Voice-based screening for SARS-CoV-2 exposure in cardiovascular clinics

Abhinav Sharma 1,2*, Emily Oulousian 1,2, Jiayi Ni 2, Renato Lopes 3, Matthew Pellon Cheng 4, Julie Label 2, Filipe Henriques 2, Claudia Lighter 1,2, Nadia Giannetti 2, and Robert Avram 5,6

1DREAM-CV Lab, McGill University Health Centre, 1001 Decarie Blvd, Montreal, Quebec H4A 3J1, Canada; 2Division of Cardiology, McGill University, Montreal, Quebec, Canada; 3Duke Clinical Research Institute, Duke University, 300 W Morgan St, Durham, North Carolina 27701, USA; 4Divisions of Infectious Diseases and Medical Microbiology, McGill University Health Centre, 1001 Decarie Blvd, Montreal, Quebec H4A 3J1, Canada; 5Division of Cardiology, University of Ottawa, 40 Ruskin Street Ottawa, Ontario K1Y 4W7 Canada, Canada; and 6Montreal Heart Institute, University of Montreal, Montreal, 5000 Rue Belanger, Montreal, Quebec H1T 1C8, Canada

Received 14 March 2021; revised 3 June 2021; accepted 15 June 2021; online publish-ahead-of-print 16 June 2021

Aims

Artificial intelligence (A.I) driven voice-based assistants may facilitate data capture in clinical care and trials; however, the feasibility and accuracy of using such devices in a healthcare environment are unknown. We explored the feasibility of using the Amazon Alexa (‘Alexa’) A.I. voice-assistant to screen for risk factors or symptoms relating to SARS-CoV-2 exposure in quaternary care cardiovascular clinics.

Methods and results

We enrolled participants to be screened for signs and symptoms of SARS-CoV-2 exposure by a healthcare provider and then subsequently by the Alexa. Our primary outcome was interrater reliability of Alexa to healthcare provider screening using Cohen’s Kappa statistic. Participants rated the Alexa in a post-study survey (scale of 1 to 5 with 5 reflecting strongly agree). This study was approved by the McGill University Health Centre ethics board. We prospectively enrolled 215 participants. The mean age was 46 years [17.7 years standard deviation (SD)], 55% were female, and 31% were French speakers (others were English). In total, 645 screening questions were delivered by Alexa. The Alexa mis-identified one response. The simple and weighted Cohen’s kappa statistic between Alexa and healthcare provider screening was 0.989 [95% confidence interval (CI) 0.982–0.997] and 0.992 (95% CI 0.985–0.999), respectively. The participants gave an overall mean rating of 4.4 (out of 5, 0.9 SD).

Conclusion

Our study demonstrates the feasibility of an A.I. driven multilingual voice-based assistant to collect data in the context of SARS-CoV-2 exposure screening. Future studies integrating such devices in cardiovascular healthcare delivery and clinical trials are warranted.

Registration

https://clinicaltrials.gov/ct2/show/NCT04508972.

* Corresponding author. Tel: +1 514 934 1934 x35414, Fax: +1 514 938 7396, Email: abhinav.sharma@mcgill.ca

© The Author(s) 2021. Published by Oxford University Press on behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Introduction

Artificial intelligence (A.I.) driven voice-based technologies have significant potential to streamline and optimize delivery of healthcare and conduct of clinical trials.1–3 Numerous commercial voice-based devices are on the market and are poised to potentially play a larger role in healthcare environments.2 Consumer data suggested that the use of voice assistant grew by 103% from 21.5 million in 2018 to 43.7 million in 2020.4 However, the current major use case of these technologies is for search engine queries.5 There is a lack of data surrounding the evaluation of such technologies to within a healthcare environment and furthermore, the reliability, accuracy, and concordance with manual or healthcare provider data collection has not been extensively explored. The SARS-CoV2 pandemic has provided an opportunity to test the feasibility of using A.I. driven voice-based devices to collect patient level information. Furthermore, such devices may facilitate optimal healthcare environment. For instance, within a healthcare setting, screening for risk factors or symptoms relating to SARS-CoV2 exposure is often done directly by healthcare personnel, using standardized institution-specific questionnaires. Such methods are inconsistent and inefficient in screening multiple people. Amazon Alexa6 (‘Alexa’) is a voice-based A.I. enabled personal assistant that can activate cloud-based ‘Skills’ through verbal triggers.7 The ability to create skill that comply with privacy regulation [e.g. the Health Insurance Portability and Accountability Act (HIPAA) in the USA] may enable devices to be integrated into clinical environments. There is limited evidence in the evaluation of voice-based technologies in healthcare settings or in the context of clinical trials.1,2,6–8 By testing the utility and accuracy of using the Alexa in screening for symptoms or risk factors relating to SARS-CoV2, we will obtain greater evidence into the ability to integrate such devices into clinical care and future clinical trials. We prospectively evaluated the ability of Alexa to screen for risk factors or symptoms relating to SARS-CoV-2 exposure (VOICE-COVID-19 Study; NCT04508972).9

Methods

The device utilized was the Amazon Echo Show 8. With Amazon Web Services (AWS), an Alexa Skill was developed using the Amazon Alexa Developer Kit.7 Questions relating to risk factors for SARS-CoV-2 exposure were based on Health Canada recommendations (Table 1).10 The content of the skill was initially developed in English and then translated into French. As Alexa can only have one default language we set this in French but entered skill questions in both English and French. This resulted in a skill where the participant could select either French or English. The resulting English voice was spoken with a slight French accent.

Keywords

Voice-based technologies • Amazon • Alexa • SARS-CoV2 • COVID-19
Data capture using the Alexa

An Alexa skill uses a vocal interaction model and application logic to determine the participant’s request (Figure 1). The voice algorithm is based on natural language understanding (NLU) and automatic speech recognition (ASR). The ASR involves the recognition and translation of speech into text. However, as vocal intonations and inflections vary widely, the device needs NLU to first rearrange spoken data into a machine-readable format. The skill built for this study is a pre-determined sequence of screening questions driven by a vocal request: answering yes/no, numbers for age, etc. The questionnaire is activated through an utterance (a specific phrase): ‘Alexa, ___’. The patient’s words are streamed to the Alexa service in the cloud, ASR and NLU will respond through voice recognition and structure the information into a request, and the screening process will begin. Each response is stored as a slot value and will act as a request for the next question.

Recruitment and screening process

The participants were recruited from 1 October to 31 December 2020. Potential participants were recruited by verbally discussing with the individual upon entry into the cardiovascular clinic. No compensation was provided to the participant. Among participants who verbally consented, the screening questions were initially asked by the healthcare professional. Subsequently, the participant would then move towards the Amazon Echo device and initiate the screening process with the device. The first question asked by the Alexa was whether the participant wanted to proceed in French or English. Subsequent questions were asked about the purpose of visit to the clinic (Table 1). Then the Alexa then proceeded to initiate the screening questions using the language the participant selected. The screening questions asked by the healthcare provider and the Alexa were the same: any international travel in past 14 days out of Canada; in close contact with a person with COVID-19 without personal protective equipment; or presence of any symptoms (in the prior 14 days) including cough, fever, runny nose, sore throat, diarrhea (Table 1). All responses regarding the screening questions were binary (yes/no) and if participants provided additional details this would not be recorded.

Furthermore, if the Alexa device could not understand the participants response, the Alexa would indicate that the response was not understood and would ask the participant to re-respond. The participant was only able to respond to Alexa once the question was completely delivered else Alexa would restart the question. People entering the cardiovascular clinic at the McGill University Health Centre were first manually screened by healthcare personnel. The entry location had a small enclave that enable a more private discussing with the healthcare personnel but still had ambient noise from the surrounding areas. If participants responded ‘yes’ to the screening questions either with the healthcare provider or the Alexa, following completion of the post-survey questions, the healthcare provider would then move the participant to an isolation area and conduct a more detailed history.

Post-screening survey

After completion of the screening process, the Alexa Skill was rated by participants on a five-point scale. The scale was based on a previously validated scale of user preference and app engagement. Participants were subsequently asked about any privacy concerns related to data storage or use of the Alexa application by asking the following question with a binary yes/no response: ‘Do you have any privacy concerns regarding the use of the Amazon Alexa to screen for symptoms’.

Data flow, consent, and study ethics approval

The high-level data flow and study architecture for the VOICE-COVID-19 study are described in Figure 1. When the researchers are defining the specific questions and triggered, it gets uploaded onto the AWS cloud platform through the AWS S3 bucket. This will integrate into the AWS Lambda server which will transfer information into the Aws DynamoDB. With DynamoDB researchers can create database tables that can store and retrieve any amount of data and serve any level of request data flow. This study was approved by the ethics board at the McGill University Health Centre. Verbal consent (written consent waiver was granted) was obtained from all adults participants 18 years and older (or from their legal representatives) prior to screening by the healthcare provider.

Statistical analysis

Categorical variables were presented as counts (percentages) and continuous variables were presented as median, 25th and 75th percentiles. The primary endpoint of the current study was the interrater reliability of Alexa vs. manual screening using simple and weighted Cohen’s Kappa statistic. To facilitated pragmatic data collection, data collection on the specific comorbidities of patients were not collected. Data were analysed using SAS version 9.4 software (SAS, Cary, NC, USA). Statistical significance was based on a P-value of <0.05.

Coding availability

Coding for the Amazon Alexa skill will be made available by requests to the corresponding author.

Results

Demographics

In total, 215 participants were screened [mean age of 46.1 ± 17.7 years; 118 (55%) females; Table 2]. The range of age was 16–84 years. There were no participants that were offered the opportunity to participate who refused. There were 66 (31%) French speakers and the remainder were English. There were 114 (53%) patients or family

---

**Table 1 Specific questions asked by the Amazon Alexa**

| Baseline questions asked by the Amazon Alexa: |
|---------------------------------------------|
| (1) Which language do you prefer to continue in? You can select French or English. Dans quelle langue préférez-vous continuer? Vous pouvez sélectionner le français ou l’anglais. |
| (2) Do you work at this hospital? |
| (3) What is the purpose of your visit? You can indicate work, appointment, visit, delivery, pickup or other. |

**Screening questions for exposure to SARS-CoV2**

|   |   |
|---|---|
| (1) Do you work at this hospital? |
| (2) What is the purpose of your visit? You can indicate work, appointment, visit, delivery, pickup, or other. |
| (3) Have you travelled outside of Canada in the last fourteen days? |
| (4) Have you been in close contact with a confirmed case of COVID-19 without protective personal equipment? |
| (5) Do you currently have any of the following symptoms: cough, fever, runny nose, sore throat, and diarrhoea? |

*If the participant works in the hospital, this question is skipped.*
members and the rest were hospital workers. In total, 645 screening questions were delivered by Alexa.

**Screening results**

There were 13 responses missing due to participant withdrawal ($n=2$) and loss of wireless internet connection (due to our hospital network) after starting. In 15 questions (out of 645; 2%), the study coordinator had to clarify how the respond to the Alexa due to the participant interrupting the Alexa delivered questions. In 8 attempts (out of 645; 1.2%), the Alexa did not understand the response and asked the participant to repeat the response. There were 14 positive responses to screening questions. Alexa mis-identified one English response—one screening question was responded to as ‘yes’, while the Alexa capture this as ‘no’. The overall simple Cohen’s kappa statistic reflecting the correlation between Alexa and the healthcare personnel screening was 0.989 [95% confidence interval (CI): 0.982–0.997]. The weighted Cohen’s Kappa correlation was 0.992 (95% CI 0.985–0.999) (Figure 2). In the oldest tertile of age (range 57–84 years), the concordance was perfect with Cohen’s kappa correlation of 1.0. There was perfect concordance in French with a Cohen’s Kappa 1.00 (95% CI 1.00–1.00); for English, the concordance was 0.96 (95% CI 0.886–1.000).

**Post-screening survey**

The median time for completion of the Alexa screening questions was 40 s (interquartile range 33–44 s). The mean (with standard deviation) participant rating (out of five) was as follows: 4.7 (0.7) for appropriately asking the questions; 4.7 (0.7) for ease of language comprehension; 4.5 (1.0) for ease of device use; and 4.4 (0.9) for overall rating (Table 3). No participant expressed any privacy concern.

**Figure 1** Architecture diagram for VOICE-COVID study data flow. AWS, amazon webservice; CSV, comma separated value; NLU, natural language understanding. 1. With natural language understanding (NLU), the data transferred from the Alexa device can deduce through artificial intelligence what a speaker is uttering. 2. The intent and slots will define the utterance into the coded values that have been defined by the researcher. 3. When the utterance data is transferred to the AWS cloud platform, it gets integrated into AWS Lambda which is a serverless compute service that enables the running of the question algorithms.

**Table 2** Baseline characteristics of the participants

| Demographics                               | N (%) |
|--------------------------------------------|-------|
| Number of subjects                         | 215   |
| Preferred language                         |       |
| English                                    | 149 (69.3%) |
| French                                     | 66 (30.7%) |
| Mean age (SD) (Range), years               | 46.1 (17.7) (16–84) |
| Male                                       | 97 (45.12%) |
| Work at this hospital                      | 101 (47) |
| Purpose of visit                           |       |
| Work                                       | 101 (47.0%) |
| Appointment                                | 76 (35.4%) |
| Visit                                      | 27 (13.6%) |
| Delivery                                   | 1 (0.5%) |
| Pick up                                    | 1 (0.4%) |
| Other                                      | 9 (4.2%) |
| Travelled outside of Canada in the last 14 days | 5 (2.3%) |
| Have been in close contact with a confirmed case of COVID-19 without protective personal equipment | 4 (1.9%) |
| Currently have any of the following symptoms: | 6 (2.8%) |
| cough, fever, runny nose, sore throat, diarrhoea |

SD, standard deviation.
Discussion

There has been limited evaluation of the use of voice-based devices in healthcare settings. In the present study, we demonstrated that the Amazon Alexa demonstrated near perfect concordance with screening questions compared to healthcare personnel. Furthermore, it appeared that there was no decrement in the ability to be screened by Alexa among older participants. The ability to use commercial voice-based devices to facilitate SARS-CoV-2 screening may represent a strategy to reduce burden on healthcare personnel and minimize their exposure to potentially infected individuals.2 Furthermore, these results suggest that future studies evaluating the utility of integrating such devices into routine clinical care and clinical trials in non-SARS-CoV-2 related settings are warranted.

There are several aspects to the functionality of voice-based technologies. The hands-free SARS-CoV-2 screening is built on the Alexa Skills Kit’s interactive voice-driven interface.13 The ability to activate the skill without requiring additional step, and just through voice, makes the use of such technologies ideal for populations where data capture via electronic devices is challenging, such as amongst older participants.

As healthcare delivery and future clinical trials become more pragmatic, there will be increasing need to leverage digital technologies to facilitate rapid data acquisition. Our result demonstrate the potential for such devices to collect medical grade data with high degree of accuracy and reliability. Such technologies may have the potential to reduce the burden of healthcare delivery and trial conduct by enabling accurate remote data collection. In addition, leveraging voice-based A.I. to supplement such routine screening can possibly enable more effective workforce utilization. Our results provide important feasibility information to identify that data can be accurately collected through the Alexa. Extending the use of such voice-based
technologies in clinical and non-clinical environments (such as schools, workplaces, or home settings) to evaluate clinical workflow and burden warrants further evaluation.

Our study also highlights the practical elements faced with such devices (Table 4). Voice-based technologies generally need continuous internet access. As we experienced, loss of wireless internet access may impair device functionality. Additional testing into functioning in ‘offline’ modes will be needed to ensure continuous function in clinical settings. Furthermore, while privacy concerns are significant issues, especially around devices and technologies related to identifying those at risk of SARS-CoV2 exposure, in our study, no participant expressed any privacy concern. Such privacy concerns will need to be explored with larger deployment of such voice-based technologies.

Expanding on the results of our study, there is great need in cardiovascular clinical studies and care delivery for accurate remote patient monitoring. Currently, most remote monitoring strategies focus on either telephone or video-based follow-up along with additional wearable devices or patches. Voice-based assistants may provide one potential tool to facilitate remote patient follow-up, especially in settings where directed synchronous interaction with a patient (through a phone call or via video-platform) is limited. While our study does not directly demonstrate the use of Alexa in these use cases, the high degree of concordance of data collected by Alexa to standardized questions, compared to a healthcare provider, suggests that further exploration of such technologies are warranted within cardiovascular studies and care delivery.

**Study limitations**

There are several limitations. Our results cannot be extrapolated to other voice-based devices. In the one case where the Alexa mistook the response, an English response was recognized as French and mis-coded. Future studies will have to evaluate the accuracy of data capture in other languages. Challenges and solutions in using voice-based device are shown in Table 4. Our study was a non-randomized study but provides significant evidence to consider randomized studies of voice-based assistants in similar clinical environments. A larger sample size with more diverse population would enable greater generalizability of our findings. The time required to complete the screening by the healthcare provider was not completed. Further studies comparing duration of interaction with Alexa and healthcare providers are needed. Our entry location still had ambient noise from the surrounding areas which may have impacted on the ability of the Alexa to correctly identify responses; however, these results highlight real-life environments in which such devices can be used. Currently, the participant responds to the questions delivered by the Alexa. Future studies enabling participants to ask questions to the Alexa regarding COVID-19 represents a novel opportunity to explore the ability of such devices to provide medical information.

**Conclusion**

Our results demonstrated the feasibility of a multilingual voice-based assistant to accurately capture medical grade information and facilitate accurate SARS-CoV-2 exposure screening. Such results suggest the feasibility of integrating voice-based technologies in healthcare delivery and future clinical trials. Given the strong concordance of data collected by Alexa and the healthcare provider, evaluating the use of such technologies to facilitate cardiovascular studies and delivery of clinical care is warranted.

**Acknowledgements**

Funding for VOICE-COVID was supplied by the MUHC-Foundation, Canso Investment Counsel Ltd, and in-kind development support was provided by Amazon Web Services (AWS).

**Conflict of interest:** Amazon Web Services (AWS) was the vendor that assisted in making the Alexa Skill. AWS did not influence study design or have access to the primary data nor decision to publish the manuscript. Authors have no disclosures pertaining to this manuscript. A.S. reports receiving support from the Fonds de Recherche Santé Quebec (FRSQ) Junior 1 clinician scholars program, Alberta Innovates Health Solution, European Society of Cardiology young investigator grant, Roche Diagnostics, Boeringer-Ingelheim, Novartis, and Takeda. There are no other relevant disclosures.

**Data availability**

Coding for the Amazon Alexa skill will be made available by requests to the corresponding author. Primary patient data collected for this analysis will not be available.

### Table 4 Challenges and solutions in the evaluation of voice-based devices

| Challenges                                                                 | Solutions                                                                                           |
|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| Accuracy in data collection has not been assessed for devices              | Academic–industry collaborations to validate accuracy of devices for data collection                 |
| Privacy concerns regarding collection of data                             | Limit data collection and adhere to local and national data privacy laws                              |
| Data capture of voice-based devices in noisy environments                  | Field testing of device in real-world settings is critical before deployment                          |
| Multiple language skill development may remain challenging                | Ensure multi-language support is available for the device and enabling validation/training of device algorithms for multiple languages |
| Need for continuous wireless internet access                              | Enable testing of possible ‘offline’ modes of device                                                |

---

et al.
References
1. Sharma A, Harrington RA, McClellan MB, Turakhia MP, Esaph ZJ, Steinhubl S, Mault JR, Mamdani MD, Lothar Roessig, Chandross KJ, Green EM, Patel B, Hamer A, Olgin J, Runsfeld JS, Roe MT, Peterson ED. Using Digital Health Technology to Better Generate Evidence and Deliver Evidence-Based Care. J Am Coll Cardiol. 2018;71(23):2680-2690.
2. Sezgin E, Huang Y, Ramtekkar U, Lin S. Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic. npj Digit Med. 2020;3(1):122.
3. Marquis-Gravel G, Roe MT, Turakhia MP, Boden W, Temple R, Sharma A, Hirshberg B, Slater P, Craft N, Stockbridge N, McDowell B, Waldstreicher J, Boursa A, Bansal S, Wing J, Meunier C, Kassahun H, Coran P, Battaile L, Patrick-Lake B, Hirsch B, Mehta R, Muse ED, Chandross KJ, Silverstein JC, Silcox C, Overhage JM, Califf RM, Peterson ED. Technology-Enabled Clinical Trials. Circulation. 2019;140(17):1426-1436.
4. Voice Assistant Use Through Smart Hearables Consumer Report - Voicebot.ai. https://voicebot.ai/2020/11/19/voice-assistant-use-through-smart-hearables-earbuds-has-doubled-since-2018-apple-airpods-extend-market-share-lead-new-report/ Accessed April 27, 2021.
5. McCaffrey M, Hayes P, Hobbs M, Wagner J. Prepare for the voice revolution. PwC Consum Intell Ser. 2018:1-12. https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/pwc-voice-assistants.pdf.
6. Alexa Healthcare Skills. https://developer.amazon.com/en-US/blogs/alexa/alexa-skills-kit/2019/04/introducing-new-alexa-healthcare-skills. Accessed September 21, 2020.
7. Alexa Skills Kit Official Site. https://developer.amazon.com/en-US/alexa/alexa-skills-kit/frequently-asked-questions. Accessed September 21, 2020.
8. Whitelaw S, Mamas MA, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. Lancet Digit Health. 2020;2(8):e435-e440.
9. The VOICE-COVID-19 - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT04508972. Accessed September 21, 2020.
10. CANADA COVID-19 Self-Assessment Tool. https://ca.thrive.health/covid19/en. Accessed May 3, 2020.
11. What Is Automatic Speech Recognition! - Alexa Skills Kit Official Site. https://developer.amazon.com/en-US/alexa/alexa-skills-kit/2019/04/introducing-new-alexa-healthcare-skills. Accessed March 8, 2021.
12. Levine DM, Co Z, Newmark LP, Groisser AR, Jay Holmgren A, Haas JS, Bates DW. Design and testing of a mobile health application rating tool. npj Digit Med. 2020;3(1):1-7.
13. About Voice Interaction Models | Alexa Skills Kit. https://developer.amazon.com/en-US/docs/alexa/ask-overviews/voice-interaction-models.html. Accessed March 8, 2021.
14. Bengio Y, Janda R, Yu YW, Ippolito D, Jarvie M, Plat D, Struck B, Krastev S, Sharma A. The need for privacy with public digital contact tracing during the COVID-19 pandemic. Lancet Digit Heal. 2020; doi:10.1016/S2589-7500(20)30133-3.
15. Sana F, Isselbacher EM, Singh JP, Heist EK, Pathik B, Armoundas AA. Wearable Devices for Ambulatory Cardiac Monitoring: JACC State-of-the-Art Review. J Am Coll Cardiol. 2020;75(13):1582-1592.
16. Bhatia A, Maddox TM. Remote Patient Monitoring in Heart Failure: Factors for Clinical Efficacy. Int J Hear Fail. 2021;3(1):31.
17. Bayoumy K, Gaber M, Elshafeey A, Mhaimeed O, Dineen EH, Marvel FA, Martin SS, Muse ED, Turakhia MP, Tarakji KG, Elshazly MB. Smart wearable devices in cardiovascular care: where we are and how to move forward. Nat Rev Cardiol. 2021. doi:10.1038/s41569-021-00522-7.
18. Yuan N, Pevnick JM, Botting PG, Elad Y, Miller SJ, Cheng S, Ebinger JE. Patient Use and Clinical Practice Patterns of Remote Cardiology Clinic Visits in the Era of COVID-19. JAMA Netw Open. 2021;4(4):e214157.