Journal of Digital Imaging

Integrating Dictation with PACS to Eliminate Paper

Thomas E. Warfel, M.D., Ph.D., and Paul J. Chang, M.D.

Five years ago our department migrated from a film-based imaging environment to a PACS-based environment. We discovered that our reliance on paper tracking forms for clinical history and dictation information was hindering our practice. Integrating dictation with PACS was one of three key components we needed to free ourselves from the tyranny of paper, the other two being an online worklist and an online patient history. We discuss our evolution to a (mostly) paperless reading room environment, our implementation, general performance, and future development plans, focusing on integrated dictation.

KEY WORDS: Integrated dictation, report authoring, workflow, remote collaboration, asynchronous collaboration

PRIOR TO THE INTRODUCTION of picture archiving and communications system (PACS) at our institution, studies to be read would arrive in the reading rooms with a paper requisition form clipped to each study's film jacket. The requisition form would have scheduling information printed by the radiology information system (RIS), plus handwritten patient history notes obtained by a nurse or the technologist at the time of study acquisition. Using a telephone-based dictation interface, a radiologist would scan the barcode of each requisition form into the dictation system and then dictate the report for that study. On the rare occasion a barcode could not be scanned, the radiologist manually typed the study's accession number into the dictation interface. (Fig 1).

Five years ago our department migrated from film to a PACS-based imaging environment. Initially, we continued using paper requisition forms for dictation (Fig 2), but studies appeared at the ordering clinicians' PCs (as well as our workstations) minutes to hours before the requisition forms would reach the reading rooms. If we waited for the tracking forms to arrive, our clinical relevance diminished because the clinicians would simply do their own interpretations, relying on us to call if something life-threatening was later seen. Why wait for our interpretations if they could see the images before we read them? Even worse, radiologists wasted time performing two reads of many exams: one person issuing a "wet read" over the phone for the clinician, followed by another radiologist doing the "real read" later when the paper tracking form arrived.

Attempts to make radiologists hand-type the accession numbers for each case into the dictation unit proved inefficient (especially for high-volume studies like chest radiographs) and led to inaccurate reporting, as even single-digit keypad mistakes resulted in an incorrect report being sent out. With a daily workload of 500–600 studies, roughly 10–15 erroneous exams were generated per week. Even worse, if the (erroneous) exam was of the same exam type the patient was scheduled for, there was a chance the error would not be detected. Finally, manual keypad entry is both stressful and time-consuming for the radiologists. Assuming 300 chest studies per day (typical for our institution) at 6 seconds to key each study, approximately 30 min each day was wasted on data entry in

From the Department of Radiology, University of Pittsburgh Medical Center, 3705 Fifth Ave., CHP Suite 3950, Pittsburgh, PA, 15213, USA.

Correspondence to: Thomas E. Warfel, M.D., Ph.D., tel.: (412)-647-7288; fax:(412)-647-7795; e-mail: warfelte@upmc.edu

Online publication 18 February 2004
doi: 10.1007/s10278-003-1715-2

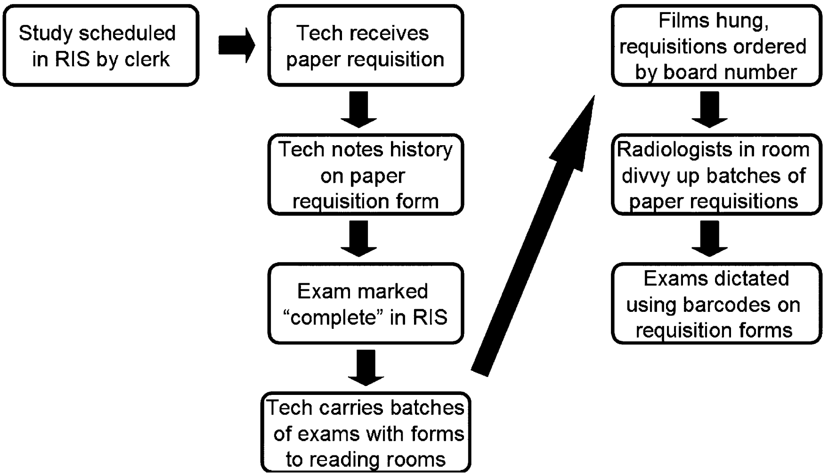

Fig 1. Filmed exams with paper requisition forms.

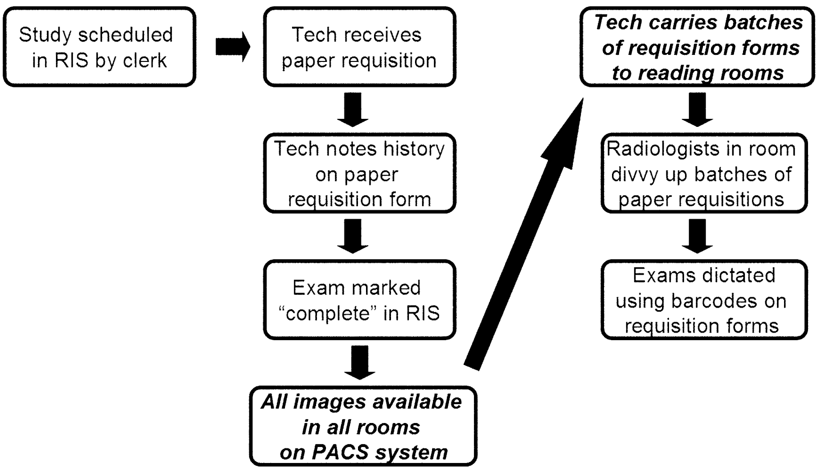

Fig 2. Filmless (PACS) with paper requisition forms.

one reading room (in addition to time spent redictating erroneous exams).

Our solution required development of three components; an online worklist that accurately reflected the list of studies to be dictated, an online history-taking tool for nurses and technologists to enter patient data, and tight integration of report dictation with our PACS system (Fig 3).

This article discusses our development of an integrated dictation system that seamlessly interfaces with our PACS and existing dictation infrastructure. This integration allows a single sign-on for both PACS and dictation authentication and guarantees that the study being viewed on the screen is the study being dictated. Our intent was to minimize both development and implementation costs and to utilize existing infrastructure where possible. We are presently still investigating voice recognition, multimedia reporting,and structured reporting technologies as part of a larger "report authoring" project. The following is a description of our initial integrated dictation system as it is utilized today.

MATERIALS AND METHODS

Integrating dictation into a PACS system requires several components: an audio input device, a means of signaling dictation control to the system (ie, record, play, rewind), software for performing the audio capture, and software for communicating the captured audio to a transcription

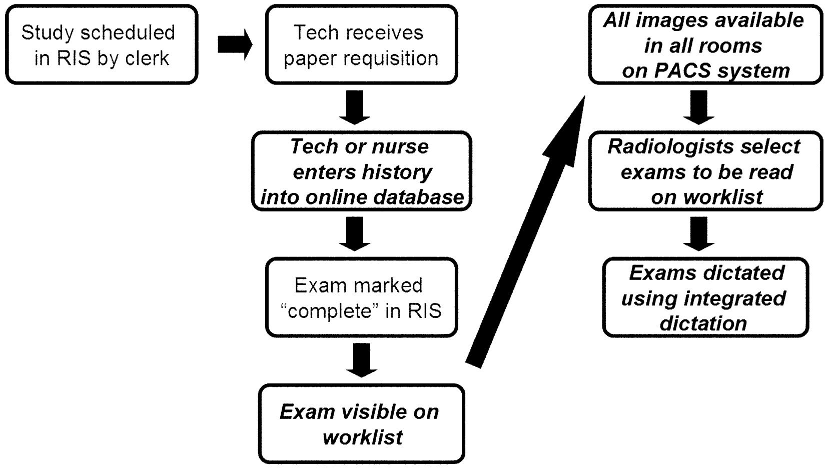

Fig 3. Filmless (PACS) with paperless reading room.

infrastructure. Issues include error handling and detection during audio acquisition, as well as reliable transfer to the transcription infrastructure. Furthermore, the software must be supportable and maintainable across a large number of distributed radiology workstations.

We utilize commercially available USB dictation microphones for physical audio acquisition, proprietary web-distributed ActiveX components for performing client-side audio capture and integration with our PACS software, and a proprietary middleware server layer that relays audio to the existing radiology dictation infrastructure. Our existing dictation system is Dictaphone Enterprise Server, and our RIS (into which the transcriptionists type directly) is IDX-RAD version 9. Because the hospital had a large investment in the existing infrastructure (Dictaphone Enterprise Server, IDXRAD, and training for transcriptionists in their use), the overriding design goal was to minimize capital expenditure while maximizing functionality. This approach also allowed us to focus development effort on the user interface and freed us from worrying about transcription management infrastructure.

Neither CCOW nor IHE were available at the time the system was developed; neither our transcription system, our RIS, nor our PACS supported either technology (although they theoretically could have been purchased as customizations for additional money and additional time delays). Furthermore, current CCOW implementations require both a consistent means of identification across various hospital applications, as well as (reliable) dedicated context servers. Today, CCOW implementations are available for our PACS system, and our institution is moving towards a common identification system; the next version of our integrated document authoring system may use CCOW, enabling its reuse in other medical reporting arenas beyond radiology.

RESULTS

We first implemented integrated dictation in the chest reading room at UPMC Presbyterian University Hospital in July 2002. Clinicians using the system for dictating chest radiographs (approximately 200–300 per day at our institution) reported approximately one-half hour of time saved each day by not having to manually type accession numbers into the old dictation system. We have reduced the number of duplicate dictations (compared to the paper tracking form/PACS dictation method) but do not have firm statistics for this improvement. The system was iteratively developed to enable better integration with the online worklist and has since been deployed to all reading rooms across the UPMC main campus in April 2003. As Langer[1] predicted, we saw a significant decrease in average time-to-dictate for completed studies when the radiologist no longer had to wait for a technologist to bring requisition forms to the reading room. This effect was most pronounced for studies completed in the late afternoon; they are now being dictated the same day of the study rather than done as part of the "morning rush" on the following day.

DISCUSSION

Audio Input and Control Signaling

Traditional dictation systems rely on an omnidirectional handheld microphone with dictation controls integrated into the handset; many of these are essentially touch-tone telephones. A number of commercial microphones exist which rely on a separate PC sound card for audio digitization and communicate button state to the PC through either joystick or RS-232 ports. More recently, both Philips and Dictaphone have begun commercially offering USB (universal serial bus) microphones that integrate a microphone, sound card, and a set of programmable pushbuttons (Fig 4). By integrating the audio digitization circuitry into the handset, the long antennalike leads that would otherwise trail from the microphone to the sound card are eliminated, thus greatly reducing electrical noise. Furthermore, the audio gain for each microphone is uniform from PC to PC; using the same type of USB microphone across all PCs eliminates the variabilities in sound card gain and quality, regardless of whether the PC is a Compaq, Dell, or no-name clone. The control buttons of the USB device are accessed as a USB HID (human interface device) within Microsoft® windows, providing an interrupt-driven signaling mechanism and eliminating the overhead of endlessly polling the joystick port, which in turn simplifies interface programming and improves application performance. Finally, a USB connector is a single plug without any screws; this facilitates both the installation and (when necessary) replacement of the device.

Audio Capture

While Microsoft Windows provides a simple MCI (media control interface) for sound capture, this facility is unreliable when attempting multiple audio insertions and deletions. Furthermore, the software for audio capture must not cause a perceptible delay to the workstation during times of high CPU load (such as when viewing multiple correlated MRI slices), nor may the capture software drop audio during times of high load. We developed a proprietary ActiveX audio acquisition component to circumvent these limitations; it coexists unobtrusively with our third-party PACS software. The acquisition program establishes the existence of the USB microphone, determines that the PACS user has permission to dictate, verifies that space for a dictation is available on the hard drive, and then presents a simple user interface when the user activates dictation (Fig 5).

Fig 4. Philips (left) and Dictaphone (right) USB microphones.

Communication of the Captured Audio to the Existing Dictation Infrastructure

This element has proved the most challenging to implement. Our department uses a Dictaphone Enterprise dictation server that offers limited ability to import WAV files via the proprietary Dictaphone EVIS TCP/IP interface. Unfortunately, the EVIS interface provides limited error reporting, and, when problems do occur at the Dictaphone Enterprise server, the radiologists cannot send dictations to it. To improve availability, we developed a highly reliable middleware tier that receives dictation audio from the PACS workstations, performs automatic gain control and background noise reduction, and relays the "cleaned" dictation to the Dictaphone Enterprise server. Studies can continue to arrive at (and be buffered by) our middleware server regardless of the state of the Dictaphone Enterprise server. Even if there is a catastrophic loss of dictations on the Dictaphone machine (as happened when a hard drive failed three years ago), audio data can be resent from the middleware server when the Dictaphone system is

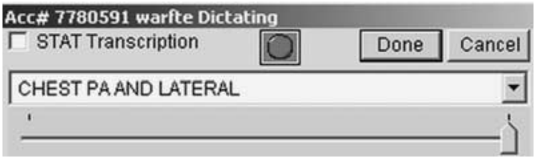

Fig 5. PACS audio capture interface display.

running again. An HL7 interface updates the status of studies on our RIS from "completed" to "dictated" as the dictations arrive at the middleware server (Fig 6).

The middleware layer is composed of several parts: a relay program that conveys dictations from the PACS workstation to the middleware server, two redundant middleware servers (one primary and one secondary), and a copy of the Dictaphone EVA software on each middleware server to relay data to the Dictaphone Enterprise server. Each middleware server has ECC (error corrective coding) memory, as well as redundant power supplies and redundant Ethernet connections. Both units are on UPSs (uninterruptible power supplies), and both have RAIDed disks that allow continued operation even in the event of a hard drive failure. These servers store the incoming reports, perform background noise reduction using the NCT Active ClearSpeech™ software, resample and reformat the data to be compatible with Dictaphone, and finally forward the data via the Dictaphone EVIS utility to the Dictaphone Enterprise server. If either the primary or the secondary unit fails, the other continues alone. Dictations remain on the middleware servers until the final (signed) report is available in the RIS, after which the audio recording is deleted. The original audio file is kept on the PACS workstation where it was created for three days following successful transmission, then it is automatically deleted. While this may seem excessive, the radiology report is our final product, not the images alone. If we cannot produce reports, we are not producing a useful product. Hence, the multitiered redundancy is necessary to prevent loss of data.

HIPAA Compliance Issues

The current implementation performs clear-text network transactions but runs only on PCs that are in physically restricted areas. Also, at present, the audio recordings kept locally for three days are also in the clear. We are presently converting the report format to XML and adding support for more complex multimedia documents, as well as implementing a public-key/shared-key hybrid system for encrypting dictations at the client. This will protect both the locally stored data and network transfers so that radiologists may use the software to dictate from remote sites that may not be physically secure.

Our middleware server maintains a distinct private and public key pair, and the public key is embedded in our PACS audio capture software. When a user logs into the PACS workstation, a pseudorandom "session key" is generated at the workstation, and then it is encrypted with the public key embedded in our audio capture software and registered with the middleware dictation server. All user dictations during that login session then are encoded using that session key. This bundle of encoded audio plus metadata is sent to the middleware server, as well as stored locally for three days. At the server, incoming dictations are decoded in two stages: The session key is looked up from the previous login registration, and the dictation data are decoded using the (known) session key. While this method does not allow nonrepudiation, it does protect the confidentiality of dictations both on the client and across the network and could be extended to support nonrepudiation as part of the login process. Unfortunately, the link from the middleware server to the Dictaphone Enterprise server is under the control of the EVIS software and subject only to Dictaphone's "security through obscurity" protocol. While we are investigating alternatives, this network link occurs on a short hop within our internal hospital network and would be difficult to eavesdrop.

Problems Observed

If a technologist completes a study but fails to send the study to PACS, there is no indication in the reading room that a study needs to be dictated. At present, these studies are detected by an evening batch-mode query run on the RIS to look for completed imaging studies

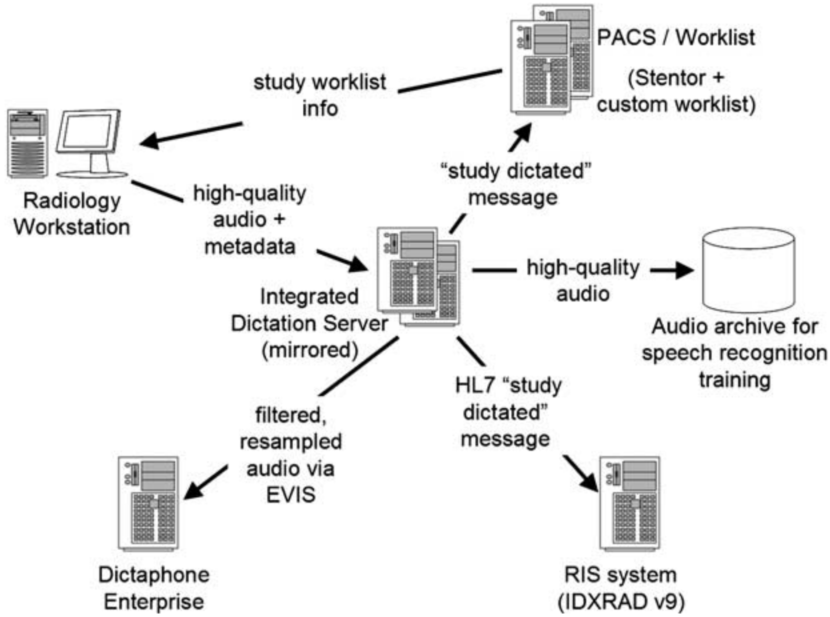

Fig 6. UPMC Integrated Dictation Implementation.

without images in the PACS. It has been suggested that this problem would not exist if the worklist were to be driven by the RIS instead of the PACS. Unfortunately, uncompleted studies which are sent to PACS (which occur more frequently than completed studies which are not sent) would then not be visible to the radiologists and a similar query to look for scheduled studies with images would be needed. Whether the list is driven by PACS or by RIS, similar problems emerge. With the PACS-driven list, uncompleted studies are caught by radiologists at the time of dictation and only the rare unsent studies are in limbo. We are investigating ways of detecting and correcting this human error more rapidly. Ultimately, some form of integration between a DICOM storage commit on PACS, integrated with "study completed" on the RIS, will be needed to fully resolve this problem.

While our integrated dictation system has been deployed throughout UPMC Presbyterian University Hospital, we have not completely divorced ourselves from the legacy dictation (telephone) stations. We still receive a number of outside films to be dictated which are not visible on our PACS system and hence are not visible on the worklist. We are working on ways of adding these non-PACS-image studies to the existing worklist.

Time to Completion

While the combination of integrated dictation with an online worklist and online medical history has shortened our average *time to dictation*, our average time to report sign off has not improved as much as desired. Because the RIS requires separate sign on and times out after a few minutes of inactivity, there is a time/frustration penalty in looking "too early" for reports to be signed. As a result, people tend to wait until the end of the day to sign their reports, delaying the finalization of reports that had been transcribed several hours before. We recently deployed an additional tool (integrated into the worklist) that displays the number of reports waiting to be signed by the user. This has improved compliance among the willing, but there is still no incentive driving people to sign promptly beyond a sense of moral duty.

Future Directions

In addition to a simple display of outstanding studies, we are experimenting with alternative

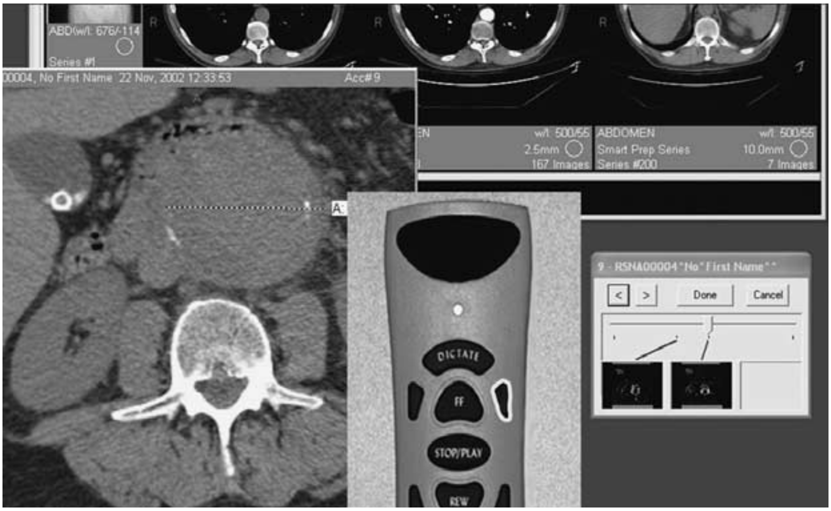

Fig 7. Integrated dictation can be easily extended to enable image tagging.

(more intrusive) notification options, including automated e:mail alerts, automated e:mail to a secretary with instructions to call the reading room and nag the radiologist, and even automated pager alerts when studies to be signed age more than 15 hours. We are also developing a report editor component to integrate with the PACS worklist and dictation tools, eliminating the need to continually relaunch a separate application for signing. Hopefully by lowering the "pain threshold" for signing while increasing the visibility of reports to be signed, we can improve our report completion time as well.

Switching to an XML representation of our report will simplify the inclusion of images thereby facilitating true multimedia report authoring, hi particular, we wish to enable identification of key images and tie them to particular segments of the dictation. Because there are multiple programmable buttons on the dictation microphones, "tagging" key images becomes as simple as pressing a button on the dictation microphone (Fig 7).

In addition, we can construct more complex reports interactively, using structured report generators to create text for routine studies, or even embed dictated free text as a small component within a (larger) structured report for procedures.

Finally, we are actively investigating batch-mode speech recognition technologies and anticipate having a pilot system running within at least one reading room later this summer. Because we will be running in batch mode, we are able to take additional processing time for more confusing or complicated dictations and will be able to dynamically adapt our language model based on the additional clinical information available.

CONCLUSION

There are many advantages to integrating dictation with PACS, among them increased radiologist efficiency and reduction of errors caused by dictating the wrong study. We are gaining between 30 and 50 min per day of radiologist work time by the elimination of manual data entry into the legacy dictation system. Further, adding integrated dictation to an online worklist enables collaboration between radiologists at remote sites by keeping the online worklist current. Key technical factors to be considered when constructing and integrating dictation software include providing a consistent user interface and consistent sound quality across different reading rooms and hardware. Security and reliability are of critical importance—if the dictation system fails, the department grinds to a halt. A combination of early error detection and critical component redundancy enhances reliability. This combination allows us to improve radiologist dictation efficiency while minimizing system downtime despite the increased technological complexity.

ACKNOWLEDGMENTS

We would like to thank Kenneth Crane (UPMC Radiology Informatics) and Audrey Grube (UPMC Radiology Informatics) for developmental design and coding; David Lionetti (UPMC Radiology Informatics) for assistance with PACS/worklist interface Linda Gmitter (UPMC ISD) for assistance with HL7 interfacing between IDXRAD, Dictaphone Enterprise Server, and Stentor; and Timothy Fagert (Dictaphone / PowerScribe) for assistance with early microphone development and debugging efforts.

REFERENCE

1. Langer, SG: Impact of speech recognition on radiologist productivity. J Digit Imaging 15:203-209, 2002