# Electronic Health Record Interactions through Voice: A Review

Yaa A. Kumah-Crystal[1]   Claude J. Pirtle[1]   Harrison M. Whyte[2]   Edward S. Goode[2]   Shilo H. Anders[1,3]
Christoph U. Lehmann[1]

[1] Department of Biomedical Informatics, Vanderbilt University Medical Center, Vanderbilt University, Nashville, Tennessee, United States
[2] Department of Computer Science, Vanderbilt University College of Arts and Science, Vanderbilt University, Nashville, Tennessee, United States
[3] Department of Anesthesiology, Vanderbilt University Medical Center, Vanderbilt University, Nashville, Tennessee, United States

Address for correspondence Yaa A. Kumah-Crystal, MD, MPH, Department of Biomedical Informatics, Vanderbilt University Medical Center, 3401 West End Avenue, Suite 630, Office 647, Nashville, TN 37203, United States (e-mail: yaa.kumah@vanderbilt.edu).

## Abstract

**Background**   Usability problems in the electronic health record (EHR) lead to workflow inefficiencies when navigating charts and entering or retrieving data using standard keyboard and mouse interfaces. Voice input technology has been used to overcome some of the challenges associated with conventional interfaces and continues to evolve as a promising way to interact with the EHR.

**Objective**   This article reviews the literature and evidence on voice input technology used to facilitate work in the EHR. It also reviews the benefits and challenges of implementation and use of voice technologies, and discusses emerging opportunities with voice assistant technology.

**Methods**   We performed a systematic review of the literature to identify articles that discuss the use of voice technology to facilitate health care work. We searched MEDLINE and the Google search engine to identify relevant articles. We evaluated articles that discussed the strengths and limitations of voice technology to facilitate health care work. Consumer articles from leading technology publications addressing emerging use of voice assistants were reviewed to ascertain functionalities in existing consumer applications.

**Results**   Using a MEDLINE search, we identified 683 articles that were reviewed for inclusion eligibility. The references of included articles were also reviewed. Sixty-one papers that discussed the use of voice tools in health care were included, of which 32 detailed the use of voice technologies in production environments. Articles were organized into three domains: Voice for (1) documentation, (2) commands, and (3) interactive response and navigation for patients. Of 31 articles that discussed usability attributes of consumer voice assistant technology, 12 were included in the review.

**Conclusion**   We highlight the successes and challenges of voice input technologies in health care and discuss opportunities to incorporate emerging voice assistant technologies used in the consumer domain.

### Keywords
► communication
► natural language processing
► human–computer interaction
► clinic documentation and communications

## Background and Introduction

### Traditional Electronic Health Record Interface

The electronic health record (EHR) serves as a depository of longitudinal patient health information.[1] EHR adoption increased in the early 2000s in the United States largely due to federal incentives tied to the Health Information Technology for Economic and Clinical Health Act of 2009 and the Meaningful Use initiative.[2] EHRs offer benefits over traditional paper records with features such as ubiquitous remote access, digital storage of information making data searchable, and storage of data elements in discrete coded structures. In theory, these features should lead to more efficient entry and retrieval of relevant patient information.[1]

While EHRs promise efficient data storage and retrieval, current state EHRs suffer from usability challenges leading to workflow inefficiencies and end-user dissatisfaction.[3–6] These usability problems undermine a key expectation for the EHR, which is to help users find information necessary to deliver care easily.[3,4,7] User concerns about EHR usability became so pervasive that Meaningful Use incentives were reallocated to improve usability.[8] Despite this effort, usability assessments found that commonly used certified EHRs lacked adherence to the Office of the National Coordinator certification requirements and usability testing standards.[9]

One usability challenge in EHRs pertains to the inefficient navigation of interfaces and records using keyboard/mouse interactions. In a paper record, a provider familiar with the physical handling of paper records may identify and manually modify the list of a patient's medical problems more efficiently than in an EHR. Healthcare providers frequently cannot operate the EHR simultaneously while engaging in patient care. For example, when visiting patients during inpatient rounds, a provider may be able to write down relevant notes quickly on paper between rooms, but the time required to connect to an EHR and find the appropriate screen to document introduces inefficiency and delays.

The keyboard and mouse are the standard input devices for the EHR. Typing on a keyboard is limited to 80 words per minute (WPM), and when using a mouse, the WPM decline further. During patient encounters, physicians using EHRs in exam rooms spend one-third of the time looking at and navigating through the electronic record.[10] It is unknown how this compares to the historical method of documenting on paper, but providers often complain that attention required for keyboard and mouse use introduces new behavioral patterns such as "screen gaze" leading to decreased eye contact and impaired patient engagement.[11,12]

Problems associated with information quality have been linked to keyboard use due to poor word processing capabilities involving incorrect spelling, "copy-paste," and "empty phrases" associated with predefined macros.[13,14] The physical aspects of the keyboard and mouse may contribute to interaction inefficiencies like selecting incorrect information and incorrect documentation in EHRs.[15]

Handheld devices (e.g., tablets and mobile phones) have gained popularity allowing EHR access through touchscreen interfaces.[16] Several studies found that physicians view tablet use in clinical settings favorably because they are lightweight, portable, and increase clinician efficiency.[17] Noted negatives included inadequate keyboards complicating text entry[18] and infectious risks.[19]

The inefficiencies and usability challenges imposed by the standard keyboard and mouse when interacting with the EHR resulted in an interest in alternative modalities, such as voice input. This article reviews the literature and evidence on voice input technology to facilitate work in the EHR and health care, the benefits and challenges of implementation, and potential future opportunities.

## Methods

### Data Sources

To identify relevant literature, we searched MEDLINE using a combination of the following phrases: "Dragon," "Dictaphone," "dictation," "EHR," "EMR," "interactive voice response and speech," "IVRS," "macro," "Nuance," "Tangora," "Vocera," "voice assistant" in combination with the terms "Voice" or "Speech." We also searched "IBM," "Microsoft," "L&H" in combination with the term "dictation." In addition, we reviewed the references of identified key articles for additional literature. We searched Google for the phrase "voice assistant" to identify relevant consumer articles discussing the emerging technology (►Fig. 1).

### Inclusion and Exclusion Criteria

There were no publication date restrictions. We included articles that discussed voice technology systems in production use for patient care. We excluded articles not published in English, not discussing the use of voice technology in health settings, or with no discussion about the fidelity of the voice recognition technology.

### Study Selection and Data Extraction

Members of the investigation team (Y.K., C.P., H.W., E.G., S.A.) reviewed the complete texts of eligible articles for relevance, abstracted information, and provided a recommendation (include/exclude) followed by a group review. Discrepancies between the initial and group reviews were resolved through consensus. The search identified 683 articles that were reviewed for eligibility. The references of included papers were also reviewed. Of 61 included papers that discussed the use of voice tools in health care, 32 detailed voice technology use in production environments. Of 31 reviewed consumer articles discussing usability attributes of consumer voice assistant technology, 12 articles were included.

## Results

### Speech as a Communication Modality

Speech is a natural method of communication that distinguishes humans from other animals.[20,21] Speech provides faster communication than typing or writing. Speech averages between 110 and 150 WPM,[22] while typing averages 40 WPM and handwriting 13 WPM.[23] Speech is the preferred time-saving modality for people with poor typing skills.[24] Speech
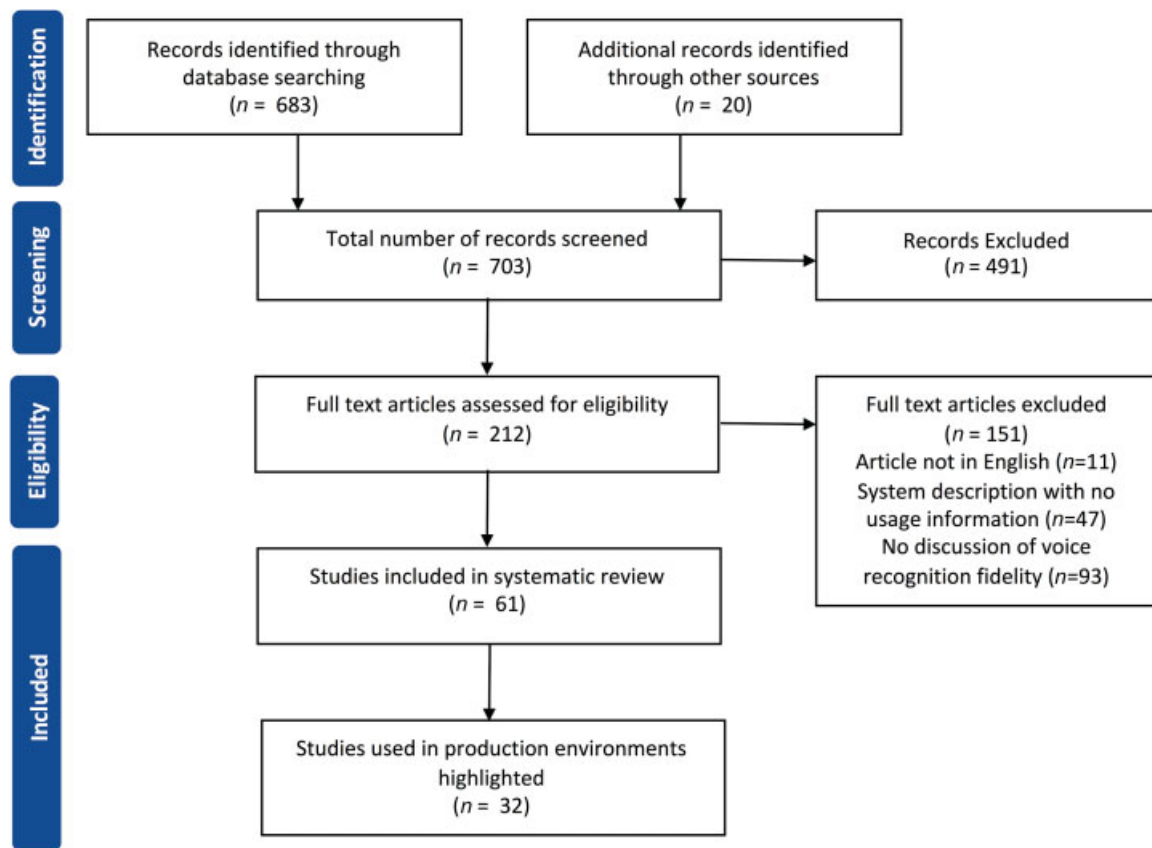
**Fig. 1** Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram for systematic review of electronic health record interactions through voice.

provides a more accessible interface for people with disabilities preventing keyboard or mouse use.[25]

When writing or typing, a user's tactile activity leads to the final output of content. With speech, however, the content is delivered using sound as a medium. With speech communication, an audible message has to be interpreted and understood by the recipient. Speech, therefore, introduces a new source of error that stems from the misinterpretation of the spoken words. While humans can use the context of discourse to appraise the communicated information, when computers are tasked with the interpretation of human speech, the lack of human context can lead to nonsensical transcriptions. The benefit of speech as a faster and more natural way of communication is offset by the requirement to manage the inaccuracies resulting from erroneous interpretation. This tension encapsulates the issues surrounding speech as a communication application in the EHR and health care, where speech interfaces providers to documentation and conveying commands, and patients to menu navigation.[26–28]

### Voice for Documentation

The foremost use of voice in health care has been speech recognition (SR) for dictation (►Table 1).[29] With SR software, a user speaks words into a microphone, and the spoken words are transcribed into electronic text. Use of dictation for transcription in health care was described in the radiology literature in the 1980s and adopted in other specialties subsequently.[26,30] Initially, SR required discrete speech input with the user pausing after each word. These systems were initially less efficient than standard dictation methods but showed promise as emerging modalities.[31,32] With improved technology in the late 1990s, continuous voice recognition became the new standard.[27]

Early research in SR[27] explored opportunities for cost reduction and timesaving compared with traditional dictation methods using transcriptionists. Comparing computerized speech dictation to transcription resulted in cost savings and faster completion.[33–41] Despite cost savings over transcription, savings were limited compared with previous methods like typing or writing on paper.[26] Further, the introduction of voice recognition was associated with significant upfront costs. Software and hardware combinations cost around $250,000 for larger health care institutions and more than $15,000 per user for smaller installations in 2002 (~$340,000 and ~$20,000, respectively, in today's dollar adjusted for inflation).[42–44] The substantial implementation effort needed and technical issues such as connectivity problems and software delays were limiting factors.

Using voice recognition, words misunderstood by the software can be problematic especially if the user is required to perform time-consuming corrections.[45] Leaving a misinterpreted word uncorrected may result in unclear documentation, embarrassing errors, and patient safety issues.[45–54] The provider burden to modify misunderstood words has

**Table 1** Voice for documentation

| Author | Year | Setting | Key findings |
|---|---|---|---|
| Motyer et al[47] | 2016 | Radiology | Occurrences of error in 4.2% of reports with potential to alter report interpretation and patient management |
| Hanna et al[58] | 2016 | Emergency radiology | Template usage decreased audio dictation time by 47% |
| Ringler et al[48] | 2015 | Radiology | Occurrences of material errors in 1.9% of reports that could alter interpretation of the report<br>Errors decreased with time |
| du Toit et al[49] | 2015 | Radiology | Occurrence of clinically significant error rates of 9.6% for SR versus 2.3% for dictation transcription |
| Dela Cruz et al[114] | 2014 | Emergency department | No difference in time spent charting or performing direct patient care with SR versus typing<br>Less workflow interruptions with SR than typing |
| Williams et al[34] | 2013 | Radiology | Radiologists using human editors dictated 41% more reports than those who self-edited |
| Hawkins et al[59] | 2012 | Radiology | Use of prepopulated reports did not affect the error rate or dictation time |
| Basma et al[50] | 2011 | Radiology | Reports generated with SR were 8 times as likely to contain major errors as reports from transcriptionist |
| Chang et al[51] | 2011 | Radiology | Occurrence of 5% nonsense phrases in the reports |
| Hart et al[35] | 2010 | Radiology | Time for document to completion was 2.2 d with SR versus 6.8 d with transcriptionist |
| Kang et al[36] | 2010 | Pathology | Median turnaround time was 30 min with SR versus over 3 h with transcriptionist |
| Bhan et al[37] | 2008 | Radiology | SR took 13.4% more time to produce reports than with transcriptionist<br>Efficiency improves with English is a first language, use of headset microphone, macros, and templates |
| Quint et al[52] | 2008 | Radiology | 22% reports contained potentially confusing errors |
| McGurk et al[53] | 2008 | Radiology | SR reports had 4.8% error versus transcribed reports with 2.1% errors<br>SR errors more likely with noisy areas, high workload, and nonnative English speakers |
| Kauppinen et al[38] | 2008 | Radiology | Reports available within 1 h was 58% with SR versus 26% with transcriptionist |
| Pezzullo et al[45] | 2008 | Radiology | SR reports took 50% longer to dictate than transcriptionist reports<br>90% of SR reports contained error prior to sign off versus 10% transcriptionist reports with errors prior to sign off |
| Thumann et al[33] | 2008 | Dermatology | SR records took 15 min per page and letters were sent after 3.2 d versus transcriptionist method of 24 min per page and letters sent after 16 d |
| Rana et al[39] | 2005 | Radiology | SR reporting time was 67–122 s faster than with transcriptionist<br>Radiologist spent 14 s more time when using SR<br>No difference in major errors |
| Issenman et al[24] | 2004 | Outpatient pediatric | Time required to make correction for SR was 9 min versus 3 min with transcriptionist |
| Zick and Olsen[54] | 2001 | Emergency department | SR accuracy was 98.5% versus 99.7% with transcriptionist<br>Number of corrections for SR per chart was 2.5 versus 1.2 with transcriptions<br>SR turnaround time was 3.6 min versus 39.6 min with transcriptionist |
| Chapman et al[40] | 2000 | Emergency department | SR turnover time was 2 h 13 min versus 12 h 33 min with transcriptionist |
| Lemme and Morin[41] | 2000 | Radiology | SR turnaround was 1 min versus 2 h for transcriptionist |
| Ramaswamy et al[56] | 2000 | Radiology | SR turnaround was 43 h versus 87 h with transcriptionist |
| Massey et al[31] | 1991 | Radiology | SR report generate time was 10.0 min versus 6.5 min with transcriptionist |
| Robbins et al[32] | 1987 | Radiology | SR took 20% longer to dictate<br>12% wording was beyond the SR lexicon scope |

Abbreviation: SR, speech recognition.

been a cause of dissatisfaction with SR technology. Because computers have limited capabilities to format and correct grammar, providers spend more time correcting mistakes using computer transcription than with human transcription.[33] In a 2003 study, computer transcription made 16 times more errors than human transcription including misrecognized words, unspoken words mistakenly inserted in the text, words recognized as commands, and commands recognized as words.[55] Since the time required to edit text is about twice the time needed to dictate,[27,29,45,56] the main reason for discontinuing SR use for 70% of users was the time required to correct errors.[27] This limitation can lead to a hybrid approach of voice recognition in conjunction with mouse and keyboard further reducing efficiency.

While purported benefits of voice dictation in the EHR are speed and productivity, comparing self-typing versus a hybrid approach with SR resulted in increased documentation speed (26% more characters per minute) but also in increased document length (almost doubling). Overall productivity was decreased, but participant mood improved.[57] Template use to guide user input may help to reduce the variability of the data entered using speech and reduce errors.[27,58,59]

The speed advantage of using dictation is especially apparent when users are not particularly adept at using a keyboard and mouse. When words are understood correctly, dictation software may help users with spelling and may reduce the need to correct mistyped words.[25,60]

SR platforms can process a maximum number of WPM while maintaining accuracy (usually slightly greater than 100 WPM depending on the platform). User and software training, domain-specific dictionaries, and medical vocabularies can improve accuracy.[61] However, training requires additional time investment creating a barrier for many users. Vendors of newer voice recognition software platforms advertise that little to no training is required.[62]

A user's accent may complicate SR. Accents modulate word meaning[63] but may make transcription more difficult and require accent-adapted dictionaries or additional training.[37,53,64]

The accuracy of modern voice recognition technology has been described as high as 99%.[65] Some reports state that SR is approaching human recognition.[66,67] However, an important caveat is that human-like understanding of the context (e.g., "arm" can refer to a weapon or a limb. Humans can easily determine word meaning from the context.) is critical to reducing errors in the final transcription.

Other limitations of SR tools include challenges with small modulations in tone and speech rate that may result in transcription errors. Users have to minimize disfluencies such as hesitations, fragments, and interruptions (e.g., "um") that the software might misinterpret resulting in reduced accuracy.[33] Users must be conscious of their speech behavior and cadence when dictating which can make the interaction harder and less natural and may distract a speaker from the content. Technical issues such as managing connectivity, software delays, computer performance, and the time required to load the product and preparing it for use may also be limiting factors. As computing systems become more efficient, these issues may be solved, but for now remain influential factors when considering an investment in current systems.

SR for documentation can pose problems for performance reporting which requires structured data.[68] While dictation software supports the expressivity of free text documentation well,[69] navigating through structured fields that contain selection options can be cumbersome. A provider may dictate in the note that they provided smoking cessation counseling to a patient, but the lack of structured data will require manual or data mining efforts to extract information for reporting purposes. Quality of primary care evaluation found that physicians, who dictated their notes, had lower quality of care scores than physicians using structured and free text documentation.[12] A hybrid approach of dictating while keyboarding through structured fields may lead to duplication of efforts and user frustration. Incorporation of voice user interfaces that allow for voice command navigation through structured fields may help with this process.

## Voice for Commands

In addition to SR for data entry, there are voice recognition tools in production that allow users to make commands via voice (►Table 2). One such tool uses SR for data retrieval by users placing commands to their EHR through macros. Macros allow a single instruction to expand automatically into a set of instructions to perform a set of particular tasks. In the case of voice recognition, macros allow the user to associate a voice command with a sequence of mouse movements and keystrokes that are executed when the macro is verbally initiated.[44]

Voice-facilitated data entry such as dictation performs to different degrees of satisfaction than voice-facilitated commands for data retrieval contributing to different perceptions of utility. For data entry, such as free text note dictation,

**Table 2** Voice for commands

| Author | Year | Setting | Key findings |
|---|---|---|---|
| Friend et al[80] | 2017 | Perioperative environment | 89% of calls placed were understood by the system on the speaker's first attempt |
| Salama and Schwaitzberg[84] | 2005 | Operating room | All voice commands were understood by the system Voice commands were faster than nurse assist |
| Simmons[42] | 2002 | Physical therapy practice | Macro creation decreased time required for dictation |

a boundless corpus of information is communicated to and recognized by the system. Thus, there are many more opportunities for the system to misinterpret the user's spoken word and create errors. Fields with input restrictions (e.g., fields that only accept numeric values) are more likely to be completed accurately because only a limited set of variables has to match. In data retrieval via voice commands, the scope of requests the system can fulfill is narrower and has a lower potential for misinterpretation when attempting to fulfill these requests.[70,71] However, speech for dictation is a more widely utilized modality than speech for commands, because of the extra work to create, and support voice command interfaces.[27]

A benefit of macro commands is that the user can execute complex commands to navigate the EHR and import large bodies of text with a short voice trigger. Departments with substantial repetitive dictation such as radiology have seen benefits from voice recognition associated with macros.[29,44] Some institutions have applied voice recognition for documentation with standardized templates, such as autopsies and gross pathology descriptions, using synoptic preprogrammed text associated with key descriptive spoken phrases.[33,36,72] Early adopters embracing voice recognition developed templates that formatted a report into standardized sections and macros that inserted a body of standard text into report sections, for example, "insert normal chest X-ray."[27,73,74]

Hoyt and Yoshihashi[27] showed that voice-initiated macros consisting of inserted text were used by 91% of users, who continued to use voice recognition. Among these users, 72% rated macros very to extremely helpful. In contrast, 41% of those, who discontinued voice recognition use, had used macros and only 17% rated macros very to extremely helpful.[27] These findings suggest that high use of macros may contribute to the perceptions of higher productivity and accuracy.[27] Personalizing macros by users versus system developers adds to perceived benefits.[44]

Langer[75] found that speech macros increased productivity and Green[76] reported that more "powerful" macros performing functions such as loading predefined templates and inserting spoken text into proper positions were important factors in the success of SR technology.[44] In dentistry, macros have been used to navigate the chart and record data when the provider cannot directly interface with the keyboard.[77] Nursing workflows have utilized macros to retrieve information from patients charts such as allergies.[78] These examples suggest benefits for SR automating routine tasks.

Disadvantages of macros include requirements to train, memorize, and understand their use. Users may have to invest time to program macros, which discourages use. Users must memorize the name of the macro or the triggering word and say it in a specific way to execute the macro. Macros do not use inherent semantic understanding and are based on execution of the saved phonemes.[74] The maintenance of macros may pose a sustainability problem due to system updates that require modification for continued usage.[44] Knowledge management of macros that may call/trigger other macros (nesting) may be complex.

A tool used in health care known as Vocera uses SR commands to facilitate tasks such as initiating phone calls, reviewing messages, and to authenticate logins. Users speak preprogrammed commands to evoke the desired action. While the system has good user acceptance, for some users the system has difficulty recognizing the person that the user is trying to contact[79,80] resulting in calls to the wrong individual. Challenges with Vocera and similar communication devices generally involved SR accuracy and concerns about privacy.[79–82]

A participant in a study evaluating Vocera for health care communication remarked, "The technology works about as well as the voice recognition on my smartphone – about 75% of the time. This is not… an effectiveness level sufficient for critical care." SR data show that 89% of calls placed via Vocera were understood on the speaker's first attempt at the time of the postimplementation survey. Recognition improved to 91% on first attempt in 2016.[80] To further increase first-time SR, users can train the system using the "Learn a Name" feature, which is helpful especially to those with strong accents, by creating custom voice copies of the pronunciation and intonation of other users' names.[79] The ability to call team names or roles instead of persons (e.g., "Team A Intern 1" or "Nursing Supervisor"), improved accuracy and recall and prevented failure when the name of the specific individual was unknown—a common occurrence in hospital settings with multiple daily hand-offs.

Voice commands can facilitate work in the operating room in conjunction with touchless gesture interaction and eye-tracking tools to aid in operations when the user is unable to simultaneously interact with a computer and maintain their sterile field.[83] Notable beneficial findings included voice functionalities supplementing gesture input.[83] Voice commands were faster than a nurse assist.[84] During laparoscopic surgery, voice-guided actions such as light source adjustments, and video capture were well understood. Additionally, voice commands allowed the circulating nurses to concentrate on patient care rather than on adjusting equipment during the surgery.[84]

### Interactive Voice Response Systems for Patients

Many people have perceptions of voice communication technologies based on personal experiences using interactive voice response systems (IVRS) in the consumer domain. IVRS in health care can facilitate interactions as patient-facing tools for phone triaging (►Table 3). The natural language processing (NLP) used in these applications has a more limited scope to facilitate the triage of very specific and simple cases.

Frequently, IVRS are viewed unfavorably because of problems with accuracy. Voice-assisted EHR systems were more prone to errors even in simple use cases.[46] Other complaints about IVRS include the restrictions on exchanges/uses permitted and the depersonalization of the customer service experience.[85] Users frequently become frustrated with IVRS because options do not reflect a person's desired response. IVRS are often used as gatekeepers (or deterrents) to collect information prior to permitting the user to speak to a person.

**Table 3** Interactive voice response systems

| Author | Year | System evaluation | Key findings |
|---|---|---|---|
| Bauermeister et al[86] | 2017 | Adherence evaluation | Difficulty recognizing users' voice responses with background noise |
| Krenzelok and Mrvos[89] | 2011 | Medication identification system | Fewer calls were made to center after IVRS implementation |
| Haas et al[28] | 2010 | Medication symptom monitoring | Some evidence of passive refusal with hang-ups on IVRS |
| Reidel et al[88] | 2008 | Medication refill and reminders | Participants expressed frustration about machine versus real person interactions |

Abbreviation: IVRS, interactive voice response systems.

Environmental factors can influence how well IVRS function. Voice requests in noisy environments can degrade the accuracy.[86] Another obstacle is that users of voice recognition products must maintain focus when using the system.[87]

People are significantly more tolerant to repeating phrases to another human than a machine.[87] While NLP of automated medical triage systems may help direct users to the correct resource, patients may prefer to speak immediately to a person for consultation. Forcing a patient to go through an automated system to speak to a clinician can cause dissatisfaction with the complete treatment experience.[88] While the technology supports many valid uses, user barriers must be considered.

IVRS technology was useful for ambulatory e-pharmacovigilance for most patients but encountered problems with "passive refusals" where patients refused to answer calls or hung up on the IVRS.[28] In an IVRS pilot study to improve medication refill and compliance, participants had a negative perception of the technology because the voice recognition system did not function properly.[88] The authors were unable to determine if the negative feedback was due to a dislike for the technology, technical flaws, or both. In a study using IVRS for medication identification, patients would provide a zip code, age, and gender which were used by the system to provide identification of the medicine. Although the accuracy of the IVRS system was set to 100%, the evaluators noticed that call volume to the system decreased overtime, and this was thought to be related to lack of human communication in the call process.[89]

While the general dislike of IVRS systems comes from their use as gatekeepers, newer tools that use voice engagement to assist patients have been viewed more favorably. With the popularization of home virtual assistant tools, some hospitals are starting to incorporate consumer voice devices like the Amazon Echo into the patient care workflow by allowing patients to place meal orders or call their nurse using the device's voice user interface.[90] There is anecdotal acceptance of these tools by patients,[90] but more research is required to demonstrate the utility of these workflows in the care process.

## Discussion

### Learning from Consumer Voice Tools
The concept of a virtual assistant to be able to retrieve information, execute commands, and communicate with a user through natural speech has existed for some time.[91] These interactions have long been imagined as science fiction entities such as the voice responsive computer system in *Star Trek the Next Generation*.[92] To date, naturalistic voice interactions with technology have overcome a necessary threshold in accuracy and utility to turn from futuristic notions to current manifestations.

A 2015 evaluation of 21,281 SR patents granted by the United States Patent and Trademark Office identified that the top 10% of patents considered seminal based on elements such as patent classifications, age of patent, etc. identified Microsoft, Nuance Communications, AT&T, IBM, Apple, and Google as the leading assignees of seminal patents.[93] Nuance Communications owns the most seminal patents in recognition technology while Microsoft dominated in linguistics technology. In 2016, at least 172 seminal patents belonging to the leading 10 seminal patent owners expired and moved into the public domain. This made SR technology theoretically more easily available to the larger market.

Advances in SR have enabled application such as voice assistant technologies to gain popularity in the consumer realm. These tools can facilitate tasks and retrieve information using natural verbal commands. Examples include voice assistants on smartphones and personal assistants like the Amazon Echo that can coordinate appliances via the Internet of Things. Voice assistants offer a more natural way to interact with technology similar to interactions with another person.

The most popular voice assistant software tools including the Amazon Echo, Google Home, Microsoft Cortana, and Apple Siri provide application programming interfaces for developers.[94] Services like Apple's Siri and Microsoft's Cortana allow users to control applications within the limited scope of their own operating systems.[94] Amazon Echo and Google Home are designed more openly and allow users to develop tools and skills to serve a diversity of external functions. The systems that employ "always on" listening modalities use local component processing to identify trigger words before sending the user's request to cloud-based service where more powerful processing of the commands can be handled. The local processing is considered more secure than the cloud-based, but the cloud infrastructure is better equipped to manage complex NLP tasks and then to return the appropriate configured actions.[95,96]

Artificial intelligence (AI) voice recognition tools use machine-learning models to improve interactions and responses over time. Historically, hidden Markov models (HMM) and Gaussian mixture models (GMM) have been widely utilized to determine the probabilities of potential next words and improve both recognition as well as response. With continuing research into machine learning, deep neural networks and recurrent neural networks (RNN) are beginning to show higher levels of accuracy than more traditional frame-works. This is due primarily to inefficiency in the way that GMMs model low-dimensional systems like SR, which can be better modeled by RNNs. Furthermore, RNNs perform well even on data sets with extremely large vocabularies, which is important to consider because of the medical vocabulary size. Neural networks have been known to predict HMM states for decades, but recent improvements in hardware and algorithms have made them significantly more efficient and applicable.[97,98] The use of semantic networks and hierarchies allow for checks of relations between concepts and their instances in a sentence, which can guide the judgments of content words and further improve accuracy.[99]

For voice recognition tools to complete user requests, vendors often use intent schemas to develop custom interactions. Intent schemas outline the request by providing the NLP with the requested task and variables. They represent the action that corresponds to a user's spoken request. Intent schemas are comprised of two properties: intents and slots. The pattern varies between systems, but generally, an intent is the action that is to be fulfilled (e.g., "Get Patient Data") and the slots represent the relevant data needed (e.g., demographics). Phrases are then mapped to each intent, as a variety of commands could ask for the same task to be completed. Slots contain a type similar to most variables and are customizable. For example, a "Get Patient Data" intent could have custom types (Blood Type), (Weight), (Patient) with values (A, B, O), (1–500), and (free text), respectively. Types must be values that can be spoken by the NLP and the user. Slots refer to variables that a user may request. Using the previous example, an example phrase would be "Tell me (Get Patient Data) what the type (Blood Type) is for Mr. Smith (Patient)." Slots are used to request and fulfill the task requested. Vendor-specific details of schema management can be found in each system's respective development documentation.

Given the flexibility that AI voice assistant tools afford compared with traditional command interfaces like macros, incorporating this new technology into the EHR might prove useful to facilitate interactions that are more natural. A significant concern due to the cloud architecture many of these tools employ is data storage and privacy when dealing with patient information and protected health information (PHI). The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is United States law that provides data privacy and security provisions for safeguarding the electronic exchange of medical information.[100] Currently, the main consumer voice assistant tools including Amazon, Google, Microsoft, and IBM may not meet all the standards of HIPAA compliance for their voice assistant modules.[101–105] A workaround to the HIPAA problem may be possible by using the NLP and machine-learning engines from the Web services to perform the machine learning and retrieval requests, but developing a platform that separates the PHI from the information that is sent to the Web service, which would, however, introduce an additional level of complexity.

## Potential Use of Voice Assistants in Health Care

Evaluation of voice technology to facilitate provider work while the patient is interacting directly with the practitioner has described positive acceptance. Dahm et al described that the work of a provider dictating a consultation letter with the patient present can be viewed as coconstructing of the dictated notes.[106] Positive aspects of this interaction model for the patient included establishing rapport with the provider, building trust, clarifying information, and aiding information accuracy. These have been associated with increased patient satisfaction[107] and decreased patient anxiety levels.[106–110] Negative aspects of this interaction model include confusing patients with technical language and patients being uncomfortable interrupting the provider to make corrections.[107,109,111]

Emerging use of voice assistants in health care include data retrieval, command execution, and chart navigation. Medical dictation software vendors such as Nuance Communications Inc. and M*modal are working with EHR vendors like Epic Systems Corporation to incorporate AI voice assistants into the EHR.[112] EHR vendor eClinicalWorks has launched a virtual assistant tool to help users navigate their EHR interface.[113] Due in part to the Meaningful Use initiative, more structured and standardly named data elements exist in the EHR that these tools may utilize. Unstructured data in the EHR may become easier to query in the future.

Errors associated with SR can result in unsafe conditions when producing content such as prescriptions or initiating action that will affect the delivery of care. Given that errors in health care information submission and retrieval may have far more serious effects than other SR applications, it is important to consider the efficacy and safety of such tools. Hodgson et al evaluated emergency department physicians, who used the Cerner Millennium EHR suite with the FirstNet ED component for keyboard and mouse or the Nuance Dragon Medical 360 Network Edition for SR.[46] They found that tasks done by voice recognition instead of keyboard and mouse had significantly more errors with approximately 138 errors compared with 32 errors in the evaluation of 8 documentation tasks including patient assignment, assessment, diagnosis, orders, and discharge.[46] This highlights the need for caution and vigilance when using these tools, and the need for specialized decision support to facilitate these workflows.

## Conclusion

Given the many usability challenges EHR users face, there is potential for emerging voice assistant tools to help users navigate the EHR more productively. There are opportunities to improve the contextual awareness of these systems to understand what users would want to communicate in different circumstances. Further research is required to understand the impact of these tools on workflow and safety. The optimal use

cases that would benefit from the dialogue type interactions of a voice assistant must be identified in addition to the use cases that could result in safety and privacy risks. It will be important to consider how key EHR interactions such as decision support can be incorporated into the voice interaction to guide best practices.

After cautiously addressing these issues, adoption of these new technologies in the EHR will help train them to recognize medical language and workflows and improve over time. There is potential to develop additional functionalities to facilitate patient care, but we must take careful steps when incorporating these tools into medical workflows to learn their strengths and eliminate the weaknesses. With proper implementation, these tools may offer a path away from the constraints and inefficiencies imposed by classic graphical user interfaces to more naturalistic voice interactions with the EHR.

## Clinical Relevance Statement

EHR interactions through voice continue to evolve as an alternative to standard input methods. Virtual assistants offer a promising approach of communicating naturally with the EHR.

## Multiple Choice Questions

1. What specialty was the earliest adopter of dictation for transcription?
   a. Endocrinology
   b. Radiology
   c. Pulmonology
   d. Obstetrics

   **Correct Answer:** The correct answer is option b. Early research in speech recognition was performed in radiology to explore opportunities for cost reduction and time-saving compared with traditional dictation methods using transcriptionists. The tools were adopted by other specialties over time, but radiology is largely credited with pioneering the use of these technologies for clinical documentation.

2. What is the primary reason for the discontinuation of speech recognition among users?
   a. Time required to develop macros
   b. Cost of maintaining the software
   c. Time required to correct errors
   d. Greater number of words per minute achieved by handwriting

   **Correct Answer:** The correct answer is option c. The main reason for discontinuing the use of speech recognition for the majority of users was the time required to correct errors. Although speaking averages more words per minute compared with typing and handwriting, the work to correct misunderstood words can be a time-consuming process offsetting the benefits of faster speech processing.

## References

1 King J, Patel V, Jamoom EW, Furukawa MF. Clinical benefits of electronic health record use: national findings. Health Serv Res 2014;49(1 Pt 2):392–404

2 Joseph S, Sow M, Furukawa MF, Posnack S, Chaffee MA. HITECH spurs EHR vendor competition and innovation, resulting in increased adoption. Am J Manag Care 2014;20(09):734–740

3 Middleton B, Bloomrosen M, Dente MA, et al; American Medical Informatics Association. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013;20(e1):e2–e8

4 Friedberg MW, Chen PG, Van Busum KR, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Rand Health Q 2014;3(04):1

5 Zhang J, Walji MF. TURF: toward a unified framework of EHR usability. J Biomed Inform 2011;44(06):1056–1067

6 Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. Risk Manag Healthc Policy 2011; 4:47–55

7 Natarajan K, Stein D, Jain S, Elhadad N. An analysis of clinical queries in an electronic health record search utility. Int J Med Inform 2010;79(07):515–522

8 Ray A. Implementation and Usability of Meaningful Use: Usability Panel. Certification Commission for Healthcare Information Technology (CCHIT); 2017 Available at: https://www.healthit.gov/sites/default/files/facas/iu_mu_raytestimony072313.pdf. Accessed November 12, 2017

9 Ratwani RM, Benda NC, Hettinger AZ, Fairbanks RJ. Electronic health record vendor adherence to usability certification requirements and testing standards. JAMA 2015;314(10):1070–1071

10 Montague E, Asan O. Dynamic modeling of patient and physician eye gaze to understand the effects of electronic health records on doctor-patient communication and attention. Int J Med Inform 2014;83(03):225–234

11 Kazmi Z. Effects of exam room EHR use on doctor-patient communication: a systematic literature review. Inform Prim Care 2013;21(01):30–39

12 Linder JA, Schnipper JL, Middleton B. Method of electronic health record documentation and quality of primary care. J Am Med Inform Assoc 2012;19(06):1019–1024

13 Tsou AY, Lehmann CU, Michel J, Solomon R, Possanza L, Gandhi T. Safe practices for copy and paste in the EHR. Systematic review, recommendations, and novel model for health IT collaboration. Appl Clin Inform 2017;8(01):12–34

14 Edwards ST, Neri PM, Volk LA, Schiff GD, Bates DW. Association of note quality and quality of care: a cross-sectional study. BMJ Qual Saf 2014;23(05):406–413

15 Sittig DF, Singh H. Eight rights of safe electronic health record use. JAMA 2009;302(10):1111–1113

16 Anderson C, Henner T, Burkey J. Tablet computers in support of rural and frontier clinical practice. Int J Med Inform 2013;82(11):1046–1058

17 Saleem JJ, Savoy A, Etherton G, Herout J. Investigating the need for clinicians to use tablet computers with a newly envisioned electronic health record. Int J Med Inform 2018;110:25–30

18 Walsh C, Stetson P. EHR on the move: resident physician perceptions of iPads and the clinical workflow. AMIA Annu Symp Proc 2012;2012:1422–1430

19 Hirsch EB, Raux BR, Lancaster JW, Mann RL, Leonard SN. Surface microbiology of the iPad tablet computer and the potential to serve as a fomite in both inpatient practice settings as well as outside of the hospital environment. PLoS One 2014;9(10):e111250

20 Hockett CF. The origin of speech. Sci Am 1960;203:89–96

21 Zuberbühler K. Linguistic capacity of non-human animals. Wiley Interdiscip Rev Cogn Sci 2015;6(03):313–321

22 Williams S. Can Dragon Speech Recognition Beat the World Touch Typing Record? IT Pro Portal; 2015

23 Bledsoe D Jr. Handwriting speed in an adult population. Adv Occup Ther Pract 2011;27(22):10

24 Issenman RM, Jaffer IH. Use of voice recognition software in an outpatient pediatric specialty practice. Pediatrics 2004;114(03): e290–e293

25 Garrett J, et al. Using speech recognition software to increase writing fluency for individuals with physical disabilities. J Spec Educ Technol 2011;26(01):25–41

26 Ajami S. Use of speech-to-text technology for documentation by healthcare providers. Natl Med J India 2016;29(03):148–152

27 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;7:1e

28 Haas JS, Iyer A, Orav EJ, Schiff GD, Bates DW. Participation in an ambulatory e-pharmacovigilance system. Pharmacoepidemiol Drug Saf 2010;19(09):961–969

29 AHIMA. Speech Recognition in the Electronic Health Record. AHIMA Practice Brief; 2013

30 Leeming BW, Porter D, Jackson JD, Bleich HL, Simon M. Computerized radiologic reporting with voice data-entry. Radiology 1981;138(03):585–588

31 Massey BT, Geenen JE, Hogan WJ. Evaluation of a voice recognition system for generation of therapeutic ERCP reports. Gastrointest Endosc 1991;37(06):617–620

32 Robbins AH, Horowitz DM, Srinivasan MK, et al. Speech-controlled generation of radiology reports. Radiology 1987;164(02): 569–573

33 Thumann P, Topf S, Feser A, Erfurt C, Schuler G, Mahler V. Digital speech recognition in dermatology: a pilot study with regard to medical and economic aspects [in German]. Hautarzt 2008;59 (02):131–134

34 Williams DR, Kori SK, Williams B, et al. Journal Club: voice recognition dictation: analysis of report volume and use of the send-to-editor function. AJR Am J Roentgenol 2013;201(05): 1069–1074

35 Hart JL, McBride A, Blunt D, Gishen P, Strickland N. Immediate and sustained benefits of a "total" implementation of speech recognition reporting. Br J Radiol 2010;83(989):424–427

36 Kang HP, Sirintrapun SJ, Nestler RJ, Parwani AV. Experience with voice recognition in surgical pathology at a large academic multi-institutional center. Am J Clin Pathol 2010;133(01):156–159

37 Bhan SN, Coblentz CL, Norman GR, Ali SH. Effect of voice recognition on radiologist reporting time. Can Assoc Radiol J 2008;59(04):203–209

38 Kauppinen T, Koivikko MP, Ahovuo J. Improvement of report workflow and productivity using speech recognition–a follow-up study. J Digit Imaging 2008;21(04):378–382

39 Rana DS, Hurst G, Shepstone L, Pilling J, Cockburn J, Crawford M. Voice recognition for radiology reporting: is it good enough? Clin Radiol 2005;60(11):1205–1212

40 Chapman WW, Aronsky D, Fiszman M, Haug PJ. Contribution of a speech recognition system to a computerized pneumonia guideline in the emergency department. Proc AMIA Symp 2000:131–135

41 Lemme PJ, Morin RL. The implementation of speech recognition in an electronic radiology practice. J Digit Imaging 2000;13(02, Suppl 1):153–154

42 Simmons J. What works: speech recognition. Talking it through. Busy physical therapy practice converts from manual transcription to voice recognition. Health Manag Technol 2002;23(02):38

43 LLC. C.M.G. US Inflation Calculator; 2015. Available at: http:// www.usinflationcalculator.com/. Accessed June 14, 2018

44 Green HD. Adding user-friendliness and ease of implementation to continuous speech recognition technology with speech macros: case studies. J Healthc Inf Manag 2004;18(04): 40–48

45 Pezzullo JA, Tung GA, Rogg JM, Davis LM, Brody JM, Mayo-Smith WW. Voice recognition dictation: radiologist as transcriptionist. J Digit Imaging 2008;21(04):384–389

46 Hodgson T, Magrabi F, Coiera E. Efficiency and safety of speech recognition for documentation in the electronic health record. J Am Med Inform Assoc 2017;24(06):1127–1133

47 Motyer RE, Liddy S, Torreggiani WC, Buckley O. Frequency and analysis of non-clinical errors made in radiology reports using the National Integrated Medical Imaging System voice recognition dictation software. Ir J Med Sci 2016;185(04):921–927

48 Ringler MD, Goss BC, Bartholmai BJ. Syntactic and semantic errors in radiology reports associated with speech recognition software. Stud Health Technol Inform 2015;216:922

49 du Toit J, Hattingh R, Pitcher R. The accuracy of radiology speech recognition reports in a multilingual South African teaching hospital. BMC Med Imaging 2015;15:8

50 Basma S, Lord B, Jacks LM, Rizk M, Scaranelo AM. Error rates in breast imaging reports: comparison of automatic speech recognition and dictation transcription. AJR Am J Roentgenol 2011; 197(04):923–927

51 Chang CA, Strahan R, Jolley D. Non-clinical errors using voice recognition dictation software for radiology reports: a retrospective audit. J Digit Imaging 2011;24(04):724–728

52 Quint LE, Quint DJ, Myles JD. Frequency and spectrum of errors in final radiology reports generated with automatic speech recognition technology. J Am Coll Radiol 2008;5(12):1196–1199

53 McGurk S, Brauer K, Macfarlane TV, Duncan KA. The effect of voice recognition software on comparative error rates in radiology reports. Br J Radiol 2008;81(970):767–770

54 Zick RG, Olsen J. Voice recognition software versus a traditional transcription service for physician charting in the ED. Am J Emerg Med 2001;19(04):295–298

55 Al-Aynati MM, Chorneyko KA. Comparison of voice-automated transcription and human transcription in generating pathology reports. Arch Pathol Lab Med 2003;127(06):721–725

56 Ramaswamy MR, Chaljub G, Esch O, Fanning DD, vanSonnenberg E. Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. AJR Am J Roentgenol 2000;174(03):617–622

57 Vogel M, Kaisers W, Wassmuth R, Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: randomized controlled trial. J Med Internet Res 2015;17(11):e247

58 Hanna TN, Shekhani H, Maddu K, Zhang C, Chen Z, Johnson JO. Structured report compliance: effect on audio dictation time, report length, and total radiologist study time. Emerg Radiol 2016;23(05):449–453

59 Hawkins CM, Hall S, Hardin J, Salisbury S, Towbin AJ. Prepopulated radiology report templates: a prospective analysis of error rate and turnaround time. J Digit Imaging 2012;25(04):504–511

60 Bruce C, Edmundson A, Coleman M. Writing with voice: an investigation of the use of a voice recognition system as a writing aid for a man with aphasia. Int J Lang Commun Disord 2003;38 (02):131–148

61 Liu F, Tur G, Hakkani-Tür D, Yu H. Towards spoken clinical-question answering: evaluating and adapting automatic

speech-recognition systems for spoken clinical questions. J Am Med Inform Assoc 2011;18(05):625–630

62 Nuance Communications, Inc. Nuance's Dragon Medical One Accelerates Adoption among Clinicians for Clinical Documentation. Nasdaq Globewire; 2017. Available at https://globenewswire.com/news-release/2017/08/24/1100145/0/en/Nuance-s-Dragon-Medical-One-Accelerates-Adoption-among-Clinicians-for-Clinical-Documentation.html. Accessed November 12, 2017

63 Cai ZG, Gilbert RA, Davis MH, et al. Accent modulates access to word meaning: Evidence for a speaker-model account of spoken word recognition. Cognit Psychol 2017;98:73–101

64 Kat LWF. Pascale, Fast accent identification and accented speech recognition. In 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. IEEE; 1999:221–224

65 Nuance Communications Inc. Nuance Launches New Version of Dragon Medical Practice Edition [Press release]; 2013. Available at: http://investors.nuance.com/news-releases/news-release-details/nuance-launches-new-version-dragon-medical-practice-edition. Accessed November 12, 2017

66 Meeker M. Internet Trends 2017. Kliener Perkins; 2017. Available at: http://www.kpcb.com/blog/internet-trends-report-2017; 1–355. Accessed November 12, 2017

67 Weller C. IBM speech recognition is on the verge of super-human accuracy. Business Insider; 2017

68 Guide BNL. Capturing high quality electronic health records data to support performance improvement, in Implementation Objective. Office of the National Coordinator for Health Information Technology; 2013

69 Rosenbloom ST, Denny JC, Xu H, Lorenzi N, Stead WW, Johnson KB. Data from clinical notes: a perspective on the tension between structure and flexible documentation. J Am Med Inform Assoc 2011;18(02):181–186

70 Husnjak S, Perakovic D, Jovovic I. Possibilities of using speech recognition systems of smart terminal devices in traffic environment. Procedia Eng 2014;69:778–787

71 Grasso M, Ebert D, Finin T. The effect of perceptual structure on multimodal speech recognition interfaces. Proc AMIA Annu Fall Symp; 1997

72 Klatt EC. Voice-activated dictation for autopsy pathology. Comput Biol Med 1991;21(06):429–433

73 Henricks WH, Roumina K, Skilton BE, Ozan DJ, Goss GR. The utility and cost effectiveness of voice recognition technology in surgical pathology. Mod Pathol 2002;15(05):565–571

74 Sistrom CL, Honeyman JC, Mancuso A, Quisling RG. Managing predefined templates and macros for a departmental speech recognition system using common software. J Digit Imaging 2001;14(03):131–141

75 Langer S. Radiology speech recognition: workflow, integration, and productivity issues. Curr Probl Diagn Radiol 2002;31(03):95–104

76 Green HD. Speech recognition technology for the medical field. J Am Acad Bus Camb 2003;2(02):299–303

77 Nagy M, Hanzlicek P, Zvarova J, et al. Voice-controlled data entry in dental electronic health record. Stud Health Technol Inform 2008;136:529–534

78 Wolf DM, Kapadia A, Kintzel J, Anton BB. Nurses using futuristic technology in today's healthcare setting. Stud Health Technol Inform 2009;146:59–63

79 Friend TH, Jennings SJ, Levine WC. Communication patterns in the perioperative environment during epic electronic health record system implementation. J Med Syst 2017;41(02):22

80 Friend TH, Jennings SJ, Copenhaver MS, Levine WC. Implementation of the Vocera communication system in a quaternary perioperative environment. J Med Syst 2017;41(01):6

81 Richardson JE, Ash JS. The effects of hands free communication devices on clinical communication: balancing communication access needs with user control. AMIA Annu Symp Proc 2008:621–625

82 Fang DZ, Patil T, Belitskaya-Levy I, Yeung M, Posley K, Allaudeen N. Use of a hands free, instantaneous, closed-loop communication device improves perception of communication and workflow integration in an academic teaching hospital: a pilot study. J Med Syst 2017;42(01):4

83 Mewes A, Hensen B, Wacker F, Hansen C. Touchless interaction with software in interventional radiology and surgery: a systematic literature review. Int J CARS 2017;12(02):291–305

84 Salama IA, Schwaitzberg SD. Utility of a voice-activated system in minimally invasive surgery. J Laparoendosc Adv Surg Tech A 2005;15(05):443–446

85 Radziwill N, Benton N. Evaluating Quality of Chatbots and Intelligent Conversational Agents. Cornell University Library. Computers and Society; 2017

86 Bauermeister J, Giguere R, Leu CS, et al. Interactive voice response system: data considerations and lessons learned during a rectal microbicide placebo adherence trial for young men who have sex with men. J Med Internet Res 2017;19(06):e207

87 Lalanne DK. Juerg Human Machine Interaction: Research Results of the MMI Program. Lecture Notes in Computer Science (Book 5440). Berlin Heidelberg: Springer; 2009

88 Reidel K, Tamblyn R, Patel V, Huang A. Pilot study of an interactive voice response system to improve medication refill compliance. BMC Med Inform Decis Mak 2008;8:46

89 Krenzelok EP, Mrvos R. A regional poison information center IVR medication identification system: does it accomplish its goal? Clin Toxicol (Phila) 2011;49(09):858–861

90 Siwicki B. Special Report: AI voice assistants have officially arrived in healthcare. Heathcare IT News; 2018

91 Harnad S. The annotation game: On Turing (1950) on computing, machinery, and intelligence. In The Turing test sourcebook: philosophical and methodological issues in the quest for the thinking computer. Dordrecht, Netherlands: Kluwer; 2006

92 Cadario F, Binotti M, Brustia M, et al. Telecare for teenagers with type 1 diabetes: a trial. Minerva Pediatr 2007;59(04):299–305

93 iRunway. Speech Recognition: Technology & Patent Landscape; 2015. Available at: http://www.i-runway.com/. Accessed June 14, 2018

94 Staff DT. Virtual Assistant Comparison: Cortana, Google Assistant, Siri, Alexa, Bixby. Digital Trends; 2017

95 Habib O. Conversational Technology: Siri, Alexa, Cortana, and the Google Assistant. App Dynamics; 2017

96 Carey S. No, Amazon Alexa Isn't Listening to Your Conversations. Tech World; 2017

97 Hinton G, Deng L, Deng Y. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Process Mag 2012;29(06):82–97

98 Graves A, Mohamed AR, Hinton G. Speech recognition with deep recurrent neural networks. Acoustics, Speech. Signal Process 2013;01:6645–6649

99 Demetriou GC, Atwell ES. Semantics in Speech Recognition and Understanding: A Survey. Proceedings of the 1994 AISB Workshop on Computational Linguistics for Speech and Handwriting Recognition. 1994 AISB Workshop on Computational Linguistics for Speech and Handwriting Recognition, April 12, 1994; United Kingdom: University of Leeds

100 Annas GJ. HIPAA regulations - a new era of medical-record privacy? N Engl J Med 2003;348(15):1486–1490

101 Services AW. Architecting for HIPAA Security and Compliance on Amazon Web Services. Amazon Web Services; 2017

102 Amazon Web Services HIPAA Compliance; 2017. Available at https://aws.amazon.com/compliance/hipaa-compliance/. Accessed November 12, 2017

103 Hiner C. Is your Cloud-based Speech Recognition HIPAA Compliant? NVOQ; 2017

104 Faas R. If You Use The New iPad's Dictation Feature For Work, You Could Be Breaking The Law. Cult of Mac; 2012

105 Glatter R. Why Hospitals Are In Pursuit of the Ideal Amazon Alexa App. Forbes; 2017

106 Dahm MR, O'Grady C, Yates L, Roger P. Into the spotlight: exploring the use of the dictaphone during surgical consultations. Health Commun 2015;30(05):513–520

107 Ahmed J, Roy A, Abed T, Kotecha B. Does dictating the letter to the GP in front of a follow-up patient improve satisfaction with the consult? A randomised controlled trial. Eur Arch Otorhinolaryngol 2010;267(04):619–623

108 Hamilton W, Round A, Taylor P. Dictating clinic letters in front of the patient. Letting patients see copy of consultant's letter is being studied in trial. BMJ 1997;314(7091):1416

109 de Silva R, Misbahuddin A, Mikhail S, Grayson K. Do patients wish to 'listen in' when doctors dictate letters to colleagues? JRSM Short Rep 2010;1(06):47

110 Lloyd BW. A randomised controlled trial of dictating the clinic letter in front of the patient. BMJ 1997;314(7077):347–348

111 Waterston T. Dictating clinic letters in front of the patient. Further research should be qualitative. BMJ 1997;314(7091):1416

112 Nuance Communications, Inc. Nuance and Epic Join Forces on Artificial Intelligence to Revolutionize Disabled Veterans Accessibility to Health IT; 2017 [Press release]. Available at https://www.nuance.com/about-us/newsroom/press-releases/nuance-and-epic-improve-accessibility-for-veterans.html. Accessed June 26, 2018

113 McCluskey PD. Meet Eva, the voice-activated 'assistant' for doctors. Boston Globe; 2018

114 Dela Cruz JE, Shabosky JC, Albrecht M, et al. Typed versus voice recognition for data entry in electronic health records: emergency physician time use and interruptions. West J Emerg Med 2014;15(04):541–547