

A National Web Conference on Improving Health IT Safety Through the Use of Natural Language Processing to Improve Accuracy of EHR Documentation

Questions and Answers

February 7, 2017

QUESTION:

Did any nurse practitioners or physician's assistants participate in your work?

ANSWER:

Thomas Payne:

No, but that's the intention over time. We limited it to internal medicine physicians writing progress notes because of the size of the study and the need to complete it all within 2 years.

QUESTION:

Do you anticipate selling your software via Cerner's SMART on Fast Healthcare Interoperability Resources (FHIR)?

ANSWER:

Thomas Payne:

We aren't in the software marketing or development business, but the commercial world is already looking into approaches similar to this. The advantage of using this approach is that we can leverage what the commercial world has produced. For example, our speech recognition engine was a commercial product—Dragon Naturally Speaking (the Medical Practice version)—and we used the Cerner electronic health record (EHR). However, we put it together in a different way, and it gave us the flexibility to add features that our physicians wanted, such as a lot of attention to formatting. The ability to extract data using FHIR through HL7 is a really intriguing way to extend this.

QUESTION:

Did you use a Citrix platform?

ANSWER:

Thomas Payne:

Yes. However, when physicians use voice-generated enhanced electronic note system (VGEENS) at the bedside, all they're using is a smartphone with the cell phone features disabled. However, when they edit and sign the note, it's on the EHR delivered to Citrix.

QUESTION:

You mentioned the bimodal distribution that you saw. Could that be due to new patients being admitted later in the day, giving you later-in-the-day usage?

ANSWER:

Thomas Payne:

That bimodal distribution on that slide displays the time between when the patient is seen on rounds—whether that's morning or afternoon—and the time that the note is available for others to see. So that particular slide would not be different, whether the patient was seen in the morning or the afternoon.

QUESTION:

Over the course of your project, you saw a dip in quality improvement and meaningful use measures. Did everything make it to the right box to be gathered into the reports that were needed for those quality improvement and meaningful use reports?

ANSWER:

Thomas Payne:

Yes. This didn't affect our quality improvement or our compliance efforts. In fact, we're still examining whether the intervention had an effect on billing. I suspect that it would enhance billing, since it's a more accurate, complete note. We have software that automatically extracts problem concepts for clinicians to choose from. So, if anything, I would expect the effect to be neutral on compliance with meaningful use.

QUESTION:

How did the physicians and patients feel during transcription—when the information was being recorded in the room?

ANSWER:

Thomas Payne:

Many physicians create their notes in the room with the patient; I do this sometimes myself. The advantage is that if I say something in the notes that is not correct, the patient is there to correct it. I didn't gather data on the reaction of patients in the study, but my sense is that we're moving toward greater involvement of patients in their notes.

QUESTION:

Are there data to show that completely new note text is more accurate and complete than reused text that has been reviewed for accuracy?

ANSWER:

Li Zhou:

We haven't analyzed the document that was not dictated by the voice recognition or transcription services. In a future study, we may look at physicians' use of templates or text, notes they can use copy-paste functions, including their old information that should be updated.

QUESTION:

There are commercial speech recognition products (like M*Modal) that provide immediate decision support regarding the completeness of documentation and allow nurses to scale large volumes of notes to search for specific terms like "sepsis." Do either of you want to comment on these types of products?

ANSWERS:

Thomas Payne:

That market is focused on one particular area of decision support. It demonstrates the potential to improve what's written and what's done for patients by analyzing the content of the note in near real-time. I'm interested in helping with differential diagnosis or management decisions, and clinical decision support (CDS) delivery at the point of care based on natural language processing (NLP) analysis of a note gives us new opportunities.

Li Zhou:

There are ongoing efforts to look at problems with notes that were not on the structured list, such as problem lists or medication lists. Speech recognition can pick up those diagnoses and medications for medication reconciliation and make program lists more accurate and complete.

QUESTION:

For medical documentation NLP research, do you have any recommendations for note annotation tools that flag differences or topics, for instance?

ANSWERS:

Thomas Payne:

Most physicians would be grateful for assistance with making sure their note is accurate and complete, and annotation for that purpose would be well regarded. It would have to be highly accurate, because physicians have a low tolerance for annotation or decision support that they see as not adding value. However, I don't have personal experience with those tools.

Li Zhou:

This is a very challenging usage problem. When transcriptionists don't recognize something in the notes, they use a question mark to flag it so that the physician can supply the missing information. We saw a lot of question marks. For those types of things, we could automatically annotate the notes and ask physicians to fill in the information.

In doing our analysis, we found that most of the errors were general user errors or stop word errors. These may not be clinically significant. We should put more effort into errors that are related to clinical information, such as medication errors and diagnosis errors.

QUESTION:

How would you compare the length and complexity of notes with regard to VGEENS versus the control?

ANSWER:**Thomas Payne:**

We're doing that in a different study. I don't have the results yet, but I suspect VGEENS notes will be shown to be shorter. Most of my colleagues regard that as a good thing. As a physician who looks at a lot of notes, I find it harder to understand what really happened and what the physician's thinking is in a very long note. A shorter note that's more to the point allows me to understand what the problem is with that patient and what we're doing about it more succinctly.

QUESTION:

Will speech recognition data entry ever reach a tipping point in which it becomes the dominant data entry process used by clinicians? If so, when do you predict this will happen?

ANSWERS:**Tomas Payne**

In the outpatient world, we may have reached that tipping point. In the inpatient world, interactive voice recognition isn't as good a fit for rounding workflow, so we haven't reached that point yet.

Li Zhou:

In the past decades, users of speech recognition technology have been mostly radiologists or emergency department physicians, and now inpatient physicians are using it too. At Brigham and Women's Hospital, all types of clinicians are adopting this technology. Once the technology gets more mature and people know how to use it, it will make a big change.

QUESTION:

When you counted insertion or deletion errors, did you compare the transcript to the original voice file to confirm that the words that were added or deleted were originally spoken?

ANSWER:

Li Zhou:

Yes. We compared the audio file and the text generated by the speech recognition engine.

QUESTION:

How prevalent are the following flow processes in current practice: 1) non-interactive speech recognition dictation, 2) review of a speech recognition-generated document by a transcriptionist-like profession, and 3) corrected documentation by the clinician?

ANSWERS:

Tomas Payne

In the clinic, the most common use of voice recognition is interactive. In the hospital, voice recognition is most often done non-interactively. Almost every note that is dictated these days is processed by a speech recognition engine, and then the transcriptionist corrects it. So that would be number 2. The third option, where it's non-interactive and the clinician corrects it, is very new here. It's an underutilized and potentially very useful mode for workflows, such as inpatient progress notes.

Li Zhou:

A physician's preference for which way to use speech recognition technology is probably based on their background—whether they're a native English speaker or, like me, a non-native speaker. When considering which mode to use, institutions also need to consider the cost of traditional transcription services versus speech recognition software that clinicians can dictate to directly.

QUESTION:

Do you have any data on the extra time spent for correcting speech recognition-only notes by physicians?

ANSWERS:

Li Zhou:

There is not much literature about the difference in time. We are currently observing clinicians dictating notes using speech recognition software in their practice; we look at the time they spend dictating versus the time they spend correcting. I heard from physicians that they spend a

large amount of time editing the notes after dictating them. It varies from clinician to clinician as well.

Thomas Payne:

One of our computer science graduate students is assisting with the editing process. Much of the editing is pretty mechanical, although some of it requires clinical judgment. He and his faculty advisor are working on methods to reduce the burden of editing for physicians.

QUESTION:

Were any of the dictated data converted to discrete data within the EHR, versus remaining as text in a dictated note?

ANSWER:

Thomas Payne:

Previously, we worked on a method that is now commercially available to process a note and to identify candidates for probabilistic entries. This method is up to the physician and outside of our study, but it's amazingly good. It includes negation and probability. It considers the section of the note in which a phrase appears (e.g., "family history" versus "inpatient," "his" versus "herself").

QUESTION:

Is the MTERMS NLP tool open source for download?

ANSWER:

Thomas Payne:

It's not open source yet, but we're working on that.

QUESTION:

The questioner noticed that words like "uh" and "um" were not transcribed as spoken. Does the system automatically exclude these types of words?

ANSWER:

Thomas Payne:

Yes. We use Dragon for this, and it does a remarkably good job of removing those sorts of phrases.

QUESTION:

Was the deployment model used for VGEENS internal for the hospital network, or developed for a mobile device?

ANSWER:

Thomas Payne:

Slide 23 shows the deployment model that is used at the bedside. This is an inexpensive cell phone with an app we developed and the buttons you see there. It was designed to be very simple. We don't expect anything more complex to fit into the busy world of hospital rounding. We'll have a paper describing this.

QUESTION:

How exactly does VGEENS combine voice recognition and transcription with the EMR? Are they integrated or side by side? The questioner assumes that the physician is checking the final note before sending it out. Were there any studies comparing notes before and after checking?

ANSWER:

Thomas Payne:

There are several levels of integration with a commercial EHR. The note that's produced by VGEENS is transmitted to the EHR inbox, where it is edited by the physician before it is signed—just as it would be if it were a dictated note. VGEENS also extracts data from that patient's record in response to voice commands. So if you say, "Please insert vital signs," VGEENS will extract the vital signs from that day, format them nicely, and put them into the note.

QUESTION:

In addition to including ICD codes and structured data in the note, is VGEENS able to include any additional structured data, like LOINC codes, RxNorm, or SNOWMED?

ANSWER:

Thomas Payne:

We represent the problem list entries using SNOWMED CT rather than ICD, since we have evidence that SNOWMED better represents the concepts used in problem lists better than ICD—whether it's 8, 9, 10, or 11. The same applies to LOINC and other nosologies. That gives our computing systems the potential to help us with advice that might be very useful for patients once the content of the narrative text is "understood" by the competing systems.

QUESTION:

What role do you see for VGEENS and other types of reports, such as discharge summaries?

ANSWER:

Thomas Payne:

We didn't intend for physicians to use VGEENS for those types of reports, but they did anyway. VGEENS is well suited to those purposes, because those documents are common and can be time consuming to create. For instance, admission notes include patient information from the history and physical, as well as outside records. Using something like VGEENS with very rapid turnaround—5 minutes—would help us summarize those voluminous records in a way that other people caring for the patient can then read and understand. There are a few mechanical things we would have to change, but that could be done.

Additional Q&A Addressed Following the Webinar

QUESTION:

What feedback did you receive from the group using VGEENS?

ANSWER:

Thomas Payne:

House staff were unfamiliar with dictation, and transcription errors were a problem. We wanted VGEENS for admission and discharge notes, questions about billing, request for template to avoid need to dictate headings, and many others.

QUESTION:

Did your follow-up user survey assess hesitancy by physicians to delay rounds by dictating at the bedside?

ANSWER:

Thomas Payne:

Some didn't like the possibility of delaying rounds. Our response is they can use VGEENS when pre-rounding if they are on a house staff team. Attendings working on their own did not have this problem.

QUESTION:

Do you plan on testing and comparing natural language processing efficacy with other clinicians, like nurse practitioners?

ANSWER:

Thomas Payne:

We would love to. This was a preliminary, limited study.

QUESTION:

User satisfaction for both groups was alarmingly low (<50 percent). Was end users' input taken into consideration for both groups? If not, why not?

ANSWER:

Thomas Payne:

Yes, but only after we received it. This was a very short study.

QUESTION:

On slide #18 (with the diagram and Specific Aim 1), is confidentiality (HIPAA) a concern when the patient is in a two-bed patient room and/or when dictated immediately after leaving the room? It seems like the physician would be required to find a private area.

ANSWER:

Thomas Payne:

This history is being taken in the same room. Ideally, there would be private beds, but that is not the case in many hospitals. Everyone tries to treat information as privately as possible.

QUESTION:

How does VGEENS compare with the Dragon Medical dictation system? Does the system need to be trained as extensively as Dragon?

ANSWER:

Thomas Payne:

VGEENS uses Dragon Medical as the automated speech recognition system. However, it goes beyond Dragon to format, include patient data on command, and integrate with the EHR.

QUESTION:

Were any of the dictated data converted to discrete data within the EHR vs. remaining as text in a dictated note in both scenarios?

ANSWER:

Thomas Payne:

Users have the option to encode problem list data into SNOMED CT, just as they do for any other note. This is true for both the intervention group and the control group.

QUESTION:

What are some ways to address the increasing security measures that compound worsening bandwidth saturation? Simple utilization of numerous speech nodes is inadequate.

ANSWER:

Thomas Payne:

In our project, the VGEENS system used encryption of transmitted voice files via https. Our application had no appreciable effect on the bandwidth of our hospital's 802.x wireless network.

QUESTION:

What are some ways to address autonomic stress-induced voice profile changes and the curse of phonetic content dimensionality?

ANSWER:

Thomas Payne:

We did not address or study stress-induced voice profile changes. We used training as recommended by Dragon Medical, the automatic speech recognition tool we used in our VGEENS system.

QUESTION:

In your opinion, how soon will patients have access to their progress notes for informed choice making and error correction?

ANSWER:

Thomas Payne:

Currently, in some institutions, certain types of notes are open to patients, so they might become a source to correct the errors. I am not aware of which institutions share their progress notes with patients.

Millions of Americans, including patients in UW Medicine, are now able to view their outpatient progress notes. See opennotes.org for more information. Viewing inpatient progress notes is less common but is occurring at New York Presbyterian, UCSF, and other sites, I believe. This is becoming more common.

QUESTION:

Can discrete data for registries be generated with speech recognition notes?

ANSWER:

Thomas Payne:

Yes. The field of NLP is growing rapidly, and permits assigning codes such as SNOMED CT to words and phrases within notes. Several articles can be found using the following PMIDs in Medline: PMID: 26687964, PMID: 23354284, PMID: 22195225, PMID: 18999050, and PMID: 22375458.

QUESTION:

How can we avoid proliferating multi-varied error checking, as we have many EHRs?

ANSWER:

Li Zhou:

It is a challenging problem. Hopefully by studying the errors, we can find more efficient ways to improve the accuracy.

QUESTION:

Did the study account for physicians' using scribes for dictation?

ANSWER:

Li Zhou:

No. We did not include physicians who use scribes.

QUESTION:

Can you identify any available NLP tools for physicians' notes?

ANSWER:

Li Zhou:

There are open source NLP toolkits and medical NLP tools (e.g., cTAKES). Our NLP tool is called MTERMS, and our team works with some physicians in our institution for handling free-text data.