SESSION 4B: PACS Technology and Design

The Implementation of Speech Recognition in an Electronic Radiology Practice

Patti J. Lemme and Richard L. Morin

For both efficiency and economic reasons, our practice (200,000 examinations) has converted all remote dictation to speech recognition transcription (PowerScribe, L & H, Burlington, MA). The design criteria included complete automation to the existing radiology information system (RIS), with full RIS capabilities immediately available following dictation. All dictations for computed tomography, magnetic resonance imaging, ultrasound, and nuclear medicine were converted from remote transcription to speech recognition over a 2-week period (following a 4-week installation phase and 8 days of training). The average turnaround time for these reports decreased from approximately 2 hours to less than 1 minute. Reports are then sent to the institutional Electronic Medical Record and are available throughout all facilities in a nominal 2 minutes. Speech recognition rates were surprisingly high, although certain phrases caused consistent difficulties and certain staff required retraining. This presents our analysis of both successful and problematic areas during our design and implementation, as well as statistical performance analyses.

Copyright © 2000 by W.B. Saunders Company

THE MEDICAL PRACTICE at Mayo Clinic Jacksonville (MCJ) consists of 270 physicians and 1,380 allied health staff. All medical and surgical specialties except pediatrics are covered. About 40,000 patients were registered in 1999 for both basic (comprehensive history and physical) and specialty examinations. Due to our demographics (Table 1), the turnaround time for radiology reports is of utmost importance and provides part of the motivation for implementation of speech recognition transcription. Additionally, the medical practice at MCJ has been automated with electronic systems (Cerner, Kansas City, MO) for more than 3 years. All typical medical history (or chart) activities are performed electronically by physicians as well as allied health staff. In addition to the MCJ

practice, Mayo (as well as community) patients are seen at St Luke's Hospital (SLH) located approximately 9 miles from the clinic campus. SLH is a 289-bed facility with access for both MCJ (n = 270) and community (n = 260) physicians.

RADIOLOGY PRACTICE

The Mayo Radiology Practice (MCJ & SLH facilities) consists of 23 radiologists, 103 technologists, and 90 support staff. About 200,000 radiology examinations are performed per year with about 130,000 examinations performed at the Clinic location.

A radiology information system (RIS) has been in place since about 1990. The first RIS was a locally developed mainframe application with report distribution via terminals throughout the Clinic building. This system was replaced in 1995 with an RIS developed by Mayo Foundation (radiology information management system [RIMS]).

Radiology interpretations are performed in both direct and remote methods. In direct dictation, the transcriptionist is seated next to the radiologist and enters the interpretation into RIMS as the radiologist speaks. This technique is used for radiography, fluoroscopy, and mammography interpretations. All other areas (computed tomography, magnetic resonance, ultrasound, nuclear medicine, and special procedures) were traditionally transcribed via standard remote dictation techniques (Digital Dictation System, Lanier, Atlanta, GA). Unlike direct dictation, with remote transcription it was necessary for the radiologist to enter RIMS sometime following dictation, review the transcription, edit, and finalize the interpretation with electronic signa-

Table 1. Demographics for Patients Seen at MCJ

Location	% of Patients
Jacksonville & NE Florida	35
Rest of Florida	33
Rest of United States	28
International	4

From the Department of Radiology, Mayo Clinic Jacksonville, Jacksonville, FL.

Address reprint requests to Richard L. Morin, PhD, Department of Radiology, Mayo Clinic Jacksonville, 4500 San Pablo Rd, Jacksonville, FL 32224. E-mail: morin.richard@mayo.edu.

Copyright © 2000 by W.B. Saunders Company

0897-1889/00/1302-1037\$10.00/0

doi:10.1053/jdim.2000.6859

ture. Not only was this a lengthy process for the radiologist, but also delayed our workflow and thus produced delays for our physicians and patients. With film-based imaging, direct-reported examinations were available (report plus films) to the referring physician within 30 to 45 minutes following radiography. Examinations using remote interpretation were available within 2 to 4 hours. This differential became much greater as we converted to electronic practice with workstation interpretation. Following implementation of the electronic practice, radiography examinations (reports plus images) were available within about 5 to 15 minutes following examination; however, the remotely dictated studies still required 2 to 4 hours due to the nature of transcription, review, and finalization.

The motivations for implementation of speech recognition were several. A primary goal was to increase radiology efficiency, as well as decrease the time between interpretation and final report availability (along with images) throughout the institution. In addition, it had become challenging to find personnel with the appropriate skills and experience.

DESIGN

As speech recognition systems were evaluated, a primary requirement was seamless integration to our existing RIMS. The design model was for the speech recognition system to be an active window on the electronic record/RIMS workstation so that on a single workstation, the radiologist could review patient information and history and authenticate the interpretation following dictation. After vendor selection (PowerScribe, L & H, Burlington, MA), the necessary interface was jointly developed using off-the-shelf hardware and adhering to software standards over a 6-month period. Once validated, it was anticipated that a phased approach with two clinic areas and two hospital areas would be used. During this time, audit checks would be performed to analyze performance.

IMPLEMENTATION

Our initial phased implementation approach was revised. All hardware and software was installed over a 1-month period and implementation was begun in all remote dictation areas at the same time following an 8-day training period. We chose to implement speech recognition with a "big bang"

approach rather than a phased approach because both training and staff rotations made isolation of an area difficult in practice. In addition, daily switches between remote dictation and speech recognition were considered to be disruptive and would delay anticipated benefits. All dictation for computed tomography, magnetic resonance, ultrasound, and nuclear medicine were converted from remote transcription to speech recognition over a 2-week period. The average turnaround for these reports decreased from approximately 120 minutes to less than 5 minutes, and has decreased (over 6 months) to about 3.5 minutes due to both user and system adaptation.

As might be expected, many challenges arose during implementation. While the general acceptance was good, several areas have required further work: workstation configuration, microphone and bar-code reader, recognition accuracy, dictation habits, and interpretation room environment. The microphone was sometimes awkward for some radiologists and presented difficulties in reading the RIMS accession number from the bar-coded worksheet. This step was necessary to load the proper demographics into the transcription window before dictation and successfully load the report to RIMS and the electronic record. As we expected, some speech patterns presented difficulties. In particular, problems included truncation of words and syllables, slurring or interjection of phrases (eg, "hmm"), pausing in the middle of a phrase, and improper punctuation. We also discovered that background conversation and noise in some of our interpretation areas were problematic for speech recognition.

CONCLUSIONS

Based on our experience, speech recognition is a valuable tool to improve overall examination turnaround time. At present, the commitment of radiologists to this technique is necessary, as the conversion is not without pain. For optimal implementation, the interpretation room design should receive attention before implementation of speech recognition. We anticipate that both software and room changes will significantly improve the system and improve both ease of use and accuracy for the radiologists. In addition, this technology is important to accommodate practice growth without concomitant growth in support personnel.