Artificial intelligence making progress in collaborative projects across Canada

Groups are working to build a healthcare system that continuously learns about the population it serves.

BY DIANNE DANIEL

If health data is the next frontier in personalized medicine, the learning health system could be the vehicle to transport us there. That's the vision of recent collaborations across Canada bringing together academics, clinicians, businesses and funding partners to unlock critical health data found in physician notes, imaging databases, electronic medical records and other digital health systems. Not only are the groups enabling researchers and clinicians to access high-quality pools of patient information that have historically been difficult to apply to decision making, they're also working to build a healthcare system that continuously learns about the population it serves.

"What we're doing is building a system so that we'll be prospective, by which I mean that every day the system will be newly updated," said Jeremy Petch, PhD, director of Digital Health Innovation at Hamilton Health Sciences (HHS) in Ontario and founder of the hospital's Centre for Data Science and Digital Health (CREATE). "As every new patient comes in and is diagnosed, we'll be able to access their data," said Dr. Petch.

In collaboration with Dr. Mark Levine, the Buffet Taylor chair in breast cancer research at McMaster University, the team at Escarpment Cancer Research Institute (ECRI) and the Toronto-based artificial intelligence (AI) firm Pentavere, CREATE – which is staffed by experts in artificial intelligence, data sciences and software engineering – is working to digitally recreate patient 'stories' for breast cancer patients currently being treated at the HHS Juravinski Cancer Centre. The platform builds on an earlier pilot of 50 patients that used IBM's Watson Health. Moving forward, the initiative is relying on Pentavere's proprietary natural language processing (NLP) engine called DARWEN to extract real world patient evidence and insights from vast amounts of unstructured clinical data.

The impetus for the project came from Dr. Levine, who, as he was winding down his highly recognized career as a medical oncologist at HHS, had the notion that AI could solve a fundamental challenge in healthcare – how to obtain a clear picture of a patient population before, during and after treatment.

"Thirty-five years ago if a patient asked me, 'How many patients with breast cancer do you see with the same problem as me every year and how do they do?' ... I would either do a chart review myself or ask a resident to do it," said Dr. Levine. "Do you know what we do now, 35 years later? The same thing, and with all of the technology now available, that's unbelievable."

The group's initial pilot demonstrated the ability of AI to accurately identify key information from multiple digital sources for 50 stage three breast cancer patients treated at Juravinski. The current effort, funded in part by Roche Pharmaceuticals and now scaled to include 3,000 breast cancer patients, is using Pentavere's technology to generate similar anonymized information that includes the date of diagnosis, how the cancer was diagnosed, the pathology of the tumour, the data and type of surgery, whether or not chemotherapy or radiation were administered, and the outcome.

"All of this data is very siloed, spread over multiple different source systems, so really the task initially is to bring that data all together in its raw unstructured form, into a data lake – it's still messy, it's still unstructured, it still has conflicting data in it," explained Pentavere medical director Dr. Christopher Pettengell. "Our technology can then be deployed within that environment and structure it to make sense of it."

The idea is to build an automated function into the HHS health system workflow so that disparate health data can be grouped into high-quality row and data sets, preparing it for analysis. Machine learning algorithms are then trained to generate accurate, de-identified patient information from the data, giving clinicians a first-of-its-kind view into outcome data specific to Juravinski cancer patients.

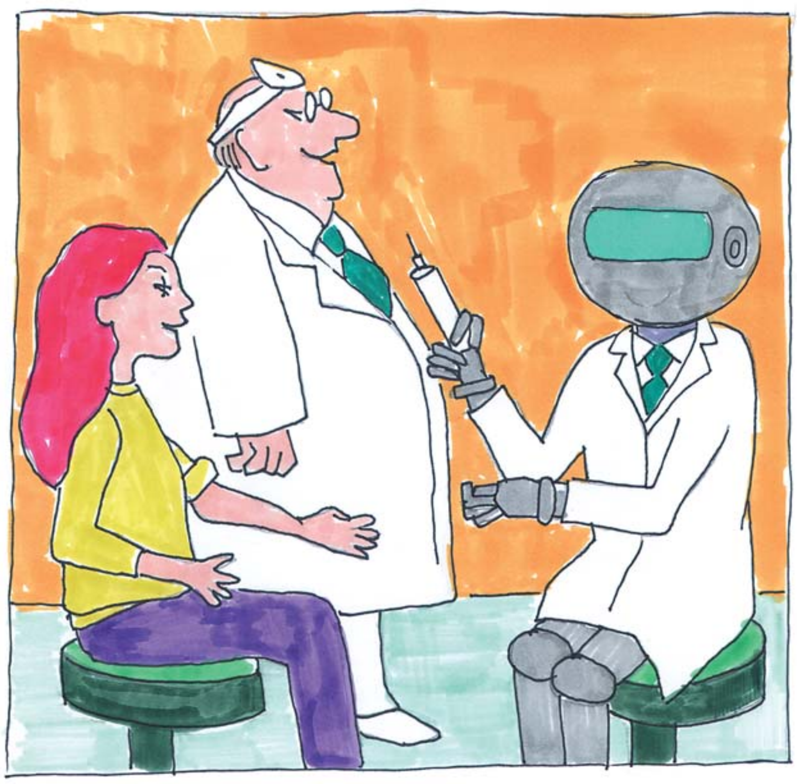

ILLUSTRATION: LINDA WEISS

"The analytics and the statistics on the data – that's something we know how to do, but we've just never really been able to access the data before because it's been locked away, and there's been no efficient way to get that data out and make it available to the statisticians and analysts," said Petch.

Though he's careful to avoid the 'hype' around the promise of what AI can deliver, Petch said the overall ambition is to better understand the cancer clinic's patient population so that the hospital can move towards a personalized care model, and ultimately make quality improvements. One possible deliverable is a learning health system dashboard utility that would enable treating clinicians to access real-world evidence about similar patient journeys within HHS, as well as recommendations based on leading-edge research or clinical trials, when deciding on the best course of treatment for the patient in front of them.

"What's unique about this is that a clinician can get a picture of their complete practice, and get a profile of their practice over all of the patients they've treated over many years, as well as the patients that their colleagues have treated, to maybe see patterns that we haven't seen before," said Petch. The approach also supports the hospital's mission to address health equity, because it analyses outcomes for entire patient populations regardless of socioeconomic status or ethnicity, removing the bias often associated with clinical trials, he added.

As work on the learning health system moves forward, the next step is to validate that the information generated by the AI model matches the information in the actual patient records. Two leading breast cancers experts, independent of the collaboration, are currently comparing key patient attributes such as histology, stage, estrogen receptor or type of surgery, and the machine learning algorithms will be refined as needed, until the match is 95 per cent or higher, explained Levine.

Another Canadian team working to 'unlock' the benefit of digital health data involves the University of Alberta Faculty of Medicine and Dentistry, Alberta Health Services (AHS) and Indianapolis-based Pinnacle Solutions Inc., a long-time SAS reseller that entered the Canadian analytics market in 2019. This past summer, the faculty launched the Data Analytics Research Core (DARC), a secure high-performance computing and on-premise data storage environment it built in partnership with Pinnacle Solutions to facilitate health data sharing between AHS and faculty researchers.

The goal is to grow the university's capacity to perform big data analysis and AI, advance e-health innovations such as personalized care, and provide student training in health data analytics. At the same time, DARC, which is supported with funding from the AHS Chair in Health Informatics Research, is also available to data analysts funded by the Alberta Strategy for Patient Oriented Research Support Unit who are invested in building a provincial learning health system.

"As medicine really becomes a data science, we're going to need to train a next generation of clinicians who know data, as well as data scientists who know clinical medicine," said vice-dean of research Lawrence Richer, a professor in the faculty's department of pediatrics who led the project. "DARC provides an environment where that can start to happen."

Historically, if a faculty researcher wanted to work with sensitive health data, they would make an individual request to AHS, the provincial custodian of patient information, and agree to work within specified security parameters. Richer found that as more requests were being made, and the sensitivity of the data increased, the custodian's willingness to share decreased. At the same time, researchers interested in applying AI and machine learning require access to much higher computer capacity than is typically available to them in a small server or desktop environment. DARC solves both challenges, he said.

Architected by Pinnacle Solutions, the integrated platform sits behind the AHS firewall at a local hospital, where it shares the hospital's network infrastructure and security protocols. In addition to featuring SAS Viya, a visual AI, analytic and data management tool that makes data analysis as easy as 'dragging and dropping,' it also supports open-source software and other free tools popular among students. The benefit is that both 'clickers and coders' are supported, said Richer.

"Not everyone can be that expert coder, and we can't continuously hire that coder every time we want access to a data set," he said.

Richer experienced the platform's ease of use firsthand when he decided to test a model he had previously requested from a data analyst to predict which children arriving in the emergency department are best served by brain imaging. "What we know is many kids present to emergency with worrisome symptoms, but only a small fraction actually have a worrisome disorder like a stroke, brain bleed or brain tumour," he said. "My hope was that we could build an algorithm to take clinical features and then give the point-of-care clinician a risk score of 'should I, or should I not' do a brain scan."

When DARC went live, Richer was able to test six different iterations of the model in one hour, even though he had never used SAS Viya prior. "I didn't need to wait for an analyst to tell me which model works best. I already have an idea of which is best – I took that as proof that we've done the right thing here," said Richer.

As of October 2021, roughly 60 users had been granted access to DARC, and Richer expects the number to grow. Approximately one-quarter of the faculty's 750 members are involved in some aspect of health data research.

One example is a big data initiative led by Dr. Daniel Baumgart, professor and director of the university's Division of Gastroenterology, who is working in partnership with Crohn's and Colitis Canada. The project aims to use DARC to deliver personalized therapies to Canadians living with inflammatory bowel disease (IBD), and is currently analysing health data from 60,000 Albertans who live with the disease to identify patterns in how they're doing over the long term and what contributes to their outcomes.

The hope is that the AI models will be able to predict which medication will be most effective and at what dose, who is more likely to experience complications related to IBD, and who is at greater risk for future hospitalizations. With AHS recently transferring terabytes of diagnostic imaging data to DARC – representing a first for researchers – the researchers will also work to refine diagnosis with imaging techniques, to assist in more precisely detecting the disease.

"Right now we're limiting access to DARC. We have a select number of teams and the demand is growing," said Richer. "I don't want to sit with a big white elephant that I can't sustain over time, but at the same time, I want to maintain what the core needs are for our core researchers who are truly innovating in the health space."

A third collaboration working to advance the concept of a learning health system in Canada is the new 'junior-senior' Research Chair in Digital Health announced by the Université de Sherbrooke in Quebec and funded by the Ministère de l'Économie et de l'Innovation (MEI). Co-directed by professor Anita Burgun, a French expert in biomedical informatics with ties to the University of Paris, and Christina Khnaisser, a health informatics researcher at Sherbrooke, the chair is overseeing an ambitious research program based on a vast network of France-Quebec collaborations including partnerships between pediatric hospitals and research centres.

The program is using a digital platform called PARS3, developed at the university by the Interdisciplinary Research Group in Health Informatics (GRIIS), to access health data in situ, meaning it doesn't have to be moved to a shared location, alleviating issues related to privacy. The goal is to build a system that can "securely and ethically analyze the data that accumulates during the patient-care trajectory," said Burgun, as they work to better understand, diagnose and treat rare diseases.

For example, medical teams will be able to securely analyze clinical records, medical images and physiological information as it is collected throughout patients' daily lives. Doctors will then be able to compare medical profiles in order to speed up diagnostic decisions and optimize treatment.