Researchers use data mining to learn more about uncommon diabetes cases

February 2 2023, by Homa Shalchi

Credit: Pixabay/CC0 Public Domain

In the ongoing research and treatment of diabetes, the focus is typically on the two forms of the disease that dominate public awareness. Type 1 has a stronger genetic component that requires insulin therapy for life; type 2 is frequently associated with obesity and lack of exercise resulting in insulin resistance. Type 2 generally occurs in adulthood and is not associated with the loss of insulin production as type 1.

But another type—a range of often unrecognized diabetes classified as

"atypical"—is gaining greater attention, thanks to the Rare and Atypical Diabetes Network (RADIANT), led by teams at USF Health, Baylor College of Medicine, the University of Chicago and Massachusetts General Hospital.

A new study by USF Health researchers Dr. Jeffrey Krischer, director of the USF Diabetes and Endocrinology Center and the USF Health Informatics Institute, and Dr. Hemang Parikh, associate professor in bioinformatics and biostatistics in the Health Informatics Institute, was recently published in the *Journal of Clinical Endocrinology & Metabolism* in collaboration with Dr. Ashok Balasubramanyam, professor of medicine, endocrinology, diabetes and metabolism at Baylor, Dr. Maria Redondo, professor of pediatrics, diabetes and endocrinology at Baylor, and Dr. Christiane Hampe, from the University of Washington. The study focuses on data mining as a framework for identifying phenotypes of atypical diabetes.

"In addition to type 1 and type 2 diabetes, there is a range of atypical forms of diabetes that affect people who cannot be categorized in the same way," Parikh said. "Some of the time, these people—children and adults—are misdiagnosed and receive different treatment than they should get."

One form of atypical diabetes—monogenic—is due to a single gene mutation. Another type results from a cluster of genetic disorders and can accompany mitochondrial disease. Another is characterized by patients who appear to have type 2 diabetes, yet present with diabetic ketoacidosis, a complication thought to occur only in patients with type 1 diabetes. Yet another affects the manner in which fat is stored.

The new paper furthers the study of these rarer forms of the disease, which cause patients' symptoms and health challenges to differ from those with type 1 and type 2. The analysis was conducted through the

sophisticated process of data mining—digging through data to discover hidden patterns.

Parikh and his team developed a data mining system as part of a program called DiscoverAD (short for Discover Atypical Diabetes). In essence, DiscoverAD relies on a two-step filtering process—first to exclude participants who meet definitions of the typical type 1 diabetes or type 2 diabetes, then to include participants with certain pre-specified atypical diabetes characteristics.

"This is followed by robust analysis to discover novel phenotypes of atypical diabetes (AD) within the filtered group," said Cassandra Remedios, M.S., an assistant in research in bioinformatics in the Health Informatics Institute. "We developed DiscoverAD to permit flexibility and efficiency so it can be applicable to various clinical settings with different types of large cohort datasets."

In the study, two distinct cohorts of patients with diabetes were investigated. The first cohort comprised Hispanic participants with diabetes from the Cameron County Hispanic Cohort led by researchers with the University of Texas Health Sciences Center. The second cohort comprised 758 multiethnic children within the Texas Children's Hospital Registry for New-Onset Type 1 Diabetes (TCHRNO-1) study. Due to the large cohort datasets, a manual review to identify and cluster phenotypes of atypical diabetes would have been extremely time-consuming, Parikh explained.

The study was conducted as part of RADIANT, which is comprised of universities, hospitals and clinics around the United States. Baylor and the University of Chicago are the national centers of the consortium, and USF serves as the data coordinating center for the entire network.

"This work demonstrates the high prevalence of atypical forms of

diabetes in varied populations. The DiscoverAD tool is an innovative and practical tool to identify such patients in different datasets. I believe this could be a foundation for developing criteria that clinicians can use to diagnose their patients with diabetes more accurately and treat them more precisely," Balasubramanyam said.

Patients with atypical diabetes are treated throughout the country, but frequently as isolated, individual cases, and that has made it difficult to amass a base of knowledge that benefits providers and patients. RADIANT addresses that challenge by creating a centralized base of data, information and resources—with the goal of leading to more effective diagnoses and better treatment plans.

"We found in our studies that atypical cases are quite high—comprising about 5% to 11% of diabetes diagnoses," Parikh said. "We also found that many people might have been misdiagnosed as either type 1 or type 2 diabetes."

A key indicator of atypical diabetes is a treatment that does not seem to be working. For instance, some diabetes patients might start losing weight quickly and inexplicably. Others may see glucose levels remain high despite receiving insulin.

"If a person is not responding in a way they should be, that could be a sign," Parikh said.

Several hundred subjects have been involved in the RADIANT study to gain a greater understanding of atypical diabetes through data mining.

"Not only does this demonstrate the potential of personalized medicine, but the analytics also define computable phenotypes that can be generalized to many data mining situations," said Krischer, who also holds the USF Health Endowed Chair in Diabetes Research.

More information: Hemang M Parikh et al, Data Mining Framework for Discovering and Clustering Phenotypes of Atypical Diabetes, *The Journal of Clinical Endocrinology & Metabolism* (2022). DOI: 10.1210/clinem/dgac632