healthcare informatics

A ‘Google’ for Clinical Notes Draws Interest

By David Raths

Those of us who cover healthcare informatics often hear clinicians and researchers talk about the problems involved in doing analytics or research on unstructured data in clinical notes. That was why I was intrigued when I saw that informatics teams at the University of North Carolina School of Medicine are implementing a tool called EMERSE (Electronic Medical Record Search Engine), which allows users to search free-text clinical notes from the electronic health record (EHR). They describe it as being like "Google" for clinical notes.

But then I noticed that the tool was actually created quite a while ago, in 2005, at the University of Michigan, and has been in use there ever since. So I reached out to its developer, David Hanauer, M.D., a clinical associate professor of pediatrics and communicable diseases at the University of Michigan Medical School. He also serves as assistant director for clinical informatics in UM’s Comprehensive Cancer Center’s Informatics Core as well as associate chief medical information officer at the UM Medical Center.

Hanauer told me that the developers of EMERSE at Michigan have a grant from the National Cancer Institute to further develop the tool and help disseminate it, with a focus on cancer centers around the country. “We are about one year into the grant,” he said. “We have spent the last year cleaning up the infrastructure to make it even easier for people to adopt. We have been working hard on technical documentation. When we started it, we had almost no documentation; now we have substantial and detailed documentation about how to implement and run it.”

The five sites implementing EMERSE as part of the grant are the University of North Carolina, University of Kentucky, University of Cincinnati, Case Western Reserve University and Columbia University.

I asked Hanauer if health systems continue to struggle with unstructured data in clinical notes. “They all absolutely struggle with it,” he said. “They have mostly been ignoring it, to tell you the truth. That is why we believe and hope EMERSE will fit well into this environment of people needing different tools.”

I also asked him to describe some of the use cases. Most generically, anybody who needs to look through the chart and doesn’t know exactly where to look can get benefit from it, he said. He described three categories of users: research, clinical care and operations. “For example, in research you could use it for cohort identification. You want to find patients who meet your needs when it comes to a research study. This is important in part because ICD codes, the go-to way people often try to identify a cohort, are often inaccurate and non-specific.”

According to the EMERSE web site, for studies in which eligibility determination is complex and may rely on data only captured within the free text portion of documents, EMERSE can be a rapid way to check for mentions of inclusion/exclusion criteria.

In another example, EMERSE also can be used to help find details about a patient rapidly, even during a clinical visit. “For example, if a patient mentions that a certain medication helped their migraine three years ago but can’t remember the name, just search the chart for 'migraine' and find that note within seconds,” the web site notes. Cancer registrars can use EMERSE for data abstraction tasks, including difficult-to-find information such as genetic and biomarker testing.

Hanauer said at Michigan, clinicians have a way to access EMERSE from their Epic EHR. “If you have a patient’s record open, you can click a button, it will log you into EMERSE and bring that patient’s context over, and you can start searching in just of a few seconds.”

In 2005, the platform was written to work with a homegrown EHR. When UM transitioned to Epic in 2012, Hanauer and team used that as an opportunity to make it more powerful. “When we went live with Epic, it became clear there were some architectural limitations that were probably going to limit the future power of the software,” he recalled. “We leveraged the design and concepts and rewrote it from scratch. But even though we were going to work with Epic, we designed it specifically so it would not be tied to any particular EHR.”

Because it deals with patient records, security and audit logs have to be taken very seriously. Every time you log into EMERSE, you come to an attestation page. “You have to declare why you are using it for this session,” Hanauer explained. “We have tried to make it as simple as possible. Almost every institution that does research now has an electronic IRB system, so we have a way you can pull a user’s IRB-approved study into the EMERSE database, and a list appears of that user’s studies only. The user can click on it, record that use, and move forward.” There also are quick buttons for common administrative use cases.

I asked Hanauer if other academic medical centers had developed similar search tools. He said some have created local tools. “The main difference with EMERSE is that it is proven it can work elsewhere. (It was used at the VA in Ann Arbor, Mich., on the VistA system.) “We have a long track record of use and have been working on the infrastructure to disseminate it,” he said. “We are giving it away at no cost, but it is almost like running a software company, where you have to have a web site, user documentation, and system administrator documentation. To me, it doesn’t make a lot of sense for others to reinvent the wheel when this is something we have invested millions of dollars in at this point.”

He stressed that although the grant project is focused on five cancer centers, they are giving the software away at no cost, and are glad to help anybody interested in getting it up and running. “One of the key challenges is that the users can’t control whether it gets deployed or not,” he said. “Our biggest challenges is not the users, who are contacting us and asking us for it, but getting this through local IT leadership, and that is a big hurdle.”

Why would CIOs be opposed to deploying this tool? “I think their plates are full and a lot of times people are looking for vendor solutions,” he surmised. “I also think that often people don’t understand what the issues are. Some people think they will just get some off-the-shelf NLP software. But I can assure you that that software will not be able to do the kinds of things that EMERSE can do. That is partly because a lot of medical documents are not in natural language. Medical documents are anything but. They are a mess.”
