Scientists have developed a machine-learning method that crunches massive amounts of data to help determine which existing medications could improve outcomes in diseases for which they are not prescribed.

The intent of this work is to speed up drug repurposing, which is not a new concept—think Botox injections, first approved to treat crossed eyes and now a migraine treatment and top cosmetic strategy to reduce the appearance of wrinkles.

But getting to those new uses typically involves a mix of serendipity and time-consuming and expensive randomized clinical trials to ensure that a drug deemed effective for one disorder will be useful as a treatment for something else.

The Ohio State University researchers created a framework that combines enormous patient care-related datasets with high-powered computation to arrive at repurposed drug candidates and the estimated effects of those existing medications on a defined set of outcomes.

Though this study focused on proposed repurposing of drugs to prevent heart failure and stroke in patients with coronary artery disease, the framework is flexible—and could be applied to most diseases.

"This work shows how artificial intelligence can be used to 'test' a drug on a patient, and speed up hypothesis generation and potentially speed up a clinical trial," said senior author Ping Zhang, assistant professor of computer science and engineering and biomedical informatics at Ohio State. "But we will never replace the physician—drug decisions will always be made by clinicians."

The research is published today in Nature Machine Intelligence.

Drug repurposing is an attractive pursuit because it could lower the risk associated with safety testing of new medications and dramatically reduce the time it takes to get a drug into the marketplace for clinical use.

Randomized clinical trials are the gold standard for determining a drug's effectiveness against a disease, but Zhang noted that machine learning can account for hundreds—or thousands—of human differences within a large population that could influence how medicine works in the body. These factors, or confounders, ranging from age, sex and race to disease severity and the presence of other illnesses, function as parameters in the deep learning computer algorithm on which the framework is based.

That information comes from "real-world evidence," which is longitudinal observational data about millions of patients captured by electronic medical records or insurance claims and prescription data.

"Real-world data has so many confounders. This is the reason we have to introduce the deep learning algorithm, which can handle multiple parameters,"
said Zhang, who leads the Artificial Intelligence in Medicine Lab and is a core faculty member in the Translational Data Analytics Institute at Ohio State. "If we have hundreds or thousands of confounders, no human being can work with that. So we have to use artificial intelligence to solve the problem.

"We are the first team to introduce use of the deep learning algorithm to handle the real-world data, control for multiple confounders, and emulate clinical trials," Zhang said.

The research team used insurance claims data on nearly 1.2 million heart-disease patients, which provided information on their assigned treatment, disease outcomes and various values for potential confounders. The deep learning algorithm also has the power to take into account the passage of time in each patient's experience—for every visit, prescription and diagnostic test. The model input for drugs is based on their active ingredients.

Applying what is called causal inference theory, the researchers categorized, for the purposes of this analysis, the active drug and placebo patient groups that would be found in a clinical trial. The model tracked patients for two years—and compared their disease status at that end point to whether or not they took medications, which drugs they took and when they started the regimen.

"With causal inference, we can address the problem of having multiple treatments. We don't answer whether drug A or drug B works for this disease or not, but figure out which treatment will have the better performance," Zhang said.

Their hypothesis: that the model would identify drugs that could lower the risk for heart failure and stroke in coronary artery disease patients.

The model yielded nine drugs considered likely to provide those therapeutic benefits, three of which are currently in use—meaning the analysis identified six candidates for drug repurposing. Among other findings, the analysis suggested that a diabetes medication, metformin, and escitalopram, used to treat depression and anxiety, could lower risk for heart failure and stroke in the model patient population. As it turns out, both of those drugs are currently being tested for their effectiveness against heart disease.

Zhang stressed that what the team found in this case study is less important than how they got there. "My motivation is applying this, along with other experts, to find drugs for diseases without any current treatment. This is very flexible, and we can adjust case-by-case," he said. "The general model could be applied to any disease if you can define the disease outcome."


Provided by The Ohio State University