Driving the development of new cardiovascular risk prediction models with big data

Jack Tu

The Cardiovascular Health in Ambulatory Care Research Team (CANHEART) big data project was launched in 2012. Under the direction of **Jack Tu**, a senior scientist at ICES and lead of its cardiovascular research program, the multi-year project uses encoded personal identifiers to link multiple population-based databases held at ICES, enabling the project team to analyze population-wide cardiovascular clinical measures and outcomes while taking into account patient sociodemographic characteristics, behavioural and traditional cardiac risk factors, comorbidities, and health services and prescription drug use.

The scale of the data is large, encompassing the entire Ontario adult population of 10 million people dating back to 2008. A 2015 editorial in the journal *Circulation* on CANHEART's work described the size and integration of the data as "impressive."

From this big data, the CANHEART team has developed a Canadian definition of ideal cardiovascular health for adults and youth. Created in 2014 in partnership with the Heart and Stroke Foundation, the CANHEART Health Index measures and monitors the heart health of Canadians against an ideal based on six health behaviours or factors associated with better cardiovascular health.

In the clinical setting, family doctors can use the index when counselling patients about their heart risk. In policy environments, the index provides a baseline for goal setting. For example, the Heart and Stroke Foundation used the index to inform its ambitious 2020 Mission Impact Goals, with an aim of reducing the risk factors for heart disease and stroke by 10% and the death rate from heart disease and stroke by 25% by 2020.

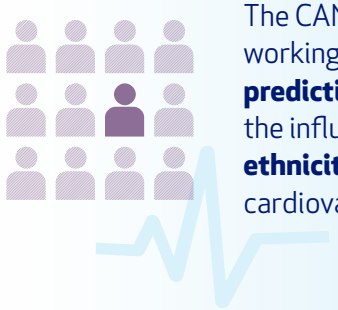
The CANHEART team is working to develop cardiovascular risk prediction models that consider the impact of ethnicity on cardiovascular health, using the robustness of an enhanced CANHEART big data cohort with contributions from the Electronic Medical Record Administrative Data Linked Database (EMRALD) and the Ontario Laboratories Information System (OLIS).

According to Tu, the CANHEART project demonstrates some of the many ways big data will be used in the future to improve cardiovascular risk prediction models that will more accurately identify all patients at risk.

"Our work so far has shown that despite progress in cardiovascular health and health care, there are significant disparities in our understanding of heart health among ethnic populations," explains Tu. "Through our big data multifactorial analyses, we're showing that cardiovascular risk is not a one-size-fits-all proposition. Our team is working to create better ethnic-specific rules to unravel the root causes of the disparities in heart health among Canada's ethnic groups."

Launched in 2012, the **CANHEART big data** initiative conducts multifactorial analyses of cardiovascular risk, outcomes and health system use for Ontario's adult population of **10 million people** dating back to 2008.

The **CANHEART Health Index** was used by the Heart and Stroke Foundation to inform its **2020 Mission Impact Goals**, and is being used by family doctors when counselling patients about their heart risk.



The CANHEART team is now working to develop new **risk prediction models** based on the influence of an individual's **ethnicity** on his or her cardiovascular health.