

Lessons Learned from Implementation of Voice Recognition for Documentation in the Military Electronic Health Record System

Save to myBoK

by Robert Hoyt, MD, FACP, and Ann Yoshihashi, MD, FACE

Abstract

This study evaluated the implementation of voice recognition (VR) for documenting outpatient encounters in the electronic health record (EHR) system at a military hospital and its 12 outlying clinics. Seventy-five clinicians volunteered to use VR, and 64 (85 percent) responded to an online questionnaire post implementation to identify variables related to VR continuance or discontinuance. The variables investigated were user characteristics, training experience, logistics, and VR utility. Forty-four respondents (69 percent) continued to use VR and overall felt that the software was accurate, was faster than typing, improved note quality, and permitted closing a patient encounter the same day. The discontinuation rate of 31 percent was related to location at an outlying clinic and perceptions of inadequacy of training, decreased productivity due to VR inaccuracies, and no improvement in note quality. Lessons learned can impact future deployment of VR in other military and civilian healthcare facilities.

Key words: voice recognition, speech recognition, electronic health records, electronic medical records, AHLTA, EHR

Introduction

Voice recognition (VR) and electronic health records (EHRs) have both entered mainstream medicine in the past decade. Currently the increased time burden of data entry into EHRs is one of the reasons that the EHR adoption rate is low. With voice recognition software continuing to improve in speed and accuracy, it could potentially improve the process of inputting data into electronic health records and thereby decrease one of the key barriers to EHR adoption. Similar to the introduction of many new technologies, VR may succeed or fail based on personal experience, training, or technical or logistical reasons. We sought to explore the factors that influence the continuation or discontinuation of voice recognition as an inputting method for an electronic health record by surveying all clinicians who volunteered to receive the software.

Background

Voice recognition is a relatively new means to enter patient data. Clinicians who tried VR in the early 1990s used “discrete” voice recognition that required the user to pause after each word. Continuous voice recognition became available around 1998 and rapidly became the industry standard. In the same time frame, specific medical vocabularies were created that greatly improved accuracy.^{1,2} The earliest adopters of VR were often radiologists and pathologists because they depended on dictation services that were associated with high costs and delays in report completion. With traditional dictation, the clinician dictates a note or report, which a transcriptionist transcribes and returns for proofreading and approval. Reports are incorporated into EHRs by either the clinician or the transcriptionist. This process usually takes several days to complete, so it is not ideal when rapid access to a record is needed. Timely completion and closure of an encounter improves the coding, billing, and payment process. Clinicians and healthcare organizations are looking for solutions to rapidly and cost-effectively generate a legible record. Early adopters have embraced VR as a potential solution and have developed templates that format a report into standardized sections and macros that insert a body of standard text into a report, such as “insert normal chest x-ray” or “insert normal gross appendix.”^{3,4} With dictation costs of approximately 6 to 20 cents per line resulting in annual costs of \$5,000 to \$15,000 per clinician, organizations are looking for less expensive alternatives.⁵

Many early studies on voice recognition have limited current applicability because of significant improvements in VR software and computer hardware. Earlier versions of VR were associated with slower speed and accuracy.⁶ Frequently, studies did not report if medical vocabularies, which improve the accuracy of VR, were used.^{7,8} Until recently, computers had inadequate processor speed and random-access memory (RAM) for optimal VR performance. The manufacturer of the VR software Dragon NaturallySpeaking (version 9) recommends that it be installed on computers with at least 1 GHz of processor speed and 1 GB of RAM.⁹ Previous VR studies have included only small numbers of clinician users, thus limiting meaningful statistical analyses. In many studies, training methods were not described and varied considerably between different implementations. Cost analyses of voice recognition conflicted, with return on investment ranging from six months to six years.¹⁰ The cost of clinician time to edit mistakes associated with VR is considerable and frequently not reported.¹¹ Studies of VR use with EHRs focused solely on inputting specialty and ancillary reports and not on typical outpatient encounters.^{12,13}

The Department of Defense (DOD) uses a system known as AHLTA as the EHR for the 9.1 million beneficiaries receiving care at its military healthcare facilities.¹⁴ Clinicians can enter data with 1) MEDCIN, a point-of-care medical terminology database, 2) dictations “cut-and-pasted” into the EHR, 3) free-text typing directly into the EHR, or 4) point-and-click condition-specific automated input methodology (AIM) templates.¹⁵ While MEDCIN provides clinical elements that are codified, clinicians have found the clinical notes slow to create and cumbersome to read. Studies have shown that clinicians prefer to create a natural narrative that can only be achieved with handwriting, dictation, or voice recognition.^{16,17} Clinicians are also reluctant to use data entry methods that reduce productivity. With its continuous improvement in speed and accuracy, voice recognition has the potential to streamline data entry. We are unaware of any published studies that evaluated data entry methods, including voice recognition, into AHLTA. Given the unanswered questions in the medical literature, we studied the implementation and use of VR at a medium-sized military treatment facility and its outlying clinics. We examined the factors associated with continuation or discontinuation of voice recognition software used to input patient data into an electronic health record.

Methods and Materials

Naval Hospital Pensacola (NHP) delivers inpatient and outpatient care to active-duty personnel, military retirees, and their families at the hospital and outpatient care to active-duty personnel at its 12 branch clinics. The NHP medical staff consists of 149 military and civilian clinicians (physicians, physician assistants, and nurse practitioners). Prior to 2008, the majority of clinical notes were handwritten with the exception of the orthopedic and internal medicine clinics, where the majority were dictated. In 2008 all outpatient clinical notes were required to be entered into AHLTA.

In early 2008 NHP offered speech recognition software, Dragon NaturallySpeaking Medical 9, to the entire medical staff on a voluntary basis in order to decrease transcription costs and to potentially improve entry of clinical notes into AHLTA. The medicine-specific package included 14 preconfigured medical specialty vocabularies and a headset microphone. Seventy-five clinicians volunteered to use the software with no penalty if they decided to discontinue use. We did not study those clinicians who did not volunteer to have the software installed. Software was installed on desktop computers with 3.4 GHz processor speed and 1 GB RAM in the clinicians’ offices and/or exam rooms. While the participants received headset microphones, they had the option to purchase handheld noise-reduction microphones. The deployment of the software and training were staggered over approximately a 12-month period. Individual VR “user profiles” were stored on a server so the voice profiles could be used on multiple computers.

Training was offered by a vendor trainer, a NHP information technology (IT) trainer, a physician champion, a clinical peer, software tutorial (self-training), or a combination of these methods. The training method was not randomized and was selected largely by availability of the trainers and/or the comfort level of the clinician. The vendor provided “train the trainer” sessions for clinicians and nonclinicians with above-average technology aptitude (“superusers”). A physician champion spearheaded the VR effort for six months before his transfer to another facility. Clinicians were told training must be completed before the software would be installed on their computers.

All clinicians who received the VR software were asked to complete a voluntary, Web-based questionnaire at least three months post implementation. The assessment consisted of 24 questions about VR user characteristics, training, logistics, and utility. The Web-based anonymous survey was developed and responses collected with the online survey tool SurveyMonkey. The assessment questions were developed and pilot tested by a team of “superuser” clinicians on the medical staff.

The research protocol was approved by the Naval Medical Center Portsmouth Institutional Review Board.

Statistical Analysis: The statistical analysis was performed with GraphPad InStat 3.10 software (San Diego, CA). We analyzed nominal data in the questionnaire in a contingency table and if cells contained an expected (not observed) value of 5 or less, then an assumption for chi-square testing was violated and categories were collapsed. As a result of collapsing categories, results were analyzed in 2-by-2 contingency tables using the Fisher exact test. If the data were ordinal, values were assigned dummy codes (0–4) and group differences were analyzed with the nonparametric Mann-Whitney test.^{18,19} Two-tailed *p*-values were used, and if the *p*-value was less than .05, it was considered statistically significant. Filters and cross-tabulation tools in the survey software were used to analyze variables related to continuation and discontinuation of VR.

The results of the post-implementation assessment was reported in percentages by category and rounded to whole numbers, so totals could be less than or greater than 100 percent.

Results

The survey was completed by 64 clinicians for a return rate of 85 percent. The following are their responses, divided into sections based on the questionnaire.

User Characteristics: Most participants were military clinicians (78 percent) located at the hospital (75 percent). Fifty-nine percent of respondents were primary care clinicians, and 41 percent were non-primary-care clinicians. Only 14 percent of participants had prior experience with voice recognition. Sixty seven percent rated their comfort level with technology in the novice to moderately comfortable range, whereas 33 percent considered themselves very comfortable to expert (Table 1).

Table 1

User Characteristics for Continuing and Discontinuing Voice Recognition

User Characteristics	Total % (n)	VR Continued	VR Discontinued	Statistical Analysis*		
				RR	CI	<i>p</i>
Volunteered to use VR Returned survey	50% (75) 85% (64)	69%	31%			
Clinician Status Physician Non-physician	83% (53) 17% (11)	72% 55%	28% 45%	1.31	.74–2.3	.29
Specialty Primary care Non-primary care	59% (38) 41% (26)	71% 65%	29% 35%	1.08	.77–1.5	.78
Military Status Military clinician Civilian clinician	78% (50) 22% (14)	74% 50%	26% 50%	1.4	.85–2.5	.11
Technology Comfort Level Novice to moderately comfortable Very comfortable to expert	67% (43) 33% (21)	70% 66%	30% 34%	1.04	.72–1.5	.78
Location Hospital Branch clinic	75% (48) 25% (16)	81% 31%	19% 69%	2.6	1.2–5.4	.0004
Training Received From trainer(s) (1–4 types)** None or only software tutorial	61% (39) 39% (25)	77% 56%	23% 44%	1.3	.93–2.0	.1
Training Experience None or inadequate Adequate	17% (11) 83% (53)	27% 77%	73% 23%	.35	.13–.93	.002

* Fisher exact tests (RR = relative risk, CI = confidence intervals)

** Trainer types: vendor trainer, hospital information technology trainer, clinical peer, physician champion

Training: A majority of participants (92 percent) received at least one type of training, with the software tutorial being the most common method (53 percent). Five participants reported receiving no training, four of whom accounted for 20 percent of

the discontinuers. Clinicians in branch clinics received less face-to-face training than clinicians at the hospital (44 percent vs. 67 percent).

Logistics: Participants used a headset (56 percent), handheld microphone (36 percent), or both (8 percent). A majority of participants used VR in the office (74 percent) and less often in one (5 percent) or multiple exam rooms (8 percent) and in multiple clinics (7 percent). Seventeen percent, all discontinuers, did not indicate where VR was used. Ninety-eight percent of users did not use it while the patient was still in the exam room. A majority of continuers (73 percent) used voice recognition immediately after seeing a patient, while only a few of the discontinuers did (12 percent). For continuers, voice recognition was the most common method to input outpatient encounters into AHLTA; 80 percent used it more than 75 percent of the time. For discontinuers, typing was the most common method, with 82 percent using it more than 50 percent of the time. Most respondents never used dictation (67 percent), MEDCIN templates (72 percent), or AIM templates (41 percent) for outpatient encounters. Continuers did infrequently use VR for medical boards, operative notes, discharge summaries, e-mail, and Microsoft Word documents, while all but one discontinuer never used it for these purposes.

VR Utility: Mann-Whitney testing revealed statistically significant group differences (p -values from .0001 to .0007) between continuers' and discontinuers' perceived utility of VR software. A majority of the continuers felt VR was very to extremely helpful (74 percent), saved 11 to more than 60 minutes per day (93 percent), improved EHR notes (93 percent), and resulted in same-day closing of the encounter more than 75 percent of the time (63 percent). In contrast, the majority of the discontinuers felt VR was only slightly to moderately helpful (70 percent), saved less than 10 minutes per day or increased documentation time (100 percent), did not improve the EHR notes (100 percent), and did not result in same-day closure of encounters (59 percent). Compared with typing, VR was rated as faster by continuers (88 percent) and slower by discontinuers (65 percent). The majority of the continuers (72 percent) rated VR accuracy at 85 to 95 percent, while a majority of discontinuers (76 percent) rated it at 80 percent or less. Macros consisting of voice commands that insert text were used by 91 percent of the continuers with 72 percent rating macros very to extremely helpful. In contrast, 41 percent of discontinuers had used macros with 17 percent rating macros very to extremely helpful.

VR Discontinuation: Twenty of the 64 clinicians (31 percent) stopped using VR software. For the clinicians in the internal medicine and orthopedic clinics who routinely dictated their notes prior to VR use, the discontinuation rate was 21 percent (4 of 19 clinicians). User characteristics significantly related to discontinuation were location at an outlying clinic and inadequate or no training. Factors not related to quitting were clinician status, military/civilian status, technology comfort level, and prior VR experience. Compared to clinicians that continued VR, discontinuers generally rated it much lower in helpfulness, accuracy, minutes saved per day, improvement in the quality of EHR notes, and the ability to close the encounter in one day. The main reasons cited for quitting were slowness of the method due, in part, to the time required to correct VR errors (70 percent), failure to recognize the user's voice (35 percent), inadequate training (30 percent), and failure to live up to expectations (30 percent).

Discussion

President Bush established the goal of universal adoption of interoperable EHRs by 2014.²⁰ Title XIII of the American Recovery and Reinvestment Act of 2009 established Medicare and Medicaid reimbursement for practitioners and hospitals who demonstrate "meaningful use" of certified electronic health records, beginning in 2011 with the goal of increasing the adoption rate.²¹ Healthcare organizations and clinicians will need to evaluate the various methods of inputting data into EHRs for maximal productivity and satisfaction.

After a successful pilot program, the Army Medical Command announced in 2009 that it would purchase 10,000 copies of voice recognition software and distribute them to 42 healthcare facilities worldwide for clinical documentation into AHLTA.²² Lessons learned from our study will have relevancy in their large-scale deployment and training strategies.

Our retrospective study assessed the user characteristics, training, and perceived impact of voice recognition on productivity. A majority of those who responded to the questionnaire were young, active-duty military physicians working at the hospital. Most were new to voice recognition but very comfortable with technology. In spite of the high comfort level with technology, 31 percent discontinued use of the software. Discontinuation was associated with location at an outlying clinic; low perceptions of training, performance, and time saved; and perception of a lack of improvement in patient note quality.

Inadequate training was perceived to be the reason for quitting by 30 percent of participants. Although it was originally stipulated that staff would not receive the software without formal training, some had only software-tutorial training available because of delays in training. Other studies of voice recognition noted training times as short as 30 minutes and as long as four hours.^{23,24} In the latter study, in spite of extensive training, only 25 percent of physicians persisted, so there is debate regarding the importance of extensive initial training. In our study the 80 percent discontinuation rate of personnel who received no training would support the need for required training. The authors are unaware of studies published in the medical literature regarding the importance of additional follow-up training for individuals who struggle with and/or quit using VR software.

Respondents used voice recognition primarily in their offices, even though they were given the option of having the software installed in the office or the exam room. With the majority of continuers and a minority of discontinuers using VR immediately after seeing a patient, it is unclear if user preference or exam room logistics (i.e., changing exam room assignments) accounted for this difference. Mobile laptops with wireless headsets and wireless connectivity were not available but potentially could have improved efficiency and acceptance of VR use.

VR accuracy was rated at less than 90 percent by 56 percent of all participants and at less than 80 percent by 49 percent of discontinuers, far lower than the “up to 99 percent” accuracy quoted by the vendor. Time wasted correcting VR notes or potentially missing an important error will continue to be a major barrier to acceptance of VR over traditional dictation. According to an American Health Information Management Association practice brief, the time required to edit is about twice that needed to dictate.²⁵ The literature confirms that clinicians consider the self-editing of voice recognition to be a burden.^{26,27} Clinicians who continued using voice recognition felt there was clear-cut benefit in terms of productivity and improved documentation into the electronic health record (93 percent). The high use of macros (voice commands that insert text) by the continuers (91 percent) could also contribute to their perceptions of higher productivity and accuracy.

The strengths of our study are a high questionnaire return rate, a large medical group studied with representation by primary and specialty care clinicians, inclusion of a hospital and associated clinics and the use of recent VR software with preconfigured medical vocabularies. The limitations of our study that may reduce its accuracy and generalizability are the small numbers analyzed in the contingency tables mandating the collapsing of categories, possible selection bias by using volunteers for VR, the fact that the selection of VR users and training methods were not randomized or standardized and the completion of questionnaires at variable times post implementation. Additionally, amplifying information about the type of VR errors encountered and the time required to correct them would have been helpful. Furthermore, the 50 percent of clinicians who did not volunteer to install VR were not studied to better understand the reasons for non-adoption. This adoption rate of new technology is probably not unexpected. According to Rogers’s innovation adoption curve, approximately 50 percent of individuals fall into the “late majority” and “laggards” categories, for which technology adoption is a challenge.²⁸

Lessons Learned: Based on our experience, we recommend the following to improve VR utility and acceptance:

- Identify and train sufficient numbers of clinical champions and technical-assistance staff to quickly address implementation issues, assist in development of macros and templates, and reassess performance success.
- Assess user characteristics (e.g., current inputting methods, experience with dictation, and technology expertise level) in the pre-implementation phase to tailor the training plan to the user’s needs.
- Require all clinicians new to VR software to complete the software tutorial and receive face-to-face training that at minimum validates the user’s proficiency to establish an accurate VR user profile (matching user’s speech to specific text). Additionally, the ability to use templates and macros at the completion of training and within the first two weeks of implementation should be demonstrated.
- Analyze failures early and consider alternate microphones/headsets, additional individual training, and transcriptionists to edit VR errors.

Conclusion

Our study reported the experiences of clinicians who continued and discontinued the use of voice recognition software at a medium-sized military facility and its outlying clinics. Continued use of VR was associated with location at the hospital, a positive training experience and a positive perception of how VR improved note quality and clinician productivity. Almost one-third of voice recognition users discontinued using the software, primarily because they felt the training was inadequate and their perception of the utility of VR was much lower, compared to those who continued using the software. While voice recognition holds great promise for timely documentation into an electronic health record, training and implementation are

challenging. The variables affecting success must be planned for prior to purchase, training, and implementation and must be followed by frequent reassessment. Future studies are needed to further delineate the key user characteristics, training methods, and logistical considerations that improve VR adoption and continuation rates. In addition, future research is needed to determine how to encourage adoption of new technologies such as voice recognition for the late majority and laggards who tend to resist innovation.

Robert Hoyt, MD, FACP, is the co-director of medical informatics at the School of Allied Health and Life Sciences at the University of West Florida in Pensacola, FL.

Ann Yoshihashi, MD, FACE, is a medical analyst at the Naval Operational Medicine Institute in Pensacola, FL.

Acknowledgement

We would like to give special thanks to Matthew Rings MD and Benjamin Rodriguez MD for reviewing and participating in the questionnaire.

Notes

1. Spikol, L. "Voice Recognition Software: A Tool for Encounter Notes." *Family Practice Management* 6, no. 2 (1999): 55–56.
2. Hoyt, Robert, Melanie Sutton, and Ann K. Yoshihashi. *Medical Informatics: Practical Guide for the Healthcare Professional*. 2nd ed. N.p.: Lulu, 2008, p. 286.
3. Henricks, W. H., K. Roumina, B. E. Skilton, D. J. Ozan, and G. R. Goss. "The Utility and Cost Effectiveness of Voice Recognition Technology in Surgical Pathology." *Modern Pathology* 15, no. 5 (2002): 565–71.
4. Siström, C. L., J. C. Honeyman, A. Mancusco, and R. G. Quisling. "Managing Predefined Templates and Macros for a Departmental Speech Recognition System Using Common Software." *Journal of Digital Imaging* 14, no. 3 (2001): 131–41.
5. Lawler, F. H., D. C. Scheid, and N. J. Vivani. "The Cost of Medical Dictation Transcription at an Academic Family Practice Center." *Archives of Family Medicine* 7 (1998): 269–72.
6. Zafar, A., J. M. Overhage, and C. J. McDonald. "Continuous Speech Recognition for Clinicians." *Journal of the American Medical Informatics Association* 6, no. 3 (1999): 195–204.
7. Zick, R. G., and J. Olsen. "Voice Recognition Software Versus a Traditional Transcription Service for Physician Charting in the ED." *American Journal of Emergency Medicine* 9, no. 4 (2001): 295–98.
8. Ilgner, J., P. Duwel, and M. Westhofen. "Free-Text Data Entry by Speech Recognition Software and Its Impact on Clinical Routine." *ENT-Ear, Nose and Throat Journal* 85, no. 8 (2006): 523–27.
9. Nuance. *Dragon NaturallySpeaking*. Available at <http://www.nuance.com>.
10. Green, H. D. "Adding User-Friendliness and Ease of Implementation to Continuous Speech Recognition Technology with Speech Macros: Case Studies." *Journal of Healthcare Information Management* 18, no. 4 (2004): 40–48.
11. Pezzullo, J. A., G. A. Tung, J. M. Rogg, L. M. Davis, J. M. Brody, and W. W. Mayo-Smith. "Voice Recognition Dictation: Radiologist as Transcriptionist." *Journal of Digital Imaging* 21, no. 4 (2008): 384–89.
12. Issenman, R. M., and I. H. Jaffer. "Use of Voice Recognition Software in an Outpatient Pediatric Specialty Practice." *Pediatrics* 114, no. 3 (2004): e290–93.
13. "Integrating Voice Recognition and EMR Cuts Transcription Time, Costs." *Performance Improvement Advisor* 9, no. 11 (2005): 130–32.
14. AHLTA. Available at <http://www.ha.osd.mil/AHLTA/>.
15. Medcomp Systems. Available at <http://www.medcomp.com>.
16. Walsh, S. H. "The Clinician's Perspective on Electronic Health Records and How They Can Affect Patient Care." *British Medical Journal* 328 (2004): 1184–87.
17. Gilbert, J. A. "Physician Data Entry: Providing Options Is Essential." *Health Data Management* 6, no. 9 (1998): 84–86, 90–92.
18. Motulsky, Henry J. *Intuitive Biostatistics*. New York: Oxford University Press, 1995.
19. *Does Your Data Violate Contingency Table Analysis Assumptions?* Available at http://www.quality-control-plan.com/StatGuide/conting_anal_ass_viol.htm.

20. Jones, K. C. "Obama Wants E-Health Records in Five Years." *Information Week*, January 12, 2009. Available at <http://www.informationweek.com/news/healthcare/showArticle.jhtml?articleID=212800199>.
21. U.S. Government Printing Office. American Recovery and Reinvestment Act of 2009 (Public Law 111-5). Available at <http://www.gpo.gov/fdsys/pkg/PLAW-111publ5/content-detail.html>.
22. Roberts, C. "Patient Records: Speak It, See It, File It with New LRMC Voice Recognition System." Landstuhl Regional Medical Center News Release. January 6, 2009. Available at <http://www.army.mil/-news/2009/01/09/15667-patient-records-speak-it-see-it-file-it-with-new-lrmc-voice-recognition-system/>.
23. Zafar, A., J. M. Overhage, and C. J. McDonald. "Continuous Speech Recognition for Clinicians."
24. Issenman, R. M., and I. H. Jaffer. "Use of Voice Recognition Software in an Outpatient Pediatric Specialty Practice."
25. AHIMA. *Speech Recognition in the Electronic Health Record*. AHIMA Practice Brief. October 2003.
26. Pezzullo, J. A., G. A. Tung, J. M. Rogg, L. M. Davis, J. M. Brody, and W. W. Mayo-Smith. "Voice Recognition Dictation: Radiologist as Transcriptionist."
27. Ramaswamy, M. R., G. Chaljub, O. Esch, et al. "Continuous Speech Recognition in MR Imaging Reporting: Advantages, Disadvantages and Impact." *American Journal of Radiology* 174 (2000): 617–22.
28. Rogers, Everett M. *Diffusion of Innovations*. 5th ed. New York: Simon and Schuster, 2003.

Article citation:

Hoyt, Robert; Yoshihashi, Ann. "Lessons Learned from Implementation of Voice Recognition for Documentation in the Military Electronic Health Record System." *Perspectives in Health Information Management* (Winter 2010).

Driving the Power of Knowledge

Copyright 2022 by The American Health Information Management Association. All Rights Reserved.