Provider Adoption of Speech Recognition and its Impact on Satisfaction, Documentation Quality, Efficiency, and Cost in an Inpatient EHR

Kshitij Saxena, MD, MHSA,[1] Robert Diamond,[1] Reid F. Conant, MD, FACEP,[2]
Terri H. Mitchell, MSN, RN-BC,[2] ir. Guido Gallopyn,[2] Kristin E. Yakimow, MS, MBA[2]

[1]Health Quest, Lagrangeville, New York; [2]Nuance® Communications, Inc., Burlington, Massachusetts

ABSTRACT

This study utilizes qualitative and quantitative methods to measure the adoption of speech recognition (SR) and its impact on provider satisfaction, documentation quality, efficiency, and cost when used for clinical documentation within the electronic health record (EHR). Qualitative surveys gauged providers' expectations and experiences regarding documentation before and after SR implementation. A new methodology was developed to measure SR adoption as a proportion of total documentation volume. Quantitative data was collected from the EHR, medical transcription and SR solutions to measure SR adoption and cost savings. Study results revealed significant improvements in satisfaction, documentation quality, and efficiency among providers as a direct result of SR use. An improved provider experience correlated to an 81% reduction in monthly medical transcription costs, an increase from 20% to 77% in electronic clinical documentation adoption, and a 74% SR adoption rate.

Keywords: speech recognition, medical transcription, electronic health records, health information technology, electronic provider documentation, clinical documentation improvement, provider satisfaction

INTRODUCTION AND SIGNIFICANCE

The US health information technology landscape has experienced remarkable transformation. Nearly all hospitals have adopted a certified electronic health record (EHR), but many report barriers, according to new data from the Office of the National Coordinator for Health Information Technology (ONC). In May 2016, the ONC released data briefs on EHR adoption and interoperability based on the American Hospital Association's annual survey. According to one brief, 96% of non-federal acute care hospitals have adopted a certified EHR, up from 71.9% in 2011, the year the federal EHR incentive program launched. Adoption of basic EHR functionality between 2008 and 2015 rose from 9.4% to 83.8%, with eight of ten small, rural, and critical access hospitals having also implemented a basic EHR.[1]

Since the EHR incentive program began, the American Medical Informatics Association (AMIA) and the American College of Physicians (ACP) have released policy meeting position papers regarding the purposes of clinical documentation and challenges that have arisen from EHRs. According to the ACP, requirements for capturing structured data should be kept to a minimum, never take the place of meaningful narrative comments, and EHRs should be leveraged for improving documentation and care.[2] AMIA meeting consensus revealed that in the transition to a more technology-enabled healthcare environment, the main purpose of documentation should be to support patient care and improve outcomes. Documentation for other purposes should be a byproduct of care delivery. Guiding principles for clinical data capture and documentation should support high quality information that is accurate, relevant, confidential, reliable, valid, complete, and secure while supporting downstream uses such as quality measurement, performance improvement, population health care delivery, policymaking, research, education, and reimbursement. Furthermore, central to a learning healthcare system, is the need for data capture and presentation methods that support clinicians' cognitive needs and should align with, but not impede, the care team's workflow and care delivery.[3]

With clinician burnout and shortages on the rise[4], the stakes for improving clinical documentation, workflow, and EHR optimization are imperative. According to a discussion paper published by the National Academy of Medicine (NAM), challenges affecting clinical practice are new payment and delivery approaches, EHRs, patient portals, and publicly reported quality metrics, all of which change how care is provided, documented, and reimbursed. Furthermore, they express concern that more than half of physicians are experiencing substantial symptoms of burnout while nurses report similarly high prevalence of burnout and depression. Work process inefficiencies such as computerized order entry, administrative/clerical burden, and documentation are cited as key contributors to burnout.[5]

As a solution to improve burnout, satisfaction, and other clinical documentation issues, the introduction and integration of speech recognition (SR) technology in EHR systems has made significant progress in the last decade. In 2004, an analysis of the implementation and impact of SR in healthcare concluded that, while SR already had a significant impact on the ability of healthcare providers to operate more cost effectively and provide better patient care, there was still much room for improvement in the way SR technology was adopted and implemented.[6] A study, conducted in 2006, regarding SR acceptance by 186 physicians found experiences were more negative than expectations, particularly with respect to technical performance, medical record quality, and time spent on documentation.[7] Lyons et al.[8] (2015) conducted a replication survey of physicians' expectations and experience with SR technology before and after implementation with emergency medicine physicians. The expectations survey was administered prior to SR training and the experience survey after six months of SR use. In this study, 82% of physicians were initially optimistic that using SR with the electronic record was a good idea. After using SR, 87% agreed that SR was a good idea. In addition, 72% of physicians expected that SR use would save time, while only 51% reported time savings. The increased acceptance of SR in this study was attributed to improvements in the technology and the electronic record. In 2014, a systematic review of SR technology in healthcare concluded that SR systems have substantial benefits and should be considered in light of costs and system selection, training requirements, transcription length, potential use of macros and templates, the presence of accented voices, and workflow patterns. Expectations, combined with the need for prolonged engagement, should be managed during implementation. Improved turnaround times of diagnostic procedure reports or similar tasks represent an important outcome as it impacts timely delivery of quality care. Finally, as technology advances, uptake of SR systems will increase by diverse health and support staff working within a range of healthcare settings.[9] Through the identification of studies published prior to December 2014, Hodgson and Coiera (2016) conducted a literature review, focused on risks and benefits, to assess the impact of SR on clinical documentation. Results discovered that many studies compared SR to dictation and transcription in radiology and heterogeneity across studies was high. Dictation and document editing time increased using SR, while lower rates for clinically significant errors were reported for dictation, which was partially due to documentation being completed by skilled transcriptionists as an added safety check. Turnaround time consistently improved when using SR compared to

dictation across all studies, while SR accuracy also improved. Economic benefits were poorly reported. SR technology is steadily maturing and offers advantages for clinical documentation, however, evidence supporting the use of SR is weak and further investigation is required to assess the impact of SR on documentation error types, rates, and clinical outcomes.[10] In 2015, a randomized control trial compared the impact of SR with self-typing based on measurement of documentation speed, volume, and user satisfaction. This study, based on 1455 medical documents, demonstrated that the availability of SR led to increased documentation speed, increased documentation amount, and higher physician satisfaction.[11]

Ensuring efficient, accurate, timely electronic documentation capture and communication continues to be a key initiative by government bodies and influential national professional associations, thereby, an ongoing focus for healthcare organizations.

Health Quest, the largest family of nonprofit hospitals and healthcare providers in Hudson Valley, New York, was no different from other organizations seeking to improve clinical documentation capture, while also seeking to increase adoption and provider satisfaction with clinical documentation and the EHR. Health Quest implemented Cerner® Millennium® in 2008 and achieved HIMSS Analytics® Electronic Medical Record Adoption Model[SM] Stage 6[12] by 2011. While working to achieve Stage 7, Health Quest sought to improve efficient, timely, accurate, electronic provider documentation capture in addition to provider satisfaction as part of their strategic healthcare information technology priorities.

In 2013, providers were experiencing workflow challenges documenting notes, diagnoses, and treatments, as well as capturing the full clinical narrative in the EHR through traditional input modalities (i.e. handwritten scanned notes, back-end medical transcription, and electronic documentation through manual keyboard [KB] entry). Handwritten notes were scanned into the EHR in batches approximately once every four to six hours, creating delays in the documentation process, making information difficult to find and causing legibility concerns. Back-end, technology-enabled medical transcription was another input modality. Providers dictated patient notes through a microphone into medical transcription speech recognition technology, which captured the voice and transcribed it into a complete patient note through both SR technology and medical transcriptionists' review. Although faster than handwritten scanned notes, this modality required approximately 90 minutes of turnaround time for notes to be available in the EHR (variation exists by note type, policy and procedure, contract requirements and can range from 24-48 hours depending on note type). Through the final input modality, providers used electronic documentation, typing their notes directly into the EHR with point-and-click functionality. This solution included the option to use documentation templates and macros based on the provider's specialty.

Through the integration of SR technology, Health Quest leadership sought to improve provider adoption of and satisfaction with the EHR at each of their hospitals, including the 365-bed flagship hospital, Vassar Brothers Medical Center (VBMC) in Poughkeepsie, New York, in addition to overcoming estimated annual transcription costs exceeding $1 million. Health Quest's primary goals were to drive more accurate, comprehensive electronic provider documentation through front-end speech recognition, improve provider satisfaction with the EHR and clinical documentation process, and increase adoption of the EHR.

In June 2014, Health Quest implemented a pilot of Cerner® Dynamic Documentation™ and Nuance® Dragon® Medical speech recognition technology [DD+SR] with five hospitalists at each hospital, including VBMC, to collect feedback and evaluate their implementation strategy. With SR, providers could speak their notes directly into the EHR in real-time, with greater detail, capturing both structured information and the unstructured clinical narrative. The expectation was that legibility issues, delayed access to clinical information for patient care, and the downstream impact of clinical documentation issues and delays (i.e., availability to nursing, case management, quality, coding, billing, and others) would be greatly reduced or eliminated.

Implementation of DD+SR consisted of technology components that enhanced electronic provider documentation, including the ability to dictate directly into the EHR using specialty-based document "templates" to create comprehensive notes. Providers could now click within a chosen field (e.g. history of present illness, past medical history, physical examination, or operative notes) and begin to dictate that section of the note. Additionally, the ability to speech-enable the EHR-based Smart Templates and auto-texts further enhanced documentation by allowing providers the option to retrieve patient exams, lab results, reports, and other EHR data, to incorporate them into the patient note along with clinical narrative. Finally, by deploying DD and SR together, providers could leverage the EHR "tagging" function, whereby sections of documentation could be tagged for simplified inclusion in reports.

Based upon pilot users' feedback, the implementation team elected to use standardized, specialty-specific auto-text templates and Smart Templates, based on documentation best practices for the continued deployment of DD+SR. The pilot was expanded to all hospitalists in July 2014, intensivists in October 2014, and DD+SR was deployed to all VMBC provider specialties during the final phase of implementation beginning March 30, 2015, which resulted in more than 150 providers trained during that timeframe.

Health Quest partnered with Nuance to conduct a research study during the final phase of DD+SR implementation, at VBMC, with three key objectives in mind. First was to support requests by ONC, AMIA, ACP, NAM, and other organizations seeking to expand upon research and best practices regarding healthcare information technology, clinical documentation improvement, efficiency, workflow, and clinician satisfaction. Second was to expand upon requests for future research on SR noted in the review of literature. The third objective was to validate and share best practices with other healthcare organizations seeking to drive more accurate, comprehensive electronic provider documentation through front-end speech recognition, improve provider satisfaction with the EHR and clinical documentation processes, while promoting EHR adoption and decreasing documentation costs.

The 31-month study covered four periods: pre-implementation (before: January 6, 2014 – May 31, 2014), pilot (June 1, 2014 – March 29, 2015), implementation (transition: March 30, 2015 – May 1, 2015), and post-implementation (after: May 2, 2015 – August 1, 2016).

METHODS

A two-pronged study consisting of *qualitative* and *quantitative* segments was conducted. The *qualitative* segment used surveys to collect information on providers' experiences and expectations regarding clinical documentation before and after deployment of SR to measure provider satisfaction, documentation quality, and efficiency. The *quantitative* segment collected data from the EHR, transcription, and SR solutions to measure note volume evolution per input modality. Additionally, the quantitative segment measured adoption and cost savings. All quantitative measurements were via retrospective, longitudinal, observational study.

Qualitative Survey Method, Instrument, Participants, and Procedure

An electronic survey questionnaire was developed and deployed to measure expectations and experience regarding clinical documentation prior to SR implementation and to evaluate the post-implementation documentation experience including SR. The provider survey instrument was a pair of related and/or overlapping questionnaires that focused on satisfaction, documentation quality, and efficiency **(Table 1)**. The surveys were delivered in two stages. During the pre-implementation period, the survey measured providers' current experience with clinical documentation and expectations of SR. During the post-implementation period, providers were surveyed regarding their experience with the use of SR to facilitate the clinical documentation process. All VBMC survey participants were voluntary and consisted of physicians, physician assistants, and nurse practitioners who were trained on and used SR after training. Respondents included a mix of employed, community, and contracted providers from various specialties. Providers were given the opportunity to complete an electronic pre-implementation questionnaire when they attended training. Post-implementation surveys were administered electronically at a minimum of nine months post-implementation through a maximum of 17 months post-implementation.

Survey Criteria	Questions
Provider Satisfaction	How likely are you to **recommend documentation tools** (pre) currently used in the EHR, (post) SR in combination with EHR to document the patient encounter to a colleague – on a scale from 0 to 10?
	(pre) Is it a good idea (post) Was it a good idea to **introduce SR** via Dragon Medical for documenting in the medical record: completely agree/somewhat agree/neutral/somewhat disagree/completely disagree?
Documentation Quality	After the introduction of SR, the **quality and completeness** of my documentation required for optimal care (pre) will (post) has: improved a lot/improved somewhat/remained the same/declined somewhat/declined a lot?
	After the introduction of SR, the time and effort I spend **answering questions or clarifications** from clinical documentation improvement (CDI), nursing, medical records, quality & coding staff (pre) will (post) has: decreased a lot/decreased somewhat/remained the same/increased somewhat/increased a lot?
Provider Efficiency	SR (pre) will (post) has **improved/optimized workflow** related to clinical documentation: completely agree/somewhat agree/neutral/somewhat disagree/completely disagree?
	After the introduction of SR, the **time spent documenting the patient encounter** (pre) will (post) has: decreased a lot/decreased somewhat/remained the same/increased somewhat/increased a lot?

Table 1: Survey Questions

To study provider satisfaction, researchers used Net Promoter Score® methodology.[13] Respondents were asked, on a scale from 0 to 10, how likely they were to recommend current documentation tools versus SR in combination with the EHR for documenting in the medical record. The NPS methodology groups respondents as follows: Promoters (score 9-10) are loyal enthusiasts who will refer others. Passives (score 7-8) are satisfied, but unenthusiastic. Detractors (score 0-6) are unhappy and can impede use or growth through negative word of mouth. NPS represents number of Promoters minus number of Detractors, divided by total number of respondents, times 100. A positive NPS indicates more Promoters than Detractors; an NPS of +50 is considered excellent. Confidence intervals on NPS and significance of the difference of the pre- and post-implementation NPS were calculated using an established methodology.[14]

To further evaluate satisfaction, provider skepticism regarding SR was assessed pre-implementation by asking whether it is a "good idea to introduce SR for documenting in the medical record." Respondents could express their level of agreement or disagreement as follows: completely disagree, somewhat disagree, neutral, somewhat agree, completely agree. The five levels were translated with "somewhat agree" and "completely agree" mapping to 1, and the others to 0 to calculate binomial proportions. Post-implementation, to assess satisfaction regarding SR, the survey asked if it was a "good idea to introduce SR for documenting in the medical record." The same level of agreement or disagreement options and calculation method was used.

To evaluate the impact on documentation quality required for optimal care with SR, the survey asked about expected improvement in documentation quality and completeness in the pre-implementation survey; and experienced improvement in documentation quality and completeness in the post-implementation survey. Respondents could express level of improvement with SR as follows: decline a lot, decline somewhat, remain the same, improve somewhat, improve a lot. The five levels were translated with "improve somewhat" and "improve a lot" mapping to 1, and the others to 0 to calculate binomial proportions.

To further evaluate documentation quality and efficiency, the survey assessed the time providers spent answering questions and clarifications from clinical documentation improvement (CDI), nursing, medical records, quality, and coding staff, with the use of SR. The pre-implementation question asked about expectations and the post-implementation question asked about the actual impact SR had on time spent answering questions and clarifications. Respondents could express level of time as follows: decrease a lot, decrease somewhat, remain the same, increase somewhat, increase a lot. The five levels were translated with "decrease somewhat" and "decrease a lot" mapping to 1, and the others to 0 to calculate binomial proportions.

To assess expected versus actual impact on provider efficiency with SR, respondents were asked to express their level of agreement or disagreement with whether SR improves clinical documentation workflow as follows: completely disagree, somewhat disagree, neutral, somewhat agree, completely agree. The five levels were translated with "somewhat agree" and "completely agree" mapping to 1, and the others to 0 to calculate binomial proportions.

To further evaluate impact on provider efficiency with SR, the survey asked about expected increase or decrease in time documenting the patient encounter with SR in the pre-implementation survey; and asked about time experienced documenting the patient encounter with SR in the post-implementation survey. Respondents could express time spent documenting the patient encounter with SR as follows: decrease a lot, decrease somewhat, remain the same, increase somewhat, increase a lot. The five levels were translated with "decrease somewhat" and "decrease a lot" mapping to 1, and the others to 0 to calculate binomial proportions.

To calculate 95% confidence intervals on the binomial proportions, the Wilson score interval method was used as recommended by Agresti and Coull.[15]

Some respondents completed both surveys, while others completed either the pre- or post-implementation survey. The methodology proposed by Bland and Butland[16] was used to compare the overlapping pre-implementation and post-implementation survey groups, to calculate the difference in proportions between the two groups, to calculate the confidence interval on the difference, and to perform a hypothesis test of equality of binomial samples with overlap. The results for the subset of respondents who participated in both surveys are presented as well; McNemar's Chi-squared test[17] was used to determine if experience is consistent with the expectation for this group.

Quantitative Methodology, Data Collection, and Data Processing

The quantitative segment is a retrospective, longitudinal, observational 31-month study, from January 6, 2014 through August 1, 2016. During this study, VBMC transitioned from handwritten scanned notes and medical transcription to electronic provider documentation within the EHR.

A goal of this study was to measure how EHR note volume evolved over time for various input methods of clinical documentation: scanned handwritten notes, medical transcription, EHR data entry with keyboard and mouse and EHR data entry complemented with SR use. The evolution and degree of adoption of the EHR with the SR input modality and its effect on the volumes of other input modalities was of particular interest in this study.

A major challenge encountered during this study was the EHR did not capture if a provider note was created with or without SR. To overcome this measurement challenge, a new methodology was developed to recognize whether a note was created with SR or not at the scale of the large note volume (more than 1 million) observed during the study. The methodology is based on the detection of co-occurrence of SR sessions and EHR note change events for the same provider simultaneously active in both the EHR and SR systems. SR sent utilization packages to its database, in ten-minute intervals, when the microphone was on and consisted of speech duration in seconds, number of words recognized, and number of commands spoken. The EHR logged changes to a medical record in its database, which consisted of the following events: when a user opened the GUI wizard to create a new note, when a note was first saved, when a note was edited, when a note was saved again (could be repeated), and when a note was signed.

Data Collection: Custom SQL scripts were developed to retrieve user information; report information; the utilization packages from SR, EHR, and medical transcription systems; and note change-event log information from the EHR for the study period. These collections resulted in 18.28 gigabytes of raw data with 143,352,168 observations.

Data Processing and Interlinking: Users in the EHR and SR datasets and subjects in the survey datasets were interlinked using first and last names based upon the following methodology: A unique identification (ID) was established for 2,307 subjects, interlinking user account information in the EHR, medical transcription (459 users), SR (429 users), and the survey groups. Consecutive SR utilization packages by the same user were aggregated into a "session", including quantification of the total number of words, total seconds of speech, as well as a session begin time and end time corresponding to when the microphone was turned on and off. Sessions without dictated speech were removed, resulting in 320,390 sessions in the study period. EHR report and action data were cleaned and merged in a reports dataset and a report event dataset.

EHR provider notes and SR sessions were interlinked with an algorithm developed for this study based on the following scenario: A provider logged into the EHR, created a new note, activated the microphone, dictated with SR in the EHR, saved and signed the note, and then turned off the microphone. This scenario would lead to the interlinking of the creation of a note in the EHR with the concurrent use of SR. Alternatively, sometimes, the provider created and subsequently saved and/or signed a note in the EHR without the concurrent use of SR. In this case, there would not be interlinking of the use of SR with the creation of that note. There were many possible variations of the exact user actions and sequence of the actions, but concurrent use of SR during the interval between creation and save/sign action of a provider note, within the same EHR session, led to interlinking indicating the use of SR as a part of the effort to create that provider note.

The SR-related user actions were captured in the SR dataset as a SR session, defined as the time interval between when the microphone was turned on and when it was turned off, with a start time (t_{begin}), an end time (t_{end}), the number of words dictated, all assigned to the respective SR user, during the time interval [t_{begin}, t_{end}]. Likewise, the EHR-related user actions were captured in the EHR dataset as a sequence of "create," "perform," "modify," "verify," and "sign" note events with corresponding time-stamps. The "note create" event during a SR session was by itself not an indicator that the note was created with SR, as notes could be created and discarded without saving. Only EHR notes that were saved during a valid SR session were determined to have been created in conjunction with SR. As such, the interlinking, methodology used the "perform" and "modify" events that indicate when a note was saved in the EHR.

Formally, researchers interlinked the SR session S and the EHR note N_k and classified N_k as created in conjunction with SR if the following

$$DragonTouch(N_k) \equiv \exists\, S \in SRSessions: Interlink\,(N_k, S)$$

$$InterLink\,(N_k, S) \equiv \exists\, t_{save}^k \in Save(N_k): t_{begin} < t_{save}^k < t_{end}\ \&\ User(N_k) = User\,(S)$$

$$Save(N_k) = \{t_{save,1}^k,\ t_{save,2}^k,\ \ldots\ t_{save,n}^k\}$$

conditions were true:

An efficient stack decoder algorithm in time complexity $O\,(2N_{sessions} + M_{events})$ to interlink EHR notes and SR sessions was developed based on the R package rstackdeque. EHR report data, SR data, medical transcription data, and interlinking information were merged into one large dataset. The dataset was de-identified; personal health information (PHI) elements and all persons' names and related IDs were obfuscated.

The R statistical software environment and RStudio integrated development environment, in conjunction with the dplyr, tidyr, readr, RecordLinkage, rstackdeque, digest, ggplot, tables, knitr packages, were used for data engineering, exploratory analysis, statistical analysis, and report generation.

RESULTS

Qualitative Survey Findings

During DD and SR training, 121 providers completed the electronic pre-implementation survey. Post-implementation, 108 surveys were completed, resulting in an average of 15 months of SR use for those completing post-implementation surveys. 53 providers completed both pre- and post-implementation surveys. For the 108 post-implementation surveys, provider participation consisted of the following: 64% physicians, 18% nurse practitioners, and 18% physician assistants.

Qualitative results, summarized in **Table 2**, display findings for both surveys, including the subset of 53 providers who participated in both pre- and post-implementation surveys:

Introduction of SR *Survey Criteria*	PRE	POST	Difference	Overlap Hypothesis Test	PRE	POST	McNemar's Chi^2 test
N=	121	108		(p-value)	53	53	(p-value)
Recommend Documentation Tools: Net Promoter Score®	-64 [-75, -52]	35 [22, 49]	99 [81, 117]	2.19e-27	-62 [-80, -45]	35 [16, 56]	3.8e-13
Good idea to introduce SR	73% [64, 80] %	95% [90, 98] %	22% [13, 31] %	9.83e-07	77% [64, 87] %	98% [90, 100] %	0.0023
Documentation quality and completeness	69% [60, 76] %	81% [73, 88] %	11.5% [1.9, 21] %	0.0182	72% [58, 82] %	75% [62, 85] %	0.56
Decreased time spent answering questions & clarifications	49% [40, 58] %	60% [51, 69] %	11.6% [0, 23] %	0.0486	55% [31, 67] %	70% [56, 80] %	0.059
Improved/optimized clinical documentation workflow	57% [48, 65] %	84% [76, 90] %	26.4% [16, 36] %	2.651e-07	60% [47, 72] %	83% [71, 91] %	0.00023
Decreased time spent documenting the patient encounter	46% [38, 55] %	57% [48, 66] %	9.8% [-2.2, 22] %	0.1088	55% [41, 67] %	58% [45, 71] %	0.65

Table 2: Qualitative Survey Results

Responses on likelihood to recommend documentation tools currently used in the EHR (pre-implementation) and SR in combination with the EHR (post-implementation) to document the patient encounter, grouped using Net Promoter Score methodology, are represented below **(Table 3)**:

Survey	N	Detractors	Passives	Promoters	NPS	SE	LCL	UCL
Pre-Implementation	121	89	20	12	-63.6	5.96	75.3	52.0
Post-Implementation	108	16	38	54	35.2	6.97	21.5	48.8
Difference	n/a	-73	18	42	98.8	9.17	80.9	116.8

Table 3: Net Promoter Score

Provider Satisfaction: Survey results revealed that before the implementation of SR, documentation tools had an NPS of -64. The strongly negative NPS, with a sufficiently narrow confidence interval (CI) of [-75, -52], confirms provider dissatisfaction with prior documentation tools. After implementation, however, SR in combination with the EHR, had an NPS of +35, representing a *99-point positive shift* in NPS post-implementation. This shift in NPS, with CI of [81, 117], indicates a dramatic increase in provider satisfaction that can be attributed to SR. The difference is statistically very significant as indicated by the very small p-value of the hypothesis test and the confidence interval on the difference. The result of all participants is consistent with the result in the subset of providers who participated in both pre- and post-implementation surveys. That is, the estimated NPS for the 53-person subset falls in the confidence interval estimated for all participants.

Expanding on provider satisfaction with the use of SR for electronic clinical documentation, the survey asked providers whether they believed it was a good idea to introduce SR for documenting in the medical record. Before implementation, expectations were high; 73% of providers, CI of [64, 80] %, thought it was a good idea to introduce SR. These high expectations were exceeded as 95% of respondents, CI of [90, 98] %, agreed in the post-implementation survey that implementing SR was a good idea. The p-value of the hypothesis test shows the increase of 22 points is statistically significant. The result of all participants is, again, consistent with the result in the subset of providers who participated in both surveys, as the results of the paired group falls in the confidence interval estimated for all participants.

Documentation Quality: Prior to SR implementation, 69% (CI [60, 76] %) of providers *expected* to see an improvement in documentation quality and completeness required for optimal care with the introduction of SR. However, post-implementation, 81% (CI [73, 88] %) of providers reported an *actual* improvement in quality and completeness of documentation required for optimal care while using SR. The difference between expectation and experience is 11.5 points, (CI [1.9, 21]), which is statistically significant for all participants. The subset of providers participating in both pre- and post-implementation surveys had expectations and experiences regarding documentation quality and completeness that were consistent with the results of all participants. Point estimates fall in confidence intervals for all participants. Yet, there is no significant difference in expected versus observed documentation quality and completeness in the subset of providers completing both surveys.

To expand upon documentation quality and efficiency, findings revealed 60% (CI [51,69] %) of providers *experienced* a decrease in time spent answering questions and clarifications from CDI, nursing, medical records, quality, and coding staff, whereas only approximately half (49% CI [40,58] %) of providers *expected* to see a decrease. The difference between expectation and experience is borderline statistically significant. In the subset of 53 providers completing both surveys, the difference is estimated to be 15%, but is not statistically significant.

Efficiency: While 57% (CI [48,65] %) of providers *expected* SR to improve/optimize workflow, 84% (CI [76,90] %) of providers *experienced* improved/optimized workflow after SR was implemented. Experience exceeded expectations by 26 (CI [16, 36]) points, which is statistically very significant. The result of all participants is consistent with the result in the subset of providers who completed both surveys.

Expanding on efficiency, 57% (CI [48,66] %) of providers reported *experiencing* a decrease in time spent documenting the patient encounter when using SR, while only 46% (CI [38, 55] %) *expected* SR to decrease documentation time. While the difference of 9.8% is not statistically significant, the result of all participants is consistent with the subset who completed both surveys. Therefore, it can be concluded that most providers observed decreased time spent documenting the patient encounter with SR which was consistent with their expectations.

Quantitative Results

This study analyzed six different note types (History and Physical [H&P], Consults, Progress Notes, Anesthesia Notes, Surgery & Procedure Notes, and Discharge Notes) and five input modalities (Dynamic Documentation™ using the keyboard [DD+KB], DD+SR, Medical Transcription [MedTrans], PowerNote without SR, and Scanned Handwritten Notes [ScanDoc]) during the study period. **Table 4** provides a breakdown of more than 1 million notes collected in the dataset, after processing and de-identification, as explained in the Methods section. Anesthesiologists at VBMC, using PowerNote, chose to continue using it in the EHR and did not migrate to DD+SR at the time of implementation.

Note Type	All	DD+KB	DD+SR	MedTrans	PowerNote	ScanDoc
				Input Modality		
Anesthesia Note	143262	4	0	0	98747	44511
Consultation Note	117736	14210	53053	39633	236	10604
Discharge Summary	67201	10624	27056	22713	6284	524
History and Physical	126039	10259	29351	14662	6546	65221
Progress Note	453711	92813	230484	0	19871	110543
Surgery & Procedure	162510	4449	14445	19236	7628	116752
All	1070459	132359	354389	96244	139312	348155

Table 4: Number of Notes per Note Type and Input Modality in the Dataset for the Entire Study Period

Note Volume Evolution: **Table 5** provides a breakdown of the mean note volume, measured in notes per week, across all input modalities by note type for the four study periods. Overall, there is a strong increase of 62.5% in note volume from 5873 (mean) +/- 276 (standard error) notes per week to 9546 +/-133 notes per week, with the increase being mostly pronounced in the post-implementation (after) period. H&P, Surgery and Procedure Notes, and Anesthesia Notes show only small variations across the study periods. Discharge Summaries increased 22% from 448.9 +/- 7.9 to 547.5 +/- 10.6 notes per week. Consultation Notes increased 47.6% from 688.6 +/- 59.9 to 1016.3 +/- 15.5 notes per week, and remarkably, Progress Notes increased 244% from 1406.6 +/- 112.9 to 4842.1 +/- 69.8 notes per week. This strong increase in Progress Note volume was not expected by the research team, considering the variation observed in H&P and Discharge Summaries that give an indication of the number of admissions and discharges per week. After consulting with the Health Information Management department, researchers found that Progress Notes were predominately handwritten and scanned into the electronic health record prior to implementation of DD+SR, but were correctly captured in the dataset. It is possible some of this variation may have been due to multiple provider notes on a single handwritten page being scanned in as a single document, whereas post-intervention these documents would be created as separate and multiple documents. At this time, however, the reasons for this unexpected outcome remains fully unexplained, a finding that future researchers may be interested in exploring further.

	Before	Pilot	Transition	After
			Period	
Note Type	Mean ± Std Err	Mean ± Std Err	Mean ± Std Err	Mean ± Std Err
Anesthesia Note	1139.8 ± 59.9	1088.3 ± 37.6	919.8 ± 103.7	996.2 ± 18.0
Consultation Note	688.6 ± 15.6	722.9 ± 21.1	734.0 ± 64.9	1016.3 ± 15.5
Discharge Summary	448.9 ± 7.9	434.2 ± 13.0	420.0 ± 38.0	547.5 ± 10.6
History and Physical	966.3 ± 37.0	912.4 ± 26.7	780.2 ± 78.4	921.0 ± 14.2
Progress Note	1406.6 ± 112.9	1993.2 ± 71.9	2979.6 ± 351.3	4842.1 ± 69.8
Surgery & Procedure	1222.7 ± 80.2	1125.6 ± 35.1	1088.4 ± 135.2	1223.3 ± 22.6
All	5873 ± 276	6277 ± 178	6922 ± 741	9546 ± 133

Table 5: Mean Note Volume, Per Note Type, in Notes Per Week

Input Modality Evolution: **Table 6** provides mean estimates for the five input modalities during the four periods of study. This research revealed a strong decline in medical transcription volume across all note types from 1467.7 +/- 28.9 notes per week to 165.4 +/- 8.7 notes per week, which represents an *88.7% decrease in transcription by number of documents*, from pre- to post-implementation of SR. Regarding the volume of scanned documents in the EHR, this study found *scanned note volume decreased by 34%* from 3228.5 +/- 222.7 notes per week before implementation to 2121.2 +/- 30.8 notes per week after implementation. *DD+KB note volume more than tripled* (236% increase) from 484.2 +/- 32.6 notes per week during the pilot to 1622.6 +/- 36.3 notes per week, while *DD+SR increased 561%* (from 713.2 +/- 53.5 notes per week during the pilot to 4726.9 +/- 85.7 notes per week).

	Period			
	Before	Pilot	Transition	After
Input Modality	Mean ± Std Err	Mean ± Std Err	Mean ± Std Err	Mean ± Std Err
DD+KB	0.0 ± 0.0	484.2 ± 32.6	793.2 ± 114.6	1622.6 ± 36.3
DD+SR	0.0 ± 0.0	713.2 ± 53.5	2206.4 ± 294.9	4726.9 ± 85.7
MedTrans	1467.7 ± 28.9	1146.2 ± 35.5	585.2 ± 80.1	165.4 ± 8.7
PowerNote	1176.8 ± 55.3	1051.2 ± 37.2	1016.8 ± 129.6	942.6 ± 19.5
ScanDoc	3228.5 ± 222.7	2974.5 ± 67.8	2320.4 ± 272.6	2121.2 ± 30.8
All	5873 ± 276	6277 ± 178	6922 ± 741	9546 ± 133

Table 6: Mean Note Volume per Input Modality

Figure 1 displays the evolution of note volume, per input modality, measured in number of notes per week. The study periods, per definition and timeline provided, are indicated by background color: before (red), pilot (orange), transition (purple), and after (green). Local regression (LOESS) trend lines are shown per input modality in addition to the 95% confidence intervals that are indicated as shaded bands around the trend lines. DD+KB (red curve) starts in June 2014, ramps throughout the year, increases steeply during the implementation period, and continues to increase afterwards. *DD+SR (gold curve) becomes the most prevalent input modality* after the implementation of SR. Scanned notes (purple curve) are the most voluminous input modality before the implementation, but the volume declines from the start of the pilot and continues to steadily decline throughout the study period. Medical transcription volume (green curve) declines throughout the study period with a sharp drop during the implementation.
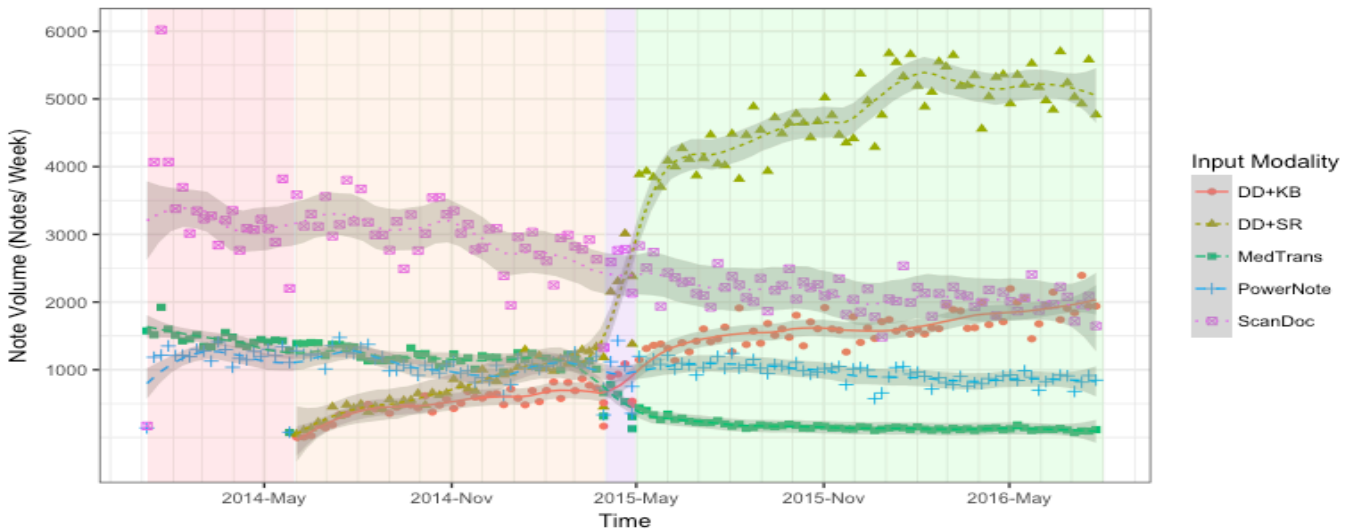

Figure 1: Note Volume Evolution per Input Modality

Input Modality Mix Evolution: **Figure 2** shows the input modality mix, as a percentage, in the four study periods. Before the pilot, 55% of inpatient provider clinical documentation was scanned handwritten notes, 25% was done using medical transcription, and 20% with the use of PowerNote. These modalities changed during the pilot, with the introduction of DD+SR. During the pilot, 11.1% of note volume was created by DD+SR and 7.5 % with DD+KB. During the transition period, there was further adoption of DD+SR. Post-implementation, *half (49.5%) of electronic provider documentation was done with DD+SR*, 17% with DD+KB, and *medical transcription was greatly reduced to 1.7%*. The adoption rate of electronic provider documentation (the combination of PowerNote, DD+KB, and DD+SR) *increases from 20% to 77% of provider notes*. The results above indicate that SR utilization evolved throughout the study period and the majority of DD note volume is captured with the use of SR.
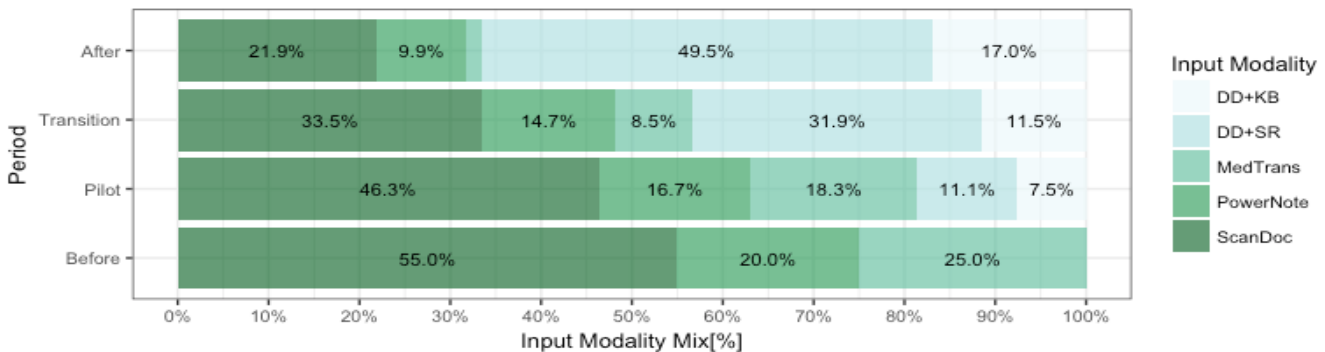

Figure 2: Input Modality Mix Evolution

Figure 3 shows the fraction of DD that is done with SR. Volume starts around 50% at the start of the pilot, ramps up during the pilot to 65%, and jumps to a significant *74.4% adoption of SR when using DD* after implementation. The steep slope of this curve during the pilot, and especially during the implementation, indicates a rapid provider adoption of SR.
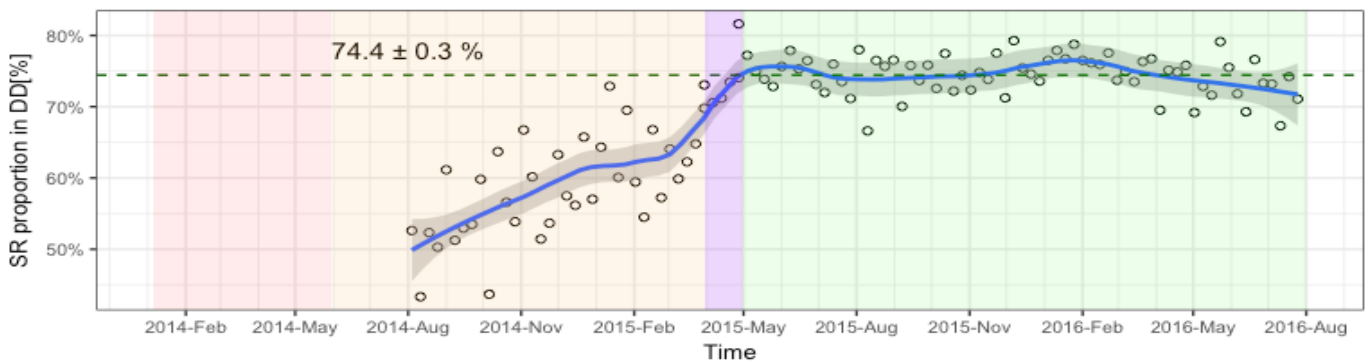

Figure 3: Adoption of SR when using DD Over Time

Financial/Cost Outcomes Analysis: Another important driver for this research was to identify the outcomes of any cost savings that resulted from the implementation of SR technology. Results, demonstrated in **Figure 4,** show the evolution of medical transcription cost and financial savings. Medical transcription costs at VBMC started to trend down during the pilot, because of the high volume of clinical documentation done by hospitalists as early adopters of DD+SR, but researchers sought to assess the overall impact of SR on medical transcription costs and determine whether cost savings existed.

Before the pilot, the mean medical transcription expense per month was $89,400, therefore, the estimated cost at that rate for the entire study period would have been approximately $2.86 million (counterfactual). The actual observed cost, however, was approximately $1.49 million, representing a total cost savings of $1.37 million for VBMC during the study period, or a monthly cost reduction of 81.3%. The financial savings were a direct consequence of the strong decline observed in medical transcription volume as SR was adopted by providers.
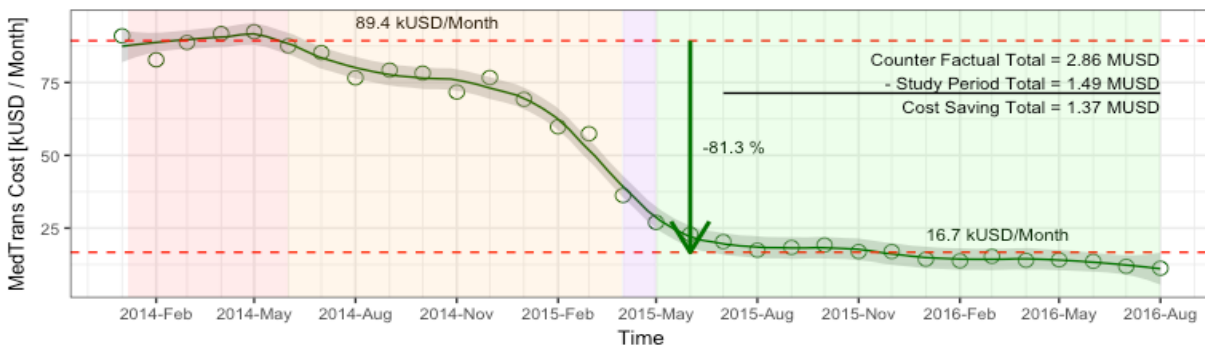

Figure 4: Evolution of Medical Transcription Cost Over Time

DISCUSSION

This is the first study of its kind, to our knowledge, based on the following: First, both qualitative and quantitative methods were utilized concurrently to measure the adoption of speech recognition and its impact on provider satisfaction, documentation quality, efficiency, and medical transcription cost when used for clinical documentation within the electronic health record. Qualitative surveys were used to gauge providers' experiences and expectations regarding documentation before and after the deployment of SR. Second, over a period of 31 months, quantitative data was collected from the EHR, transcription and SR solutions to measure adoption and cost savings. Third, a new methodology was developed and used in this study to determine which provider notes within the EHR were created with SR, to facilitate accurate measurement of note volume evolution per input modality. Fourth, preliminary analysis indicated there may be positive effects on satisfaction, efficiency, and workflow of documentation reviewers, beyond providers, but this will require additional analysis and review to confirm. Validated results revealed significant improvements in satisfaction, documentation quality, and efficiency among providers as a direct result of SR use. An improved provider experience correlated to an 81% reduction in monthly medical transcription costs, an increase from 20% to 77% in electronic clinical documentation adoption, and a 74% SR adoption rate.

According to Hodgson and Coiera's[10] review of literature, to assess the impact of SR on clinical documentation, the mean study duration was 120 days across all studies and the number of clinical documents assessed varied from 14 to more than 300,000. The quantitative segment of this study covered a period of 31 months and analyzed more than 1 million provider notes (studied six note types: History and Physical, Consults, Progress Notes, Anesthesia Notes, Surgery & Procedure Notes, and Discharge Notes; and five input modalities: Dynamic Documentation™ [DD] using the keyboard, DD with SR, Medical Transcription, PowerNote without SR, and Scanned Handwritten Notes).

Provider satisfaction was significantly improved after the introduction of SR as evidenced by the 99-point positive shift in Net Promoter Score® and 95% of providers agreeing it was a good idea to introduce SR for documenting in the medical record. This, coupled with the

studies by Alapetite, et al.[7] and Lyons, et al.,[8] shows a definite rise in both provider expectations and experiences regarding the use of SR, further validating advances made in SR technology. Furthermore, an interesting finding in this current study is that while expectations were high before the introduction of SR, expectations were significantly exceeded after the introduction of SR, which presents an opportunity for additional review and future research.

Study findings also revealed statistically significant improvement in documentation quality and completeness required for optimal care and a borderline statistically significant reduction in time spent answering questions and clarifications from CDI, nursing, medical records, quality, and coding staff. After the implementation of SR, 81% of providers reported improvement in the quality and completeness of clinical documentation, while 60% of providers reported spending less time answering questions and clarifications. Much of the previous research for SR quality and completeness focused heavily on comparing errors or recognition rates versus the quality and completeness provided by supporting clinical narrative capture, reduced time answering questions and clarifications, and its impact to quality and efficiency.[9,10] There is well-documented feedback from professional organizations to preserve the clinical narrative necessary for quality care,[2-5] which presents a key opportunity to expand upon the research findings in this study and others.

To support recommendations proposed by professional organizations,[2-5] documentation should support clinicians' cognitive needs and align with, but not impede, the care team's workflow and delivery of care. After the implementation of SR, 84% of providers reported very significant workflow improvements. This greatly exceeded expectations reported prior to SR implementation and showed remarkable progress. There is limited research focused on the impact SR has on provider and/or care team workflow,[9,10] which presents an opportunity to expand upon this study during future research.

Prior studies by Alapetite, et. al[7] and Lyons, et al.[8] indicated that 3% and 51% of participants, respectively, reported time savings using SR. This current study revealed that 57% of participants reported decreased time spent documenting the patient encounter when using SR, but was not statistically different from the Lyons, et al. study. These findings are also consistent with the Vogel, et al.[11] study. Time savings, when using SR, can be challenging to compare and quantify due to the variation in user utilization of speech-enabled Smart Templates and auto-texts, differences among specialties, and variations in note type utilized. Yet, this presents an interesting future research opportunity regarding implementation best practices.

Post-implementation, the adoption rate of electronic provider documentation increased significantly from 20% to 77% of notes, while medical transcription was greatly reduced to 1.7%. After implementation, given the choice between SR or keyboard and mouse, providers created 74% of note volume with SR when using Dynamic Documentation. This is consistent with survey findings regarding provider satisfaction with SR. These observations could be attributed to many factors, including but not limited to effective integration, implementation, training, SR accuracy, leadership buy-in, and provider satisfaction with the new solution. Furthermore, these findings suggest that SR is indeed a factor in the adoption rate of electronic provider documentation and presents an interesting future research opportunity regarding implementation best practices. While not completed for this study, the opportunity exists to augment this study, and for future research studies, to correlate provider utilization data by input method and per note type to efficiency, workflow, quality, and provider satisfaction. Another interesting research question is the adoption rate of electronic provider documentation without SR compared to with SR.

Preliminary analysis indicated, though not presented in this paper, there may be positive effects on satisfaction, efficiency, and workflow of other documentation reviewers beyond providers. This requires additional analysis to confirm, presenting an opportunity to expand this study and future studies.

Hodgson and Coiera[10] identified some form of economic evaluation in seven previous clinical documentation SR studies. Of no surprise, this current study denoted a significant decline of 81.3% per month in medical transcription costs, which translated to a total cost savings of $1.37 million at Vassar Brothers Medical Center during the study period. Furthermore, the extent and speed of the decline in medical transcription costs observed in the study, within only a few months, is noteworthy. These observations could be attributed to many factors, including but not limited to effective integration; implementation; training, which leveraged lessons learned from the pilot; leadership buy-in; and provider satisfaction with the new SR solution. A hard return on investment is difficult to calculate due to the variability involved. Variables, among others, can include the following: the impact on timely care transitions, discharges and care quality, resulting from immediate access of speech-enabled provider notes that can affect clinical and financial outcomes; financial impact from the potential reduction of delays for coding and billing through timely access to accurate, complete encounter information; efficiency gains of non-providers as a result of less time spent seeking or clarifying information, which was validated in this study as providers felt they spent less time answering questions and clarifications; and potential reduction of provider burnout and turnover rates through improved satisfaction.

LIMITATIONS

Several factors could hinder this study from being generalized to other clinical settings or information systems. Other EHR and SR systems might differ in performance and different approaches to integrating and implementing the two could also lead to varying results. Furthermore, this study was not designed to evaluate clinical effectiveness nor the direct impact on clinical outcomes of one input modality versus another. Post-implementation surveys spanned many months, potentially impacting provider perceptions of SR because of communications with and/or witnessing other providers' use of SR, which may have created unintended bias in survey results. We were unable to collect note length information to measure the effect of SR input modality on note length. The methodology to interlink EHR notes with SR is binary, and, as such, does not provide measurement as to how much SR was used as opposed to keyboard and mouse. No time-motion studies were done with providers, which may have provided greater detail around the actual decrease in documentation time, reduced clarifications, and improved documentation quality when using SR versus other input modalities.

CONCLUSIONS

For many organizations, the search for the most efficient method of capturing clinical documentation, optimizing the EHR, enhancing usability and improving provider satisfaction remains a strategic priority. Given the substantial cost of documentation, in terms of clinician time and varying input modalities, the foundational act of interacting with an electronic health record was the primary focus of this study. By linking provider notes within the EHR to SR, and evaluating both qualitative and quantitative data inputs concurrently, this study created a baseline for the impact of SR use on provider adoption, satisfaction, documentation quality, efficiency, and cost. This new methodology, and

its associated findings, have contributed to the current field of research regarding best practices for integrating and implementing SR within the EHR.

However, as SR technologies continue to extend into different healthcare settings and advance to include more robust narrative and clinical intelligence capabilities, additional studies are needed to augment this research. Furthermore, the healthcare field should evolve the understanding of the ways SR will support providers and downstream documentation users within their workflow, and how improved documentation translates to better clinical and financial outcomes.

COMPETING INTERESTS

The authors indicated with the number 2 are employed by Nuance Communications Inc., a leading provider of speech recognition solutions.

REFERENCES

1. Henry J, Pylypchuk Y, Searcy T, Patel V. Adoption of electronic health record systems among U.S. non-federal acute care hospitals: 2008-2015. ONC Data Brief, no.35, 2016. The Office of the National Coordinator for Health Information Technology, dashboard.healthit.gov/evaluations/data-briefs/non-federal-acute-care-hospital-ehr-adoption-2008-2015.php.

2. Kuhn T, Basch P, Barr M, Yackel T. Clinical documentation in the 21st century: executive summary of a policy position paper from the American College of Physicians. Annals of Internal Medicine, 2015;162:301-303. *ACP*, doi: 10.7326/M14-2128.

3. Cusack CM, Hripcsak G, Bloomrosen M, et al. The future state of clinical data capture and documentation: A report from AMIA's 2011 Policy Meeting. Journal of the American Medical Informatics Association: JAMIA 20.1, 2013, 134-140. *PMC*, doi: 10.1136/amiajnl-2012-001093.

4. NEJM Catalyst. Physician burnout: the root of the problem and the path to solutions: a collection of original content from NEJM catalyst. Massachusetts Medical Society, 2017. https://moqc.org/wp-content/uploads/2017/06/Physician-Burnout.pdf.

5. Dyrbye LN, Shanafelt TD, Sinsky CA, et al. 2017. Burnout among health care professionals: a call to explore and address this underrecognized threat to safe, high-quality care. NAM Perspectives: discussion paper, National Academy of Medicine, Washington, DC. https://nam.edu/burnout-among-health-care-professionals.

6. Parente R, Kock N, Sonsini J. An analysis of the implementation and impact of speech-recognition technology in the healthcare sector. Perspectives in Health Information Management/AHIMA, American Health Information Management Association, vol. 1, no. 5, 2004. www.ncbi.nlm.nih.gov/pmc/articles/PMC2047322/pdf/phim0001-0005.pdf.

7. Alapetite A, Andersen H, Hertzum M. Acceptance of speech recognition by physicians: a survey of expectations, experiences, and social influence. International Journal of Human-Computer Studies, vol. 67, no. 1, 2009, 36-49. doi: 10.1016/j.ijhcs.2008.08.004.

8. Lyons J, Sanders S, Fredrick Cesene D, et al. Speech recognition acceptance by physicians: a temporal replication of a survey of expectations and experiences. Health Informatics Journal, vol. 22, no. 3, 2015. journals.sagepub.com/doi/10.1177/1460458215589600.

9. Johnson M, Lapkin S, Long V, et al. A systematic review of speech recognition technology in health care. BMC Medical Informatics and Decision Making, 2014. doi: 10.1186/1472-6947-14-94.

10. Hodgson T, Coiera E. Risks and benefits of speech recognition for clinical documentation: a systematic review. Journal of the American Medical Informatics Association. 2016 Apr 1;23(e1):e169-e179. doi: 10.1093/jamia/ocv152.

11. Vogel M, Kaisers W, Wassmuth R, Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: randomized controlled trial. Journal of Medical Internet Research, vol. 17, no. 11, 2015. doi: 10.2196/jmir.5072.

12. HIMSS Analytics. Electronic medical record adoption model. 2018. Available from: http://www.himssanalytics.org/emram.

13. Satmetrix Systems, Inc. What is net promoter? 2017. Available from: http://www.netpromoter.com/know/.

14. Hopper, J. How to calculate an NPS margin of error. Versta Research Blog, 2015. Versta Research, www.verstaresearch.com/blog/how-to-calculate-an-nps-margin-of-error/.

15. Agresti A, Coull B. Approximate is better than 'exact' for interval estimation of binomial proportions. The American Statistician, 52.2, 1998, 119-126.

16. Bland JM, Butland BK. Comparing proportions in overlapping samples. Unpublished paper, 2011. Available from: http://www-users.york.ac.uk/~mb55/overlap

17. Agresti, A. Categorical data analysis. Wiley, 1990, 350-354.