Accepted Manuscript online: 2023-09-18 Article published online: 2024-03-13

Vanderbilt Electronic Health Record Voice Assistant Supports Clinicians

Yaa A. Kumah-Crystal[1] Christoph U. Lehmann[2] Dan Albert[1] Tim Coffman[1] Hala Alaw[1] Sydney Roth[1] Alexandra Manoni[1] Peter Shave[1] Kevin B. Johnson[3]

[1] Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, United States
[2] Clinical Informatics Center, University of Texas Southwestern Medical Center, Dallas, Texas, United States
[3] Department of Biomedical Informatics, University of Pennsylvania, Richards, Philadelphia, Pennsylvania, United States

Address for correspondence Yaa A. Kumah-Crystal, MD, MPH, Department of Biomedical Informatics, Vanderbilt University Medical Center, 3401 West End Ave, Ste 600, Office 630, Nashville, TN 37203, United States (e-mail: yaa.kumah@vanderbilt.edu).

Appl Clin Inform 2024;15:199–203.

Abstract

Background Electronic health records (EHRs) present navigation challenges due to time-consuming searches across segmented data. Voice assistants can improve clinical workflows by allowing natural language queries and contextually aware navigation of the EHR.

Objectives To develop a voice-mediated EHR assistant and interview providers to inform its future refinement.

Methods The Vanderbilt EHR Voice Assistant (VEVA) was developed as a responsive web application and designed to accept voice inputs and execute the appropriate EHR commands. Fourteen providers from Vanderbilt Medical Center were recruited to participate in interactions with VEVA and to share their experience with the technology. The purpose was to evaluate VEVA's overall usability, gather qualitative feedback, and detail suggestions for enhancing its performance.

Results VEVA's mean system usability scale score was 81 based on the 14 providers' evaluations, which was above the standard 50th percentile score of 68. For all five summaries evaluated (overview summary, A1C results, blood pressure, weight, and health maintenance), most providers offered a positive review of VEVA. Several providers suggested modifications to make the technology more useful in their practice, ranging from summarizing current medications to changing VEVA's speech rate. Eight of the providers (64%) reported they would be willing to use VEVA in its current form.

Conclusion Our EHR voice assistant technology was deemed usable by most providers. With further improvements, voice assistant tools such as VEVA have the potential to improve workflows and serve as a useful adjunct tool in health care.

Keywords
- EHRs and systems
- human–computer interaction
- clinical decision support
- natural language processing
- clinical data management

Background and Significance

Electronic health records (EHRs) are essential to patient care and serve as a data repository and communication tool. EHRs usually display data by type, presenting similar data like medications, notes, or laboratory results together. This data segmentation forces providers with clinical questions to perform extensive, time-consuming searches to gather the required data elements.[1]

received
March 14, 2023
accepted after revision
September 16, 2023
accepted manuscript online
September 1, 2023

© 2024. Thieme. All rights reserved.
Georg Thieme Verlag KG,
Rüdigerstraße 14,
70469 Stuttgart, Germany

DOI https://doi.org/10.1055/a-2177-4420.
ISSN 1869-0327.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Voice assistants have been used for a variety of tasks[2] including medication adherence,[3] data collection,[4] and companionship in adults.[5] While many clinicians use voice technology for nonclinical purposes, a minority also uses it in the clinical domain.[6] A review of existing voice assistant systems revealed limited development in the context of EHRs, specifically designed to address the unique needs and challenges of health care providers.[2] We developed a voice-mediated EHR assistant known as the Vanderbilt EHR Voice Assistant (VEVA) that summarizes EHR information and allows contextually aware ordering in preparation for a clinical encounter. VEVA's voice interface allows searching and summarizing health record data, which may improve workflows and reduce provider burnout.[7,8] We evaluated the usability and acceptability of VEVA by practicing physicians in guided interactions in an outpatient setting.

Vanderbilt Electronic Health Records Voice Assistant Overview

VEVA accepts voice inputs (e.g., "VEVA, give me the last A1C for Mr. Smith") as imperative or interrogative queries, translates voice into text, and then uses the natural language processing (NLP) engine to map the text to executable EHR commands. VEVA executes user queries via Fast Health Interoperability Resources (FHIR) and other application programming interfaces. VEVA's business logic synthesizes relevant results and returns them to the user as voice replies, text, and/or figures as appropriate (Demo ►**Appendix A Video 1**).

We engineered VEVA as a responsive web application for mobile devices. The user interface is comprised of a JavaScript application using the Angular framework[9] and is integrated with Vanderbilt's EHR using Substitutable Medical Applications and Reusable Technologies on FHIR resources,[10] which provide user authentication and patient context to VEVA. RESTful services[11] built in Java provide business logic and store data to an Oracle database. The third-party Nuance Florence software[12] serves as the automatic speech recognition NLP engine leveraging Nuance's subject matter expertise in medical speech recognition.[13] The technology choices for VEVA were based on a comprehensive review of the current best practices and emerging trends in the domain (VEVA Schematic ►**Appendix B**).[2,14]

Objective

The effect of EHR vice assistants like VEVA on the searching effectiveness and efficiency is not known. However, as an important first step, it is essential to understand VEVA's usability and acceptability among providers. For this study, we describe a usability and acceptability evaluation of VEVA by physicians practicing in an outpatient setting.

Methods

We recruited physicians and nurse practitioners from the Vanderbilt Pediatric Endocrine Division. Outpatient pediatric endocrinology was selected as the area of focus for this study due to the specialized nature of pediatric diabetes management, which lent itself to effective query and summarization of discrete clinical findings with the voice assistant in preparation for a clinical encounter. Providers were notified about the study and its aims during a presentation at the Weekly Pediatric Endocrine Lecture Series. Providers were consented through an institutional review board-approved process.

Providers engaged in six guided interactions with VEVA: medical summary, A1C, weight, blood pressure, health maintenance, and laboratory alert. Users phrased queries in their own words. After interacting with VEVA, users provided feedback focusing on its usability.

Provider interviews regarding their VEVA interactions were audio-recorded and coded by at least two of the seven-member research team. Our qualitative analysis followed a systematic approach, including code generation, thematic analysis, and intercoder adjudication. Identified themes were compared and discrepancies resolved through consensus discussions.

Following the interview, each provider completed a System Usability Scale (SUS) Assessment[15] and rated VEVA's effectiveness using a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5).

Results

Fourteen providers (mean age: 39, range: 29–65), including 10 physicians and four nurse practitioners, participated in the VEVA usability assessment. The 14 providers deemed the VEVA prototype highly usable (mean SUS score 81, scores of greater than 68 are considered above average across other information systems such as EHRs). The highest rated SUS item was "I thought the system was easy to use," with an average score of 4.5 [0.52 SD]. The lowest rated item on the SUS was "I found the various functions in this system were well integrated," with an average score of 3.79 [0.70 SD]. Nine out of the 14 providers (64%) indicated willingness to use the VEVA prototype in its current form assuming continued improvement iterations of the platform. Qualitative results from VEVA system primary interactions are shown in ►**Table 1**.

Discussion

We developed a novel web-based voice assistant for EHR interaction capable of receiving verbal commands, collating requested information, and presenting it to the user, thus eliminating the user's need to search for disparate EHR data. VEVA's SUS scores were better than benchmark scores across other EHR information systems, suggesting that providers perceived VEVA as usable. Most providers agreed to use VEVA in its current state in their clinical practice, whereas others suggested simple improvements.

To date, voice is used in health care predominately in one of three domains: "Voice for (1) documentation, (2) commands, and (3) interactive response and navigation for patients."[2] Speech recognition for EHR documentation is associated with significantly lower SUS scores, most likely as a function of the effort required to correct transcription mistakes, which was the main reason to abandon speech documentation for 70% of users in a 2010 study.[16] We used

 © 2024. Thieme. All rights reserved.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Table 1 Key qualitative themes identified from provider Vanderbilt Electronic Health Record Voice Assistant interviews

Summary item	Structure of summary item			User perception of usefulness: number of users			User perception of length: number of users		
	Number words	Duration (s)	Latency (s)	Positive sentiment	Neutral sentiment	Negative sentiment	Too long	Appropriate	Too short
Overview	102	37	4	11	3	0	2	10	2
	Suggested modifications *(per number of users)*: exclude glucose calculation derived from A1C (9); add medication and routes of administration (8); Add option for bullets (7); add method of insulin administration, availability of continuous glucose monitoring school absence (6); exclude language preference (5); exclude next scheduled visit date (3); exclude duration of diabetes (2); add major medical problems (2)								
A1C	40	15	0.4	10	4	0	1	10	3
	Suggested modifications *(per number of users)*: omit previous A1C comparison (7); add incorporation of normative reference ranges with the graphical visualization of the A1C (2); add colorization to improve the visualization (2)								
Weight	27	9	0.4	13	0	1	3	11	0
	Suggested modifications *(per number of users)*: add weight summary as growth chart (5); add weight percentiles (5); add Inclusion of height and growth information (4)								
Blood pressure (BP)	40	21	0.4	10	3	1	8	6	0
	Suggested modifications *(per number of users)*: add BP percentiles (8); add interpretation of BP change as concerning/not concerning (5); add bars on graph to indicate normal values (2); tabular format of data (2)								
Health maintenance	50–95	21–41	0.4	13	1	0	3	11	0
	Suggested modifications *(per number of users)*: add interactive ordering component (3); add normal laboratories (3); add overdue laboratories (2)								
Decision support (TSH alert)	8	3	—	13	1	0	–	–	–
	Suggested modifications *(per number of users)*: alert tone prior to decision support (2)								

Abbreviation: TSH = Thyroid Stimulating Hormone

VEVA for a new domain—summarization—which avoided the semantic complexity of documentation and led to higher acceptance as indicated by the SUS. Our work suggests that voice could be exploited to address the challenge of "foraging for EHR information."[17]

VEVA's translation of text to speech was occasionally inexact and occasionally experienced intermittent latency. For example, VEVA would pronounce the Roman numeral I in "Type I Diabetes" as the letter "i" instead of the number "one" resulting in the expression "Type i diabetes" instead of "Type one diabetes." This highlights the need for expanded prosodic and pronunciation training tailored to medical vocabulary and terminology. While text is typically read silently, voice assistant tools that now speak aloud clinical content must account for these context-specific vocalization considerations, which is a newer paradigm. The finding that VEVA mispronounced or misunderstood some requests while still being rated very usable reveals a discrepancy in the user experience. It suggests the concept of a flexibility threshold, where users may tolerate some degree of error if the technology otherwise proves useful at addressing other workflow needs. Further research on user expectations and that threshold of usability would provide valuable insights given the high threshold of accuracy expected for EHR interactions. Users suggested enhancing verbal laboratory test ordering processes, which could offer a more intuitive alternative to the traditional multistep methods with keyboard and mouse. Further, while voice commands are an integral feature of VEVA, additional integration of visual aids like tables and graphs could provide a more comprehensive data interpretation platform. The utility of VEVA's summaries was acknowledged, with suggestions to integrate it automatically into clinical notes. Feedback varied on the length of the prose of information delivery, with some providers favoring detailed explanations, whereas others preferred brevity. User feedback underscored the potential for user-centric refinements that could accommodate diverse user preferences and allow users more control over their experience.

Limitations

Our study has several limitations, including its single-site focus and its specialty-specific intervention, which might affect its replicability across broader domains. Environments with high EHR utilization and information needs are most poised to benefit, whereas complex workflows reliant on paper or data not readily available in the EHR may present adoption barriers. The use of VEVA was explored in a smaller sample, which was suitable for qualitative analysis but limited quantitative approaches. Although this study focused on individual provider use in preparation for clinic

 © 2024. Thieme. All rights reserved.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

encounters, exploring VEVA interactions with patients merits future research.

Conclusion

We developed a voice assistant tool for our pediatric endocrinology clinic with a SUS score in a highly usable range. Our prototype elicited noteworthy requests for improvements and additional features that enhanced our understanding of the expectations surrounding human–computer interactions with EHR voice assistant tools. With expanded usability testing, we can determine if VEVA integrates successfully into routine outpatient clinical workflows and potentially assess the future opportunities for incorporation with patient portal systems. The advent of advanced large language models, which were not available at the time of VEVA's initial design, now present new compelling opportunities for augmenting conversational agent capabilities. Applying this technology thoughtfully to VEVA iterations could open new possibilities for voice assistant architecture and improved natural language interactions. Overall, our findings highlight promising directions both for refining VEVA locally and advancing EHR voice assistants more broadly.

Clinical Relevance Statement

Voice-mediated EHR voice assistants like VEVA that search and summarize health record data have the potential to improve workflows and serve as a useful tool in health care encounters as well as reduce provider burnout.

Multiple-Choice Questions

1. Which of the following is true about the VEVA usability assessment?
   a. The VEVA prototype was deemed highly usable, with a mean SUS score of 60
   b. All 14 providers who participated in the VEVA usability assessment were MDs
   c. Providers unanimously agreed that VEVA's voice component sounded natural and high quality
   d. A total of 64% of providers indicated that they would be willing to use the VEVA prototype in its current form assuming continued improvement iterations of the platform

   Correct Answer: The correct answer is option d. According to the text, 9 out of the 14 providers (64%) indicated that they would be willing to use the VEVA prototype in its current form assuming continued improvement iterations of the platform. Therefore, option d is correct. Option a is incorrect because the mean SUS score was 81, not 60. Option b is incorrect because 3 of the providers were nurse practitioners, not all of them were MDs. Option c is incorrect because some providers commented that the VEVA voice component sounded "unnatural" and "stunted" in quality.

2. What is the primary purpose of the VEVA?
   a. To replace traditional medical consultations for patients
   b. To provide companionship to health care providers
   c. To assist health care providers with searching and summarizing EHR data
   d. To conduct automated laboratory tests for patients

   Correct Answer: The correct answer is option c. The primary purpose of the VEVA is to assist health care providers with searching and summarizing EHR data. It aims to improve workflows and reduce provider burnout by providing a voice interface for querying and summarizing health record information.

Protection of Human and Animal Subjects

This study was reviewed and approved by the Vanderbilt University Medical Center Institutional Review Board.

Funding

Research reported in this publication was supported by the Evelyn Selby Stead Fund for Innovation; The Vanderbilt Institute for Clinical and Translational Research; and The National Center for Advancing Translational Sciences of the National Institutes of Health [award number: UL1 TR003163]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of Interest

None declared.

References

1 Lasko TA, Owens DA, Fabbri D, Wanderer JP, Genkins JZ, Novak LL. User-centered clinical display design issues for inpatient providers. Appl Clin Inform 2020;11(05):700–709
2 Kumah-Crystal YA, Pirtle CJ, Whyte HM, Goode ES, Anders SH, Lehmann CU. Electronic health record interactions through voice: a review. Appl Clin Inform 2018;9(03):541–552
3 Luengo-Polo J, Conde-Caballero D, Rivero-Jiménez B, Ballesteros-Yáñez I, Castillo-Sarmiento CA, Mariano-Juárez L. Rationale and methods of evaluation for ACHO, a new virtual assistant to improve therapeutic adherence in rural elderly populations: a user-driven living lab. Int J Environ Res Public Health 2021;18(15):7904
4 Sezgin E, Noritz G, Lin S, Huang Y. Feasibility of a voice-enabled medical diary app (SpeakHealth) for caregivers of children with special health care needs and health care providers: mixed methods study. JMIR Form Res 2021;5(05):e25503
5 Jones VK, Hanus M, Yan C, Shade MY, Blaskewicz Boron J, Maschieri Bicudo R. Reducing loneliness among aging adults: the roles of personal voice assistants and anthropomorphic interactions. Front Public Health 2021;9:750736
6 Wilder JL, Nadar D, Gujral N, et al. Pediatrician attitudes toward digital voice assistant technology use in clinical practice. Appl Clin Inform 2019;10(02):286–294
7 Frintner MP, Kaelber DC, Kirkendall ES, Lourie EM, Somberg CA, Lehmann CU. The effect of electronic health record burden on pediatricians' work-life balance and career satisfaction. Appl Clin Inform 2021;12(03):697–707
8 Kissel AM, Maddox K, Francis JKR, et al. Effects of the electronic health record on job satisfaction of academic pediatric faculty. Int J Med Inform 2022;168:104881

 © 2024. Thieme. All rights reserved.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

9 Angular. Explore angular resources. Secondary explore angular resources n.d. Accessed August 30, 2023 https://angular.io/resources?category=development
10 Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016;23(05):899–908
11 Fielding RT. Architectural styles and the design of network-based software architectures. 2000
12 Nuance Communications Inc. Medical speech recognition solutions. Secondary Medical speech recognition solutions n.d. Accessed August 30,2023 at https://www.nuance.com/healthcare/provider-solutions/speech-recognition.html
13 Corbisiero E, Costagliola G, Rosa MD, Fuccella V, Piscitelli A, Tabari P. Speech recognition in healthcare: a comparison of different speech recognition input interactions. International KES Conference on Innovation in Medicine and Healthcare, 2023:142–52
14 Duda SN, Kennedy N, Conway D, et al. HL7 FHIR-based tools and initiatives to support clinical research: a scoping review. J Am Med Inform Assoc 2022;29(09):1642–1653
15 Brooke J. SUS - a quick and dirty usability scale. Usability Evaluation in Industry, 1996:4–7
16 Hoyt R, Yoshihashi A. Lessons learned from implementation of voice recognition for documentation in the military electronic health record system. Perspect Health Inf Manag 2010;7 (Winter):1e
17 Pirolli P, Card S. Information foraging. Psychol Rev 1999;106(04): 643

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

 © 2024. Thieme. All rights reserved.