The barriers and enablers to downloading the COVIDSafe app – a topic modelling analysis

Nicolas R. Smoll,[1,2] Jacina Walker,[1] Gulam Khandaker[1]

The novel coronavirus (COVID-19) has made its way around most of the world with each country having various levels of success controlling its spread. Contact tracing, alongside border restrictions, lockdowns, social distancing and heightened hygiene compliance are all measures that have contributed to Australia's successful control of COVID-19.[1] If the majority of contacts of a confirmed case are traced, then onward community transmission and outbreaks can be controlled. However, the identification of contacts can be difficult because manual contact tracing relies on each person's recall.[2]

Digital contact tracing, or contact tracing augmented by digital means, can improve the speed and completeness of contact tracing.[3] The rapid quarantine of close contacts should lead to the control of an outbreak if the quarantine of contacts is complete and instantaneous.[2,4] Digital contact tracing should augment manual contact tracing, as it may help to classify the unidentified cases (cases with no epidemiological link) that can make up 10–15% of cases overall.[5]

The COVIDSafe app was developed by the Australian Government as a public health initiative with the potential to revolutionise contact tracing by helping to increase the completeness and speed of close contact identification (and subsequent quarantine) to better control the spread of the illness.[3,6,7] As of early June 2020, only 24% of the Australian population had downloaded the app (6 million downloads), which is less than the estimated 56–95% population uptake required for it to be useful.[8,9]

Abstract

Objective: We report a survey in regional Queensland to understand the reasons for suboptimal uptake of the COVIDSafe app.

Methods: A short five-minute electronic survey disseminated to healthcare professionals, mining groups and school communities in the Central Queensland region. Free text responses and their topics were modelled using natural language processing and a latent Dirichlet model.

Results: We received a total of 723 responses; of these, 69% had downloaded the app and 31% had not. The respondents' reasons for not downloading the app were grouped under four topics: lack of perceived risk of COVID-19/lack of perceived need and privacy issues; phone-related issues; tracking and misuse of data; and trust, security and credibility. Among the 472 people who downloaded the app and provided text amenable to text mining, the two topics most commonly listed were: to assist with contact tracing; and to return to normal.

Conclusions: This survey of a regional population found that lack of perceived need, concerns around privacy and technical difficulties were the major barriers to users downloading the application.

Implications for public health: Health promotion campaigns aimed at increasing the uptake of the COVIDSafe app should focus on promoting how the app will assist with contact tracing to help return to 'normal'. Additionally, health promotors should address the app's impacts on privacy, people's lack of perceived need for the app and technical barriers.

Key words: COVIDSafe, contact tracing, machine learning, natural language processing, survey

The app works by the user downloading and registering basic demographic details to which an encrypted reference code is generated for the app on the particular device being used.[6] The app then uses Bluetooth® to look for other devices that have the COVIDSafe app installed and performs a 'digital handshake' when contact occurs. When contact with another person using the app is made (within approximately 1.5 metres), an encrypted reference code of each device and the time and proximity of the contact is then securely encrypted and stored on each person's phone. The app continues to check for new connections every minute or so to determine the total duration of contact with another person, with those noted as having contact for 15 minutes or more deemed as close contacts.[6]

Understanding the reasons for the suboptimal uptake of the COVIDSafe application is what motivates this study. A non-peer-reviewed closed question survey of the Australian population about the COVIDSafe app found that privacy was most commonly ranked as the most applicable reason for not downloading the app (31% of respondents). Other reasons listed as most applicable were a lack of perceived effectiveness (20%), battery usage

(19%), government trust (18%) and other miscellaneous causes.[10] However, to date there has not been any qualitative study or topic modelling analysis performed on this issue.

Therefore, to inform future health promotion campaigns, we performed a survey on the uptake of the application in regional Australia to better understand the barriers and enablers to downloading the COVIDSafe app.

Methods

A short five-minute electronic survey created in Citizen Space that was open between 1 May 2020 and 9 June 2020. A link was disseminated via emails to a number of organisations, targeting local healthcare professionals, mining groups and school communities in Central Queensland, which is a mainly regional area. We encouraged them to send links to others and excluded individuals under the age of 18. Many of those groups work in high-risk settings (high-exposure risks as well as those living/ working in close quarters). The link led to a dedicated survey page, and the responses were downloaded in comma-separated file-type for analysis.

The questions consisted of a series of demographic questions and the keystone question: *Did you download the COVIDSafe app?* The survey logic then presented questions that were related to the keystone question. If the respondent had responded "No" to the keystone question (voluntary response) they were then asked: *Why did you elect not to download the app?* If they had answered "Yes" to the keystone question (voluntary response) they were asked: *Why did you download the COVIDSafe app?*"

We received Human Research and Ethics Committee approval for this project (LNR/2020/QCQ/65752). All analyses were performed using R, version 4.1.

Statistical methods

Most responses were free text and natural language processing was used to establish the key topics listed as reasons for not downloading the application. Firstly, text was pre-processed using tokenization, stemming and the removal of stop words via recognised text mining dictionaries. Bigrams were included in the final model to improve the fit. Lemmatization was not employed and thus a 'bag-of-words' approach was used in modelling the data. The model used to identify the topics was the well-known Latent Dirichlet Allocation (LDA) model employed in the analysis of open-ended surveys. LDAs are used to identify k hidden topics (latent topics) within free text (or large datasets). This recognises that a piece of text (document) may have up to k topics, and each word will have a probability of belonging to each topic. Choosing the k number of latent topics (model selection) was performed using a topic coherence metric, aiming to maximise coherence while maintaining parsimony. We tested 1–20 topics (k). Due to the limited size of the dataset, we limited the topic choices to fewer than 10 to avoid overfitting. Finally, a hierarchical cluster analysis framework to identify themes of topics that resembled each other was used to present the model. These themes should be targeted when educating the public.

Results

We received a total of 723 responses with 93% (n=671) of the respondents being currently employed. The majority (n=588, 81%) of respondents were female. More than two-thirds (n=499, 69%) of respondents had downloaded the app. The respondents were mostly from educational settings (n=201, 28%), and healthcare settings (n=157, 22%), with some mining site representation (n=24, 3%). The top three methods by which respondents had heard about the COVIDSafe app were TV (n=391, 54%), social media (n=156, 22%) and their employer (n=64, 9%).

Of those who did not download the app (n=224, 31%), 203 provided free text data that could be used for analysis (the remaining were unanswered). The terms "app", "phone" and "government" were the most commonly used terms as seen in the word cloud of Figure 1.

The four-topic LDA model had the highest coherence and was selected as the final model. Our topics were: lack of perceived risk of COVID-19/lack of perceived need and privacy issues; phone-related issues; tracking and misuse of data; and trust, security and credibility.

Figure 1: Word cloud for responses to why persons chose not to download the COVIDSafe app.

Note:
The six most common words were "app" (n=45), "phone" (n=34), "government" (n=31), "don't" (n=26), "information" (n=26) and "trust" (n=23).

Figure 2: Word cloud for responses to why persons chose to download the COVIDSafe app.

The topic 'lack of perceived risk of COVID-19/ lack of perceived need and privacy issues' formed 45% of the responses, and presumably is due to many respondents living in rural and remote areas (see Table 1 for examples). The topic 'phone-related issues' represented actual or perceived technical difficulties relating to the app. The topic 'tracking and misuse of data' represented our population's dislike at the idea of being tracked by the government. The terms "tracking" and "government" were often found in these responses. The final topic, 'trust, security and credibility', represented a lack of trust (the word "trust" was often found in these responses) towards the government or third parties.

A similar analysis was performed to model the topics of the 499 survey respondents who elected to download the application (472 provided free text responses amenable to text mining, the remaining did not respond) and responded to the question: *In your own words, why did you elect to download the application?* A total of nine topics were found as shown in Table 2, and topics 2, 3 and 4 were collapsed into one representing 'assisting with contact tracing'. Similarly, topics 8 and 9 were collapsed into 'protect and ensure the safety of family and community', as there were analogous themes when compared with the raw responses. The two topics most commonly listed for those who downloaded the application were 'to assist with contact tracing' and 'to return to normal'. The word cloud shown in Figure 2 demonstrates how the contact tracing was the most important reason why respondents downloaded the COVIDSafe application.

Discussion

In our regional and rural population, a low exposure risk, lack of perceived need for the application and technical issues were the most important reasons people did not download the app. These results are consistent with those of a survey done by Thomas et al. (2020) that demonstrated a refusal-to-download rate of 28% of 1,500 sampled Australians (our refusal-to-download rate was 31%).[11] In their work, technical issues were 24% of the reasons why people did not download the application, which is identical to our findings. Distrust and the app being unnecessary were also reasons in the Thomas et al. survey. Our surveyed population, which was more heavily represented by rural and

Table 1: The four topics generated from the LDA model of responses of persons as to why they did not download the COVIDSafe application.

Topic	Title	Count	Per cent	Examples
1	Lack of perceived need/Low Exposure Risk and Privacy	92	45	"Because I live in a rural area and don't feel it is necessary." "Did not see the need." "No cases seemed unnecessary and not wanting my movements tracked." "Privacy reasons." "Where I live is not a hot spot, there has only been two cases and surrounding regions are quite safe." "I didn't want to or feel the need to as the media always lets the community know (whether true or not) if there has been a case."
2	Technical issues	49	24	"Do not have mobile phone that can download app." "Phone wont download it?" "My phone is too old and I can't afford a new phone." "My phone is old and is unable to download apps." "I have heard that the App doesn't actually work on iPhones yet."
3	Tracking and misuse of data	25	13	"We do not partake of social media or any tracing apps. I have no faith in the government or businesses in deleting the contact or information when it is no longer required." "No one has the right to track another without there consent." "I refuse to give the government permission to track me" "Because I would prefer to keep my private movements private, from this and any future governments, organisations or individuals." "Do not agree with my location being tracked by an app."
4	Trust, security and credibility	37	18	"Do not trust Commonwealth or State governments to secure data or to confine data use to COVID." "Because I could not identify a clear purpose for this app and have a distinct lack of trust about any external party accessing my personal information." "Just havent got around to it. A few concerns about data safety … "

Table 2: The six (collapsed from nine) topics generated from the LDA model of responses of persons as to why they downloaded the COVIDSafe application.

Topic	Theme	Frequency	Per cent	Examples
1	Returning back to normal	125	26	"For the good of the community/country, the sooner things get back up and running the better."
2, 3, 4	Assist with Contact Tracing	160	34	"Help in case of needing to be be contact traced." "To assist authorities to contact people who may have come in contact with COVID or if I was to be sick to let those I have come in contact with know." "To potentially assist with tracing."
5	Reduce risks to vulnerable populations / Health Care Worker	45	10	"100% for my mother! My mother is in the vulnerable population with a respiratory condition. We reside together. I work with a health service where we need to show our faith in the technology and amazing efforts to contain the virus." "Frontline healthcare worker."
6	Stop Spread of COVID-19	50	11	"Stop the spread." "To do my part in helping Australia recover from the pandemic" "To support the National Governments strategy to manage COVID-19."
7	Quickly informed of positive COVID-19 cases	35	7	"So I can be contacted easily or if I get it all of the people I've been in contact with know." "To be notified if someone I had contact with was diagnosed with Covid-19."
8, 9	Protect and ensure the safety of family and community	57	12	"To help protect my immediate family, partner's family and the co-workers whom I work in close proximity with every day." "Either it would be voluntary or it would become mandatory for all to have. I have nothing to hide so i opted to download it. Also it would help protect my family if needed." "To ensure health and safety." "Just so our family can stay safe and everyone else." "To help keep my family & community as safe as possible."

regional people, saw a greater proportion of issues relating to lack of perceived need/lack of perceived exposure (45% of respondents who refused to download the application). Overall, our findings are consistent with the findings of others.[12]

The barriers (and enablers) described here are consistent with the social learning theory and the health belief model.[13] The most important barrier in our regional population was a lack of health concern (motivation) and many believed that they were not susceptible (vulnerable) due to their rural/isolated location. In addition, the technical barriers to downloading and using the application (barriers to change) and perceived benefits are all key antecedents of change.[11,12] Addressing these antecedents of behaviour change are going to be critical for health promoters to drive increased uptake of contact tracing applications.

Strengths of this survey include the use of free text open responses, rather than closed questions that force participants to respond to a set of answers. In addition, we used validated natural language processing methods known as topic modelling allowing for a quantitative approach to a thematic analysis of free text responses.

There are several limitations to be noted. Our regional population may not represent the entire Australian population and we sampled professionals limited to the healthcare, education and mining settings, whereas Thomas et al. (2020) used a third-party survey firm to obtain a sample that was more cross-sectional in nature.[11] Persons living in hotspots or cities that saw large outbreaks such as Melbourne would see a greater need for this application. In-depth interviews and focus groups could identify further information on the decision-making process. However, considering the time constraints, our survey provided some useful information to develop tailored communication materials.

Limitations of the Latent Dirichlet Analysis include the pre-selection of the number of topics prior to modelling, and the possibility that defining the theme can be subjective. Defining the theme occurs when the researchers assign each theme a label, instead of 'theme 1' or 'theme 2'. In addition, it uses a bag-of-words approach (meaning that words and orders of words are interchangeable), and no sentence structure or grammar is incorporated into the model. Lastly, a response that would often be considered a 'miscellaneous' reason would get randomly allocated to topics due to the probabilistic nature of theme assignment.

Conclusion

Health promotion campaigns aimed at increasing the uptake of the COVIDSafe app should focus on key antecedents of behaviour change and should be tailored to the population (rural vs. urban). A key enabler of change was the perception that the app would assist with contact tracing to help return to 'normal'. The barriers to change were a perceived lack of risk of COVID-19 infection, privacy concerns and technical difficulties. Future app designers need to take into consideration that 25% of people refused to download or use the app because of technical reasons. Further research is warranted to examine the differences between rural and urban populations.

References

1. Kupferschmidt K. The lockdowns worked—but what comes next? *Science*. 2020;368(6488):218–19.
2. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Health*. 2020;8(4):e488-e96.
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.
4. Fraser C, Riley S, Anderson RM, Ferguson NM. Factors that make an infectious disease outbreak controllable. *Proc Natl Acad Sci U S A*. 2004;101(16):6146–51.
5. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). *J Epidemiol Community Health*. 2020;74(10):861-6.
6. Australian Department of Health. *COVIDSafe App*. Canberra (AUST): Government of Australia; 2020.
7. Servick K. COVID-19 contact tracing apps are coming to a phone near you. How will we know whether they work? *Science*. 2020;May 21:5:10pm.
8. Sophie Meixner. How many people have downloaded the COVIDSafe app and how central has it been to Australia's coronavirus response? *ABC News* [Internet]. 2020 [cited 2020 Jun 15];Jun 26:6:15am. Available from: https://www.abc.net.au/news/2020-06-02/coronavirus-covid19-covidsafe-app-how-many-downloads-greg-hunt/12295130
9. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: A systematic review to inform the control of COVID-19. *Lancet Digit Health*. 2020;2(11):e607–21.
10. Garrett P, White J, Little D, Perfors A, Kashima Y, Lewandowsky S, et al. *A Representative Sample of Australian Participant's Attitudes Towards the COVIDSafe App*. Newcastle (NSW): University of Newcastle Cognition Lab; 2020.
11. Thomas R, Michaleff ZA, Greenwood H, Abukmail E, Glasziou P. Concerns and misconceptions about the Australian Government's COVIDSafe app: Cross-sectional survey study. *JMIR Public Health Surveill*. 2020;6(4):e23081.
12. Megnin-Viggars O, Carter P, Melendez-Torres GJ, Weston D, Rubin GJ. Facilitators and barriers to engagement with contact tracing during infectious disease outbreaks: A rapid review of the evidence. *PLoS One*. 2020;15(10):e0241473.
13. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the health belief model. *Health Educ Q*. 1988;15(2):175–83.