intermediate process outcomes or biological outcomes of the endpoints of contact tracing (fig. 2).

Should we assess efficacy or effectiveness?

In clinical drug research, the concept of *efficacy* defines the effect of an intervention under optimal circumstances, which might not reflect real-life settings [13]. In practice, proximity tracing apps are not a standalone technology. They are an adjunct to classic contact tracing, and human interactions are needed to help judge the relevance of an alert and to guide and support testing and quarantine. Research should therefore be aimed at the evaluation of effectiveness of proximity tracing apps in the real-life settings in which they will be used [13]. There are different study designs that can be used to assess effectiveness at different stages of the contact tracing pathway.

Randomised controlled trials

Randomisation allows causal inference about the effects of an intervention on an outcome. Because SARS-CoV-2 is transmissible through direct contact and people are mobile, whole communities, towns or regions (clusters) should be randomised [14]. Cluster-randomised trials could compare the effects of proximity tracing apps introduced earlier or later. Outcomes could be measured at the individual level within clusters and at the population level in terms of transmission [15]. The design of cluster-randomised trials presents methodological, logistic, and ethical challenges. Where proximity tracing apps have already been made available in a whole country, it would not be possible to restrict their use to certain areas. Also, the course and geographic distribution of SARS-CoV-2 transmission might not allow for staggered app release or balanced randomisation. Even in countries with severe COVID-19 epidemics, the absolute incidence of SARS-CoV-2 infection is low, so randomised controlled trials would need to be very large.

Observational study designs

Effectiveness can be assessed using non-randomised *observational* study designs, accompanying the introduction of proximity tracing app release, and gathering data while the apps are in use. When access to proximity tracing apps has not been assigned at random, users and non-users of the apps and areas with differing levels of app usage will differ in ways that might be associated with the outcomes of interest (*confounding*) [16]. For example, app users might be more likely than non-users to wear masks, or practice more social distancing. Data about these potential confounding factors should be collected so that they can be controlled for in statistical analysis.

Figure 2: Stages in the effectiveness pathway for proximity tracing apps.
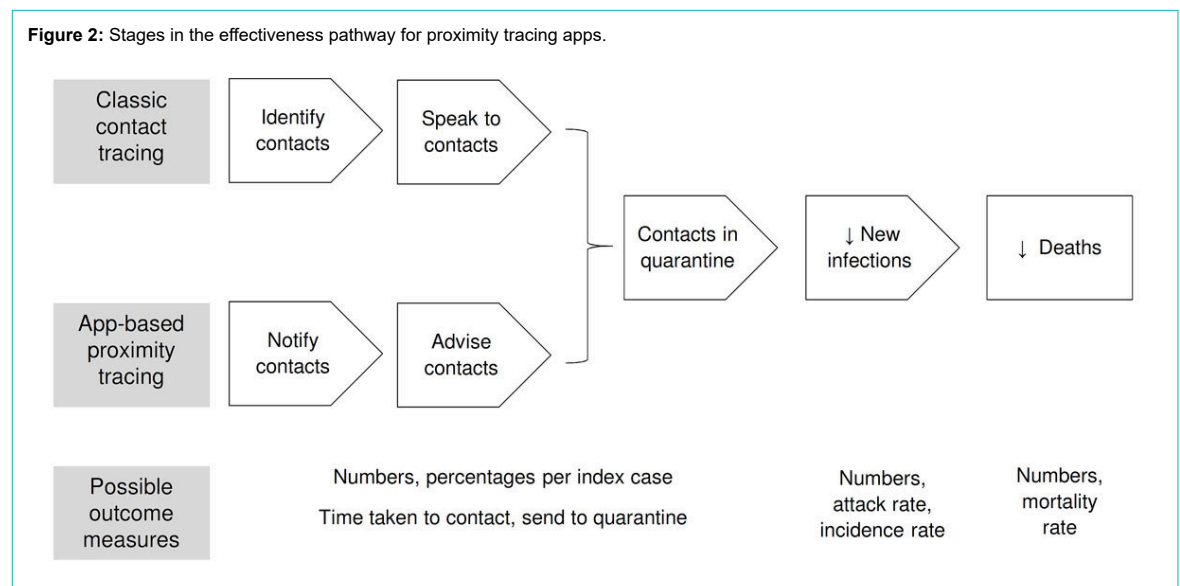
Outcomes

The choice of *study outcomes* will have implications for proximity tracing app effectiveness evaluations. These outcomes should be measurable in a *reliable* (obtaining the same result when repeating measurements) and *valid* manner (to measure what is intended) [16]. However, proximity tracing app evaluations are challenged by the fact that the main event of interest, SARS-CoV-2 transmission, is not directly observable.

There are multiple stages at which the outcomes of proximity tracing app use, and classic contact tracing as a comparator and means for data collection, can be measured (fig. 2). The goal is to contribute to reduced SARS-CoV-2 transmission and deaths from COVID-19 in a population.

The attack rate has been recommended as an outcome measure [17, 18]. The attack rate is the proportion of contacts exposed to a SARS-CoV-2 index case who become infected. To determine the influence of a proximity tracing app on the attack rate requires detailed information for each new index case about the infection status of each identified contact, and whether the contact was identified through classic contact tracing or an app alert. Most contacts who go into quarantine do not get tested for SARS-CoV-2, so large prospective cohort studies would be needed.

Studies to evaluate proximity tracing apps could also assess the intermediate steps in the contact tracing pathway before the biological endpoint of new SARS-CoV-2 infection. The frequency of each intermediate outcome can be assessed prospectively, together with information about exposure to classic contact tracing of the proximity tracing app. If the outcomes are strongly associated with the biological endpoint, they can be used as *surrogate (or proxy) outcomes* [19]. For example, proximity tracing apps could be evaluated by comparing the time-to-notification through classic contact tracing and through proximity tracing apps in persons who are notified by both systems.

Comparator

The use of proximity tracing apps is embedded in the broader TTIQ surveillance and response strategy to contain the epidemic [18, 20]. The relevant question for the effectiveness is not whether these apps alone can stop the SARS-CoV-2 epidemic, but whether their benefits increase the contribution of classic contact tracing, whilst outweighing the harms and being cost-effective and socially acceptable. In clinical research, this type of evaluation is referred to as *comparative effectiveness* studies [6].

There are also study questions and contexts in which proximity tracing apps could be compared with other interventions. For example, comparative studies aiming at deciding between different containment strategy options may warrant comparisons with more invasive measures, such as a mandate to wear masks in public transport systems or prohibition of mass gatherings and certain business activities. Although such evaluations may seem like "comparing apples to pears", they can be informative when, for example, scaling the number of prevented cases by a measure with monetary costs, such as in a cost-benefit assessment.

Mathematical modelling studies

Mathematical models of SARS-CoV-2 transmission can be used to assess the impact of interventions, including proximity tracing apps, at the population level. Mathematical modelling studies have some advantages over empirical studies for the evaluation of interventions for infectious disease control, especially where randomised controlled trials are logistically unfeasible. The outcomes that are measured in epidemiological studies, such as attack rate, give information at the level of a small cluster or community. Empirical studies cannot provide information about what is likely to happen at the population level because they are not long-term enough to capture the dynamic effects of an intervention on transmission [21].

The collection of empirical data about the uptake of proximity tracing apps, the numbers of alerts generated and the outcomes along the contact tracing pathway, including attack rates and infection incidence (fig. 2) is essential to improve the outputs of models developed with theoretical input data [3, 22, 23]. Mathematical models can also be used to examine the impacts of setting-specific features, such as how the TTIQ strategy is implemented in a country, across a range of plausible assumptions and existing data. Proximity tracing evaluation studies should be designed with mathematical modellers, so that information that they need for model parameters is captured accurately in epidemiological studies.

Economic evaluation

Comprehensive assessment of proximity tracing apps as a health technology includes evaluation of cost utility. Economic evaluation can include the direct and indirect financial costs of managing a SARS-CoV-2 epidemic, as well as the harms such as psychological stress, worsening of chronic diseases due to delayed care, breaches of privacy, or negative impacts on children's education due to school closures. Research studies should therefore investigate whether proximity tracing apps are an *efficient* use of resources when compared with other containment measures, such as mandatory mask wearing, efficiency studies should examine whether proximity tracing can achieve reductions of secondary transmissions at lower monetary or societal costs than an alternative measure. In particular, studies should assess whether proximity tracing can identify exposed persons or avert transmissions at lower costs than other pandemic containment measures.

Social, ethical and legal aspects

Ethical issues about the digital proximity tracing apps and their introduction have been discussed in some detail [4, 23–25]. Additional studies should be carried out as part of the health technology assessment, such as those outlined in pillar 3 (adherence).

Challenges for proximity tracing app research agendas

Country-specific factors

Several countries are developing or have released proximity tracing apps. The apps differ in their architecture, using centralised or decentralised collection of data for the identification of exposure risks, technical app specifications, or embedding of apps in the overall surveillance and response strategy. Results in a specific country might not be *generalisable*. This is because effectiveness, when comprehensively defined, is highly dependent on a country's

containment strategy, health system characteristics and incentives provided for individuals to enter quarantine in the case of proximity tracing app notifications. Nevertheless, to achieve greater efficiency, existing registries and ongoing studies should be leveraged to include standardised questions about proximity tracing app usage. Such data could, for example, inform research in pillars 2 (uptake) and 3 (adherence). By contrast, it is possible that more technically oriented optimisation studies do not need to be repeated for each country and setting, as many apps will be using the same application programming interface (API) specifications of Google and Apple [26]. It is therefore important to make such technical studies and data publicly accessible.

Gathering data while preserving privacy

The decentralised privacy-preserving protocols of the SwissCovid DP-3T protocol [7] limit the data that can be used for research. The design of the app follows a privacy-by-design approach and implements important data protection principles such as data minimisation and purpose limitation. Only data that are strictly needed to fulfil the primary purpose of the system, to provide a mechanism to anonymously alert the exposed contacts of an index case, leave the users' device. Proximity tracing apps that provide strong data privacy and security properties are thought to increase trust and better adoption rates by the population, an important factor in the effectiveness of proximity tracing as an intervention [7]. The strong data protection measures also encourage uptake by communities for whom sharing data with central services might be a concern.

To allow for effective evaluation of proximity tracing systems whilst respecting privacy, some countries, such as Switzerland, have regulated the use of data stored on the central server, allowing their use for statistical purposes [27]. In Switzerland, the data available for evaluations is limited to the app user-triggered confirmation of a positive RT-PCR test result, this person's non-identifiable ephemeral IDs, as well as the date of positive test and/or the date of first symptoms [27]. Other data, such as the list of a person's anonymous contacts over the past days are stored locally on app users' phones and are inaccessible. Thus, they cannot be used for analytic purposes.

Systematic evaluations of proximity tracing app effectiveness therefore need to collect data from other sources. Classic contact tracing provides data such as the number of informed contacts of index cases and COVID-19 symptom development among those identified contacts. Routine surveillance provides information about numbers of newly diagnosed cases per day, etc. There are, however, challenges for linkage of data from different sources for technical reasons, such as access restrictions, and legal reasons, such as privacy concerns regarding the use of information from classical contact tracing, or due to anonymisation, that is, lack of individual-level identifiers that would allow database connections. Potential solutions include conducting analyses at geographically aggregated levels, such as numbers of deaths or hospitalisations, to include some variables as exogenous or aggregated, e.g., percentage of app users in a given region.

The need for additional primary data collection

Primary data collection through surveys (or by making use of existing studies) will be essential for some of the research questions of the proposed research agenda. For example, to gain a deeper understanding of app uptake in specific sociodemographic groups, additional user surveys, or representative (randomly sampled) population surveys will be required. Addressing questions on app downloads (pillar 2, uptake) will be relatively straightforward, whereas assessing responses to positive tests or alarms (pillar 3, adherence) will be more challenging. In particular, there is a substantial likelihood that persons may withhold the fact that they ignored alarms or did not enter positive tests (social desirability bias). Furthermore, depending on incidence levels (the daily number of new SARS-CoV-2 cases), the probability of including persons with an active or resolved infection or persons who were warned by the proximity tracing app in a random sample survey (e.g., among app users) will be quite small.

Conclusion

Hope still rests on proximity tracing apps to help contain the SARS-CoV-2 epidemic and to sustain a version of normalised social and economic life after the pandemic lockdown in many countries. The speed of the epidemic and the threat of a second SARS-CoV-2 wave may force health authorities to take decisive actions in case infection numbers rise again substantially. Therefore, it is likely that there is only a limited time window for the optimisation of the proximity tracing app and promotion of substantial population uptake. It will be all the more important that research programmes allow data-driven, evidence-based optimisations, and information for the public about the benefits, harms and costs of proximity tracing apps.

References

1   Zastrow M. Coronavirus contact-tracing apps: can they slow the spread of COVID-19? Nature. 2020. doi: http://dx.doi.org/10.1038/d41586-020-01514-2. PubMed.

2   Thier J. Kann eine App die Corona-Pandemie stoppen? Mit dieser Frage beginnt eine für die Schweizer Wissenschaft aussergewöhnliche Erfolgsgeschichte. NZZ am Sonntag. 6 June 2020.

3   Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science. 2020;368(6491):eabb6936. doi: http://dx.doi.org/10.1126/science.abb6936. PubMed.

4   World Health Organization. Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing: interim guidance, 28 May 2020. Geneva: World Health Organization; 2020.

5   Goodman CS. HTA 101 Introduction to health technology assessment. https://www.nlm.nih.gov/nichsr/hta101/ta10103.html. Bethesda, MD: National Library of Medicine (US); 2014.

6   Armstrong K. Methods in comparative effectiveness research. J Clin Oncol. 2012;30(34):4208–14. doi: http://dx.doi.org/10.1200/JCO.2012.42.2659. PubMed.

7   Decentralized Privacy-Preserving Proximity Tracing. (DP-3T). Decentralized Privacy-Preserving Proximity Tracing - White Paper. Accessed 4 July 2020. Available from: https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf.

8   Swiss National Covid-19 Science Task Force. Digital Proximity Tracing. Policy Briefs. Accessed 4 July 2020. Available from: https://ncs-tf.ch/de/policy-briefs/digital-proximity-tracing-15-may-20-en/download.

9   Swiss National Covid-19 Science Task Force. Contact Tracing Strategy 2020. Accessed 4 July 2020. Available from: https://ncs-tf.ch/de/policy-briefs/contact-tracing-strategy-26-april-20-en-3/download.

10   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 Rep. 2020;69(14):411–5. doi: http://dx.doi.org/10.15585/mmwr.mm6914e1. PubMed.

11   Swiss Federal Office of Statistics. SwissCovid App Monitoring 2020. Accessed 4 July 2020. Available from: https://www.experimental.bfs.admin.ch/expstat/de/home/innovative-methoden/swisscovid-app-monitoring.html.

12   Osterberg L, Blaschke T. Adherence to medication. N Engl J Med. 2005;353(5):487–97. doi: http://dx.doi.org/10.1056/NEJMra050100. PubMed.

13   Singal AG, Higgins PD, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol. 2014;5(1):e45. doi: http://dx.doi.org/10.1038/ctg.2013.13. PubMed.

14   Hayes RJ, Alexander ND, Bennett S, Cousens SN. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. Stat Methods Med Res. 2000;9(2):95–116. doi: http://dx.doi.org/10.1177/096228020000900203. PubMed.

15   Kahn R, Rid A, Smith PG, Eyal N, Lipsitch M. Choices in vaccine trial design in epidemics of emerging infections. PLoS Med. 2018;15(8):e1002632. doi: http://dx.doi.org/10.1371/journal.pmed.1002632. PubMed.

16   Pearce N, Checkoway H, Kriebel D. Bias in occupational epidemiology studies. Occup Environ Med. 2007;64(8):562–8. doi: http://dx.doi.org/10.1136/oem.2006.026690. PubMed.

17   European Centre for Disease Prevention and Control. Mobile applications in support of contact tracing for COVID-19 - A guidance for EU EEA Member States 2020. Accessed 4 July 2020. Available from: https://www.ecdc.europa.eu/en/publications-data/covid-19-mobile-applications-support-contact-tracing#no-link.

18   Servick K. COVID-19 contact tracing apps are coming to a phone near you. How will we know whether they work? Science [Internet]. Accessed 4 July 2020. Available from: https://www.sciencemag.org/news/2020/05/countries-around-world-are-rolling-out-contact-tracing-apps-contain-coronavirus-how.

19   McLeod C, Norman R, Litton E, Saville BR, Webb S, Snelling TL. Choosing primary endpoints for clinical trials of health care interventions. Contemp Clin Trials Commun. 2019;16:100486. doi: http://dx.doi.org/10.1016/j.conctc.2019.100486. PubMed.

20   Salathé M, Althaus CL, Neher R, Stringhini S, Hodcroft E, Fellay J, et al. COVID-19 epidemic in Switzerland: on the importance of testing, contact tracing and isolation. Swiss Med Wkly. 2020;150(11-12):w20225. doi: http://dx.doi.org/10.4414/smw.2020.20225. PubMed.

21   Garnett GP, Cousens S, Hallett TB, Steketee R, Walker N. Mathematical models in the evaluation of health programmes. Lancet. 2011;378(9790):515–25. doi: http://dx.doi.org/10.1016/S0140-6736(10)61505-X. PubMed.

22   Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al.; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health. 2020;8(4):e488–96. doi: http://dx.doi.org/10.1016/S2214-109X(20)30074-7. PubMed.

23   Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. Lancet Digit Health. 2020. doi: http://dx.doi.org/10.1016/S2589-7500(20)30137-0.

24   Büchler A. Contact Tracing als Instrument der Pandemiebekämpfung. Zentrale Gesichtspunkte aus der Perspektive der Ethik (Stellungnahme Nr. 33. Bern, 6. April 2020). Bern: Nationale Ethikkommissionim Bereich der Humanmedizin NEK; 2020.

25   Parker MJ, Fraser C, Abeler-Dörner L, Bonsall D. Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. J Med Ethics. 2020;46(7):427–31. doi: http://dx.doi.org/10.1136/medethics-2020-106314. PubMed.

26   Apple / Google. Privacy-Preserving Contact Tracing 2020. Available from: https://www.apple.com/covid19/contacttracing.

27   Swiss Federal Council. Coronavirus: Verordnung für Proximity-Tracing-App verabschiedet, Unterstützung für Kultursektor verlängert. 2020. Available from: https://www.bag.admin.ch/bag/de/home/das-bag/aktuell/medienmitteilungen.msg-id-79103.html

Appendix 1

Glossary of keywords

Keyword	Explanation Descriptions in quotation marks reflect unaltered quotations from the respective publication.	Reference
Acceptability	Acceptability can be defined as whether stakeholders "find the App likeable, including its interface and navigation features."	4
Comparative effectiveness	"The generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policymakers to make informed decisions that will improve health care at both the individual and population levels."	1
Comparator	"Participants in the control arm of a study. […] In a controlled trial, a participant in the group (receiving placebo, no treatment, an active comparator, or standard of care) that serves as a comparator for the experimental intervention."	1
Confounding	"A variable (or characteristic) more likely to be present in one group of participants than another that is related to the outcome of interest and may potentially confuse (confound) the results."	1
Effectiveness	"The extent to which an intervention works under real-world conditions (i.e., in practice). Effectiveness studies involving drugs examine whether they work when they are used the way that most individuals take them. […] Effectiveness studies ask the question, Does it work?"	1
Efficacy	"The extent to which an intervention produces a beneficial result under ideal conditions (i.e., in clinical trials). Efficacy trials ask the question, Can it work?."	1
Efficiency	"The ratio between health system inputs (costs, in the form of labour, capital, or equipment) and either outputs (e.g. number of patients treated) or health outcomes (e.g. life years gained)."	3
Endpoint	"Medical events occurring as a result of disease or treatment (e.g., stroke, disability, hospitalization)."	1
False-positive classification	"A test result that is positive for a person who does not have the disease." With respect to PT apps: a risk exposure notification when in fact there was no risk.	1
Feasibility	Feasibility can be defined as whether implementation of the App can be "easily and conveniently done, accounting for advantages and disadvantages to integrating the application into routine workflow."	4
Generalizability (external validity)	"The extent to which results provide a correct basis for generalizations to other circumstances (e.g., populations, settings)."	1
Observational study	"A study in which investigators observe the course of events and do not assign participants to the intervention."	1
Primary endpoint	Endpoint addressing the main research question. They guide study planning and sample size calculations.	2
Randomisation	"The process of randomly assigning participants to one of the arms of a controlled trial. Ensures that participants have an equal and independent chance of being in each arm of the study."	1
Reliability	"The extent to which an instrument, scale, or other type of measurement or procedure yields consistent and reproducible results."	1
Secondary endpoint	Endpoints measured in addition to primary endpoints. "Secondary endpoints are generally not sufficient to influence decision-making alone, but may support the claim of efficacy by demonstrating additional effects or by supporting a causal mechanism."	2
Sensitivity of classification	"The proportion of time a diagnostic test is positive in individuals who have the disease or condition." With respect to PT apps: the proportion of exposure notifications when there was indeed a relevant exposure to the virus.	1
Specificity of classification	"The proportion of time a diagnostic test is negative in individuals who do not have the disease or condition." With respect to PT apps: the proportion of contacts without subsequent exposure notification when there was no relevant exposure.	1
Study question (hypothesis)	"A conjectural statement of the relation between two or more variables. A proper hypothesis should be pre-specified, measurable, have theoretical or empirical support, be clearly articulated, and testable by an appropriately designed study."	1
Surrogate (or proxy) outcomes	"Measurements of a patient's physical or biomedical status used as a surrogate for, or to infer the degree of, disease […]. Surrogate endpoints correlate with clinical outcomes but the relationship is not necessarily definitive."	1
Target population	"The population to which the investigator wishes to generalize."	1
Usability	Usability can be defined as whether the App "functioned in a way that enhanced productivity or led to unproductive tasks due to errors."	4
Validity	"The degree to which a result (of a measurement or study) is likely to be true and free of bias (systematic errors)."	1

1   https://amcp.org/sites/default/files/2019-03/AMCP%20CER_glossary%208.2011.pdf2   https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6881606/pdf/main.pdf3
https://pubmed.ncbi.nlm.nih.gov/10213735/4 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5192299/