



CAN YOU HEAR



Taking another look at turning voice into data

by Dana Hinesly

ALMOST SINCE THEIR CREATION, speech recognition (SR) programs have been overestimating their capabilities, fascinating casual observers but aggravating those hoping to benefit from their claims.

“To this day, we suffer from the over-promised, underdelivered technology, primarily because the applications are very hardware intensive—and the desktop of ten years ago couldn’t cope to deliver the accuracy that had been demonstrated and that people expected,” explains Nick van Terheyden, MD, chief medical officer of Philips Speech Recognition Systems (Atlanta). The company’s SpeechMagic speech-enables the medical IT solutions of leading international healthcare companies, such as Agfa Corp (Ridgefield Park, NJ), Sectra (Linköping, Sweden), Eastman Kodak Co’s Health Group (Rochester, NY), and Philips Medical Systems (Andover, Mass). Also, the software is used by more than 200 integration partners in the Philips global network, such as Crescendo Systems Corp (Laval, QC), Dolby (Concord, Ohio), Epic Systems Corp (Verona, Wis), Health Care Technology (Marietta, Ga), and MedQuist Inc (Mt Laurel, NJ).

Although the seamless and perfect SR systems portrayed in movies and on television are still more fiction than science, the technology has progressed in recent years by leaps and bounds, thanks to both improved algorithms and computer processors capable of the speeds required to run them.

Making Good on a Promise

The past few years have seen a surge in the adoption of SR programs by medical professionals, for a number of reasons. One significant motivator has been the simultaneous decrease in the number of qualified medical transcriptionists and a steady increase in the demand for patient documentation. This shift in available resources has had an undeniable impact on the dictation-to-completed-report turnaround that radiologists can promise.

Matthew J. Bassignani, MD, medical director for RIS and medical director of the

ME NOW?

University of Virginia (UVA) Imaging Center (Charlottesville, Va), has firsthand experience with this predicament. When he came to the center 5 years ago, SR had yet to be introduced into the facility—and the situation was dire.

“We just couldn’t hire enough transcriptionists to handle the 325,000 studies that we did every year, and we had a three-week turnaround time, most of which was spent waiting for transcription,” Bassignani says, adding that the facility’s radiologists decided during a faculty meeting that change was inevitable. “We realized that none of us could look our clinical colleagues in the face or claim that we were performing any sort of service, because it was taking so long for us to produce the reports.”

Another driving force behind growing SR adoption is the resurgence of interest in developing a truly comprehensive and valuable EMR. As government initiatives work to develop protocols and standards, SR manufacturers are creating software that will usher in this electronic future.

“The only way we can enable access to the information in the EMR is to digitize it,

and if you can automate the data-gathering process—using speech as one of the enablers—you’ll move the EMR to be the central key repository of patient information that clinicians can act on and use,” says van Terheyden. “This type of accessible database is essential, because medicine in its current form is unsustainable; the idea that the physician can be the font of all knowledge and manage the whole process without any technology is completely unrealistic.”

Perhaps the single biggest reason that more people have taken to SR programs is the most obvious: They actually work now. Not only is accuracy improving—many programs boast rates of 99%—but the software is being rolled into comprehensive solutions designed to streamline the physician’s entire job.

“Clinicians don’t want just a speech-recognition system; they want a workflow solution,” says Kulmeet Singh, director of healthcare strategies at Nuance Communications Inc (Burlington, Mass)—the former ScanSoft and the manufacturer of Dragon Naturally Speaking, one of the SR industry’s

dominant programs. “Radiologists express interest in SR solutions that integrate with existing PACS, deal with multiple accession numbers, and retrieve patient demographics—combining it all into the final report.”

Hesitate No More

While they likely dream of an overarching, data-management SR program, radiologists are creatures of habit—they are human, after all—and many bristle at the prospect of the change in workflow this type of system brings with it.

“It is change, but you have to see the benefits and the value that the change will bring,” says Andrew W. Litt, MD, associate professor and vice chairman of financial affairs in the department of radiology at New York University Medical Center. “I think this technology is an absolutely critical part of providing good radiology service today, because radiologists are here to serve other physicians—and the better we serve them, the more we contribute to the overall healthcare of their patients.” At Litt’s facility, radiologists access

>> do you hear what i hear?

JUMP INTO THE BRAVE new world of speech recognition (SR) or stay with transcription? In 2004, the University of California, Los Angeles (UCLA) radiology department was faced with a strategic decision on how to replace an obsolete 14-year-old dictation system. Upgrading to a new dictation system would incur the least capital cost and disruption to the department. SR, on the other hand, offered seemingly attractive potential for long-term savings in transcription costs and improvement in service through faster report turnaround times.

But would SR work in a large academic institution with a diversity of accents and complex workflows? Experiences at other institutions were mixed. From a financial perspective, we needed to reach a self-edit rate using SR of about 43% per year to show savings over a dictation/transcription system in the first year, using a 3-year capital lease model for the capital investment. We felt that this percentage might be too optimistic, and a more realistic rate would be 20% in the first year, 40% the second, and 80% the third. Over the 3-year period, we could achieve savings of more than \$100,000, but with no savings the first 2 years.

It would all hinge on the adoption rate for SR. The decision was made in late 2004 to upgrade to Dictaphone’s PowerScribe SR system.

INTEGRATION

We recognized from the outset that integrating SR into the RIS and PACS to create a seamless workflow was crucial to success. Our Centricity PACS from GE Healthcare (Waukesha, Wis) is tightly integrated with our RIS—the ImageCast system from GE Healthcare/IDX Systems Corp (Burlington, Vt). The radiologist workflow is driven by the RIS worklist on the PACS workstation. To maintain the integrated workflow, we decided that the SR application had to run on the PACS workstation, instead of on a separate PC, so that the radiologist needed to log in only once to the PACS, and the RIS would drive both the SR and PACS.

We were challenged by the existing integration capability between the three systems. Although basic order and result HL7 interfaces were available at the back

end, integration at the user-interface level was nonexistent or hopelessly inadequate. It took several months of work by vendors for the three systems before we felt we had a reasonably seamless system with which we could go live.

The physical environment of the workstation required enhancement. To protect the screen space of the PACS workstation, we added another monitor to our standard two-monitor PACS workstation. This third color LCD monitor displays the SR application for the radiologist to self-edit without obscuring the view of the PACS images.

WORKFLOW

One of the major decisions regarding workflow was whether to restrict all editing to the SR application only. This might sound like a trivial question, but it is not. When a report is created or edited on the SR system, it can be set to upload to the RIS, where it replaces any previous version of that report. The SR system, on the other hand, is unable to receive reports originating from the RIS. If you allow a report to be edited in the SR application and also in the RIS, different versions could reside on the two systems. How do you control changes? This is a real problem when there is more than one author on a report, such as a report dictated by a resident who needs an attending co-signer.

One option would be to require all edits to be done in SR. However, this was not a viable option, at least not during the transition period. Not all the radiologists would have been trained to use SR; and besides, there was no way—at the application level—to prevent users from editing reports in the RIS. We ended up with a workflow requiring different hybrid workflows for attending radiologists and residents.

GO LIVE

With more than 100 radiologists at UCLA, we just did not have the resources to take the big-bang approach, nor did we feel it was the right strategy. Unlike previous implementations of the RIS and PACS, SR was a big unknown to almost all of our radiologists and to the IT group as well. We had no real experience on which to predict

SR technology through the RadWhere Suite of software from Commisure Inc (New York).

Once an SR program is in place, the changes to workflow are immediately apparent, turning today's standard approach on its head. Instead of simply recording audio and sending it off for transcription, dictating with SR means that the report is converted directly into an electronic document right in front of the radiologist's eyes. Necessary edits are made, and the report is electronically signed and forwarded to the referring physician—all in the same sitting.

"This self-edit approach—what we call 'once and done'—gives the provider complete control," says Don Fallati, senior VP of marketing for Dictaphone Corp (Stratford, Conn), recently acquired by Nuance. "We've put a lot of work into making the completion process as comfortable as we possibly can, increasingly using voice commands to navigate through and make changes to the document."

To help speed through these additional steps, SR solutions include tools that can actually shorten the dictation process, such as

preset templates and "trigger" words that, when spoken, insert entire blocks of copy into a report. Making use of these shortcuts can help clinicians dictate even faster than real time, because they're able to avoid repeating "boilerplate" copy.

"We call them standardized reports, and I use them all the time so I don't keep dictating the same things over and over," says Paul M. Williams, DO, diagnostic radiologist at Northeast Regional Medical Center (Kirksville, Mo). "Also, if there's something in the template I don't like or if there's something I want to add, I can easily change it."

MedQuist supplies the dictation, transcription, SR, and coding systems used by the team at Northeast Regional.

Benefits Beyond the Obvious

Radiologists experienced with SR technology are quick to point out that the system's greatest benefits go well beyond its ability to circumvent delays due to medical transcription.

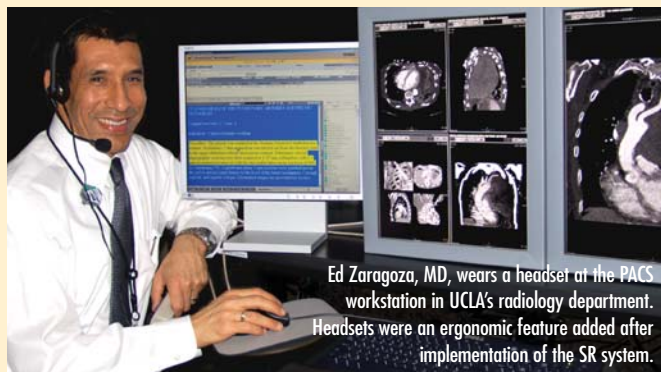
"When there's even 24 hours' delay between dictating the report and getting it

back from the transcriptionist, all you are really doing is grammar checking—because if, for instance, I said 'right' and I should have said 'left,' I'm not going to remember that a day later," UVA's Bassignani says. "Doing the report in real time means that I've definitely cut down on those errors."

Not only are reports becoming more accurate, but they also are being delivered with impressively fast turnaround times. "For patients who go directly from our imaging center to their doctor's office, we expedite reports and deliver them within one hour," Bassignani explains. The standard distribution time between the study's completion and when referring physicians have reports in-hand is less than 22 hours. "Our practice has completely changed with speech recognition, and we're now getting referrals from physicians from outside UVA. Clinicians in the area look at our imaging center as the preferred place to send patients," he says.

This type of service also can translate to a dramatic reduction in phone calls from referring physicians, who no longer have to

A firsthand account of implementing speech recognition at UCLA by Hwa Kho, PhD, MBA



how each radiologist would take to it. We decided to execute the implementation at a very deliberate pace. We had to win the battle one section at a time, radiologist by radiologist, making sure we always had enough IT resources to support them.

The Multi-Specialty Section—a group of five radiologists based in our Santa Monica-UCLA Medical Center, which is about 3 miles from the main campus in Westwood—was chosen as the beachhead for the implementation. This section reads more than 50,000 exams per year across all subspecialties, representing a range of computer skills and voice accents.

Live use started in September 2005. Each radiologist would spend about 40 minutes reading the training scripts to build their voice models. Then, they proceeded to read live cases with the help of a trainer, learning as they went along. By the time they'd read their first 20 cases, the radiologists were usually quite proficient in the basic skills and were able to continue on their own. The trainers continued to spend a lot of time with the radiologists in the reading rooms the first 2 weeks, answering questions, pointing out mistakes, and teaching more advanced features of the application, such as building report templates that populate data elements from the HL7 interface.

CURRENT STATUS

More than 70% of the radiologists throughout the UCLA Medical Center have been trained using the same approach as the one employed at the Santa Monica facility. The existing dictation/transcription system continues to be available for radiologists to use so that they can switch freely from one system to another if they chose. The speed of adoption of the new technology varied from radiologist to radiologist, but overall, we are very pleased with the progress.

Just 5 months into the implementation, about 40% of all reports are now self-edited via SR. This is ahead of our initial, now very conservative, expectation of 20% the first year. Among the very first group of radiologists trained, the Multi-Specialty Section, the adoption rate is even higher: more than 95%. Plus, the number of reports sent to the transcription service has been reduced by half. So, breaking even financially in the first year now seems like a realistic goal.

The improvement in report turnaround time—that is, the elapsed time from completion of an exam to the time a final report is available—has been significant. This is particularly true when looking at the portion of reports turned around within 2 hours. Historically, the largest contribution to the turnaround time was the time it took a radiologist to review and sign a transcribed report. With a transcription service, this is an inherently inefficient process, as the radiologist must keep checking for his or her reports to sign.

Ed Zaragoza, MD, the medical director of Santa Monica-UCLA Radiology, says that he feels liberated with the system. "It is such a relief to be able to leave on a Friday afternoon and not to have to worry about logging in from home to sign my reports," Zaragoza says. With this inefficiency eliminated, the turnaround time has decreased exponentially. Among the Multi-Specialty Group, the median turnaround time is 2 hours, and practically all reports are finalized within 24 hours. The improvement in service quality has been noted by referring physicians and is helping UCLA's standing in the very competitive local medical-imaging market.

continued on page 30

spend time trying to track down results from radiology. Phone calls aren't eliminated entirely with referring physicians receiving reports promptly; however, it means that when the phone *does* ring, it's for a reason.

"I don't get phone calls anymore about normal or minimal-abnormality reports," NYU's Litt observes. "I get more focused calls, so I'm spending my time focusing on patients where there's really something significant to talk about; it's much more of a consultative type of relationship."

Radiologists at the UVA Imaging Center have noted a similar change since implementing SR technology.

"I used to get a call about every single CT scan, and that's dropped off a lot," Bassignani recalls. "Now when I get calls, it's a clinician with a question about my report, which is all value-added, because that's my role: I'm a consultant to the clinician."

Reducing clerical duties and increasing the speed of reporting also can help radiologists avoid becoming superseded by technology that makes images available to any

clinician with Internet access.

"It doesn't do a lot of good for me to interpret the study if nobody knows about it," Litt says, expressing the growing concern shared by many radiologists that referring physicians are obtaining their images and moving forward with treatment before the report arrives. "There is value in my being able to interpret the study and communicate those results to another physician so he or she can make whatever management or therapy decisions are required."

Holding Out

No matter how good SR software becomes, it will never be ideal for everyone. To benefit from SR technology, users must be able to speak clearly—heavy accents are not as problematic as poor diction—and they need to be able to follow the standard rules of grammar. Also, ideal candidates are comfortable with computers, but even "computer-phobic" individuals can succeed with SR, just with a bit more coaxing.

"You must have a physician champion,

someone who will keep the faculty involved, letting them know what's coming and keeping them engaged, so that they feel like they're a partner and not that it's happening *to* them," advises Bassignani, who spearheaded the effort to bring Dictaphone's PowerScribe for Radiology to his facility. To build interest in the new software, Bassignani sent his colleagues a series of e-mails to provide some tips and tricks that could be used with the SR program. "When PowerScribe was installed and the doctors attended training, they already had some familiarity with the program and which features they wanted to learn more about."

An Ever-Growing Technology

The success that many radiologists are seeing with SR has caught the attention of others in the medical community, from general practitioners and cardiologists to orthopedists and mental health professionals. But the transition won't be easy, and arriving at the accuracy rates comparable to those experienced by radiologists will take time.

OUTSTANDING ISSUES

Pleased as we are with the progress thus far, the journey has not always been smooth; several challenges, both technological and workflow, remain. Chief among the technological hurdles is the integration of the user interface among the SR application, RIS, and PACS. A lot of work remains to be done to improve its robustness, efficiency, and seamlessness.

At the system level, placing the SR application on the PACS workstation invariably leads to CPU contention between SR and PACS applications, leading to cutouts of the dictation when the user is scrolling through large data sets and dictating at the same time. The radiologist using SR quickly learns to scroll, then dictate. The phenomenon of CPU contention is potentially solved by the increased processing power, and we are actively testing this hypothesis. The inability to upload tables and other, more complex formatting information through the HL7 interfaces also is a problem that will need to be addressed globally at the HL7 level.

Workflow issues arise from functional limitations of the application. For example, dictations are only stored temporarily on the SR system. An addendum cannot be added to a dictation that has been purged, forcing the user to use transcription or self-type directly in the RIS.

One issue we had not anticipated was ergonomics. In general, users were unimpressed with the design of the hand microphone and the need to hold it close to the mouth over long periods of time, which started generating complaints of upper extremity nerve impingement syndromes. We experimented with a headset and a foot pedal to replace the microphone. The headset was well received, but the foot pedal was too cumbersome to master. We are now experimenting with an innovative compromise: The radiologist wears a headset and uses the hand microphone, instead of the foot pedal, to navigate through the SR application. This allows the user to rest the arm in a relaxed position while holding the microphone.

What has not been a real problem is the recognition rate. Most radiologists achieved excellent recognition rates of consistently more than 95%. The caveat is that they must be willing to put in the time to read the training scripts and train the system during their adaptation phase. Sometimes, they must modify their speech patterns somewhat. Radiologists with an accent *do* have to put in more

effort, but many of them have been surprised at how well the system recognizes their dictation. As Zaragoza is fond of saying: "Speech recognition can be likened to parenting. If the parent spends the time to nurture and teach a growing child, then the results can be spectacular." Success in SR truly is a matter of patience and practice.

LESSON LEARNED

SR is a "disruptive" technology. New skills must be learned and workflows readjusted, and the radiologist's traditional workflow is disrupted in the process. The good news is that it can be done. The reward is a quantum leap forward in the control of the authoring process and the speed of report turnaround time.

Three key factors are mandatory for success:

- 1) **Integration.** SR is a mature product on its own, but without proper integration with the RIS and PACS, it is not possible to realize productivity gains. Another aspect of that integration, which is frequently overlooked, is the human-machine integration. Vendors need to pay more attention to ergonomics.
- 2) **Training.** New skills are best mastered through rehearsed, repetitive use. It is crucial that enough trainer resources be allocated to support the novice user through the critical start-up learning period.
- 3) **Flexibility.** Workflows will need to be changed to either circumvent functional limitations of the system or to leverage its strengths. Everyone must be open to new processes and be innovative.

In closing, SR *is* the future. Its benefits of decreased clinical report turnaround time and shortened billing cycles are too compelling to be ignored. Implementing SR successfully is challenging on a site-specific basis, because of the many potential challenges. Still, those who fail to adapt to this new technology will most likely be left behind in the competitive marketplace. ■

Hwa Kho, PhD, MBA, is director of imaging informatics at UCLA Medical Center (Los Angeles) and is a member of the Medical Imaging Editorial Advisory Board.

“As you enter a much more complex, expanded domain—general medicine being a great example—where you can be talking about many different body systems and in different ways,” Philips’ van Terheyden says, “it’s much more challenging. Speech recognition is a statistical process, and we improve that statistical model with more data. So as we accrue more data and go through a process of applying corrections, those models get refined and start to become very accurate.”

Charting and other required documentation presents a unique challenge. Simply converting voice to text isn’t enough when dealing with an entire patient history and care record. Currently, Philips is fine-tuning a program that allows physicians to dictate “freestyle” while the software analyzes the meaning of his or her speech, assigns value to what is said, and automatically populates the correct section of the medical report. Dictaphone also is tackling this challenge with what it calls natural language processing (NLP).

“NLP reads text and understands it while attempting to deal with the ambiguities of language, noting, for example, the difference between having been prescribed a medication and being allergic to it,” Fallati says. “The data extraction is tuned to the top four

most sought-after pieces of information by caregivers: medications, allergies, procedures, and problems.”

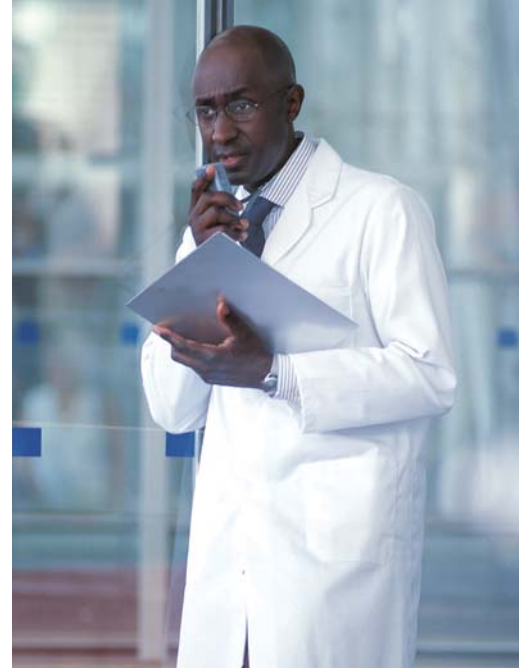
Both systems operate on a similar philosophy: Words without meaning are useless in today’s medical environment. Assigning contextual value to content and designing a program to understand the physician’s intended meaning by discerning the fine distinctions inherent to spoken communication moves the idea of easily accessible patient information much closer to reality.

“It represents the transition from interest in technology to real, useful, valuable support of clinical activity, which is essential if we’re going to deliver high-quality care,” says van Terheyden.

Clinicians already sold on SR technology are eager for such advances and have a detailed wish list for future SR programs—including the addition of drafting-type features so that attending physicians could educate residents by returning “edited” studies; SR-generated files with automatically assigned ICD-9 coding; and SR systems tightly integrated with the RIS/PACS, allowing it to offer such features as automatically loading previous study images or providing one-click access to the patient’s medical record.

Without a doubt, these types of fundamental shifts in the aim of SR technology will prove to be a daunting and time-consuming task—but one that manufacturers believe is realistic and most likely just around the corner. **MI**

Dana Hinesly is a contributing writer for Medical Imaging.



We Sell all Makes and Models of Ultrasounds

Huge Inventory of Parts!

GE, ATL, Acuson, Siemens, HP and more

Ultrasound Service Specialists

- Full Service/PM Contracts Available
- Ultrasound Training Courses Available
- Transducers, Circuit Boards
- Monitors, Cameras & Printers

INTERMED Inc.

Technology Management

CORPORATE OFFICE: 13351 Progress Blvd.
Alachua, Florida 32615

Toll Free: 800.805.5754
Tel. 386.462.5220

www.intermed1.com
sales@intermed1.com



Are you receiving your own free subscription to *Medical Imaging*?



Radiologists, technologists, and other imaging professionals involved in the selection, integration, and management of medical imaging technology and services rely on *Medical Imaging* for the latest news, issues, and trends.

Subscribe or renew online at
www.submag.com/sub/mc?pk=hse

what's good for the goose ...

Radiologists aren't the only ones putting speech recognition (SR) technology through its paces. The Logiq 9 from GE Healthcare (Waukesha, Wis)—employed by more than 12 sonographers at Baptist Memorial Hospital-DeSoto (Southaven, Miss)—boasts the option of using voice commands to control the unit during an exam.

"As with anything, there's a learning curve; but after that, it really sped up our workflow, particularly when working with patients where you need both hands," explains Vicki Pyles, a senior sonographer at Baptist Memorial. "For example, when doing a venous Doppler exam, you can actually leave the patient's side—and the machine—to go to the foot of the bed and work, which really helps."

Performing ultrasound exams on patients can be physically demanding work. Bending over, crouching down, and reaching across beds to reach patients—all while keeping one hand on the control board—can take its toll.

"When I use the voice commands, it means I'm not reaching for things on the machine, so I can get comfortable without reaching back and forth," says Cindy Owen, a diagnostic ultrasound services consultant at Baptist Memorial. "I've been scanning for more than 20 years, and I have problems with my neck and my back. I think [speech recognition] is going to allow me to stay in the field, because I don't have to stress my body as much."

A PERFECT FIT

Not only does SR free users to move around the patient, but it gives them a hand—literally.

"It is really helpful to have your hands free when you're with the patient, because you're adding gel and all kinds of things with the other hand," Owen says. "So, using voice commands allows you to truly multitask, especially in the neonatal nursery, where babies are not going to hold still for you; you need one hand to hold the baby and the other to hold the probe—and you don't have another hand to work the machine."

Neonatal intensive care isn't the only department where sonographers benefit from hands-free operation. A patient's bedside is often prime real estate where a host of medical apparatuses, such as IV poles and heart monitors, fight for space. When performing exams in patient rooms, trying to squeeze in one more piece of equipment can be impossible in some cases and vexing in most. Using voice commands means sonographers don't have to worry about "fitting in"—as long as the probe reaches the patient, they can do their job.

REAL-WORLD EXPERIENCE

As one of GE Healthcare's clinical sites, the Baptist Memorial staff actually helped refine the Logiq 9's SR software, realizing that not all commands are created equal.

"When we first started testing the system, we realized our southern slang was confusing the

system—in the South, words that should be one syllable become two," Pyles laughs, giving the word "freeze" as an example. "The word becomes too long, and the machine wouldn't understand it, which is a problem, because that's one of the most common commands."

GE Healthcare's engineers tackled the dilemma by adding options. The first fix was programming "stop" as an alternative word for "freeze." But the concept was adopted, and now, many of the controls have more than one command.

The Baptist Memorial team continues to work closely with GE Healthcare's developers, providing suggestions for improving the system's practical application, such as less-cumbersome cursor navigation and the ability to add new commands and words to the system.

COMFORT EQUALS CARE

The ability to control the ultrasound system with their voices has not only improved staff member's workflow and working conditions, but it also has made an impression on patients.

"I have better rapport with the patient, because I'm not reaching for things on the machine," Owen notes. "Patients also like the 'high-tech' aspect and feel like they're really being scanned on a top-of-the-line system, which gives them even more confidence that they're getting a quality exam." —DH



INTERNATIONAL

Call in the Specialist

Power Protection Specialists

- Over Three Decades of Excellence in TVSS
- Superior Support and Service
- Commitment to Quality

Contact LEA today and ask to speak with a medical imaging specialist!

1-800-654-8087

www.leaintl.com

lea@leaintl.com

smthhs



Please visit www.medicalimagingmag.com/advertisers