**PRINT-Mapping the future for early diagnosis of rare diseases**

Rare diseases might seem individually uncommon but collectively 1 in 10 people across the globe have a rare disease according to the World Health Organization. Undiagnosed rare diseases pose a significant challenge in health care, often leading to delayed treatment and poorer patient outcomes.

When specialists or primary care providers are confounded by a patient's symptoms, it can send them back to their medical books. But, Adam Cross, MD, a pediatric hospitalist and clinical informaticist who leads the Children's Innovation Lab at the Jump Trading Simulation & Education Center, is collaborating on research to better help primary care providers make a rare disease diagnosis quickly. The goal is to create an AI-backed, point-of-care support tool that can prompt testing for confirmation for the most relevant diagnoses.

Dr. Cross says typically genetic testing is the only way to confirm a rare disease. That testing is expensive and not readily available in primary care settings, particularly in rural communities. Patients have long waits and often must travel in hopes of getting a diagnosis.

A Jump ARCHES grant is funding a collaborative effort with the University of Illinois Urbana-Champaign (UIUC), with co-lead researcher Jimeng Sun, PhD, to use generative AI and medical knowledge graphs to improve diagnosis of rare diseases.

"We are trying to create a model that will take the notes that are already being written by primary care physicians and look for unique sets of signs and symptoms that might suggest a patient has an undiagnosed rare disease." Dr. Cross adds, "That approach can empower the physician with information regarding those diseases and prompt further testing if they choose to do so."

Dr. Cross believes rare diseases, despite their complexity, leave identifiable patterns within electronic medical records. The project introduces Automated Rare Disease Mining (AutoRD) to help unearth patterns when analyzed using machine learning models and mapped using knowledge graphs. He believes the Auto RD method can potentially enable the early detection of these diseases. These patterns could be drawn from clinical signs and symptoms and medical history, in addition to demographic features such as race/ethnicity, age and geographic location.

In one year, researchers have been able to develop robust machine learning models to extract signs and symptoms of rare diseases from notes within electronic medical records that have been stripped of personal, identifying information. They've also been constructing comprehensive knowledge graphs to visually map the relationships between various medical conditions and symptoms of rare diseases from existing resources.

"We did this so that we could give people a better visual example of just how complex these diseases are and how they're all connected," Dr. Cross observes while viewing an early version of a knowledge graph created by researchers that looks like a celestial constellation. "It's kind of beautiful when you see the simplicity and the complexity of all the different symptoms, but the simplicity of the structure. It gives you a sense of just how incredible these new technologies really are in terms of helping us find patterns in the chaos."

The effort involves sophisticated machine learning approaches that have changed with lightning speed, even since Dr. Cross and his UIUC colleagues began their work.

"Even in that year, there have been more advanced and more powerful methods that have come out that we've since adopted, and we continue to improve our own methods based on what's being developed globally."

The research is leveraging the abilities of a supercomputer at UIUC which has a federal health privacy-compliant section for secure data storage and processing of massive amounts of health care data. The next focus is to work on the predictive ability of the models. Dr. Cross is convinced that dynamically linking patient-specific data with the broader context of medical knowledge on rare diseases can uncover patterns and correlations that might otherwise remain hidden.

As the approach is refined and improved, he also doesn’t rule out the possibility of discovering new rare diseases.

“As we go along, it is certainly possible with the technology that we've created, maybe we'll start to see patterns of signs and symptoms that don't have a known diagnosis. Or maybe there are subtypes of diseases that are out there that really seem to have their own unique progression, and we could also maybe discover those in subsequent phases of the project.”

Dr. Cross, who also serves as an assistant professor at the University of Illinois College of Medicine in Peoria (UICOMP) and is an adjunct professor at UIUC, says he and his fellow researchers have already published one scientific paper, and another has been submitted, showing the results of their research so far.