

Open source machine learning tool could help choose cancer drugs

November 7 2018, by John Toon



Sample tubes from sequencing equipment are shown in Georgia Tech's Petit Institute for Bioengineering and Bioscience. Credit: Rob Felt, Georgia Tech

The selection of a first-line chemotherapy drug to treat many types of cancer is often a clear-cut decision governed by standard-of-care

protocols, but what drug should be used next if the first one fails?

That's where Georgia Institute of Technology researchers believe their new open source decision support tool could come in. Using machine learning to analyze RNA expression tied to information about patient outcomes with specific drugs, the open source tool could help clinicians chose the chemotherapy [drug](#) most likely to attack the disease in individual patients.

In a study using RNA analysis data from 152 patient records, the system predicted the chemotherapy drug that had provided the best outcome 80 percent of the time. The researchers believe the system's accuracy could further improve with inclusion of additional patient records along with information such as family history and demographics.

"By looking at RNA expression in tumors, we believe we can predict with high accuracy which patients are likely to respond to a particular drug," said John McDonald, a professor in the Georgia Tech School of Biological Sciences and director of its Integrated Cancer Research Center. "This information could be used, along with other factors, to support the decisions clinicians must make regarding chemotherapy treatment."

The research, which could add another component to precision medicine for [cancer](#) treatment, was reported November 6 in the journal *Scientific Reports*. The work was supported in part by the Atlanta-based Ovarian Cancer Institute, the Georgia Research Alliance, and a National Institutes of Health fellowship.

As with other machine learning decision support tools, the researchers first "trained" their system using one part of a data set, then tested its operation on the remaining records. In developing the system, the researchers obtained records of RNA from tumors, along with with the

outcome of treatment with specific drugs. With only about 152 such records available, they first used data from 114 records to train the system. They then used the remaining 38 records to test the system's ability to predict, based on the RNA sequence, which [chemotherapy drugs](#) would have been the most likely to be useful in shrinking tumors.

The research began with ovarian cancer, but to expand the data set, the research team decided to include data from other cancer types – lung, breast, liver and pancreatic cancers – that use the same chemotherapy drugs and for which the RNA data was available. "Our model is predicting based on the drug and looking across all the patients who were treated with that drug regardless of cancer type," McDonald said.

The system produces a chart comparing the likelihood that each drug will have an effect on a patient's specific cancer. If the system were to be used in a clinical setting, McDonald believes doctors would use the predictions along with other critical patient information.

Because it measures the expression levels for genes, analysis of RNA could have an advantage over sequencing of DNA, though both types of information could be useful in choosing a drug therapy, he said. The cost of RNA analysis is declining and could soon cost less than a mammogram, McDonald said.

The system will be made available as open source software, and McDonald's team hopes hospitals and cancer centers will try it out. Ultimately, the tool's accuracy should improve as more patient data is analyzed by the algorithm. He and his collaborators believe the open source approach offers the best path to moving the algorithm into clinical use.

"To really get this into clinical practice, we think we've got to open it up so that other people can try it, modify if they want to, and demonstrate

its value in real-world situations," McDonald said. "We are trying to create a different paradigm for cancer therapy using the kind of open source strategy used in internet technology."

Open source coding allows many experts across multiple fields to review the software, identify faults and recommend improvements, said Fredrik Vannberg, an assistant professor in the Georgia Tech School of Biological Sciences. "Most importantly, that means the software is no longer a black box where you can't see inside. The code is openly shared for anybody to improve and check for potential issues."

Vannberg envisions using the decision-support tool to create "virtual tumor boards" that would bring together broad expertise to examine RNA data from patients worldwide.

"The hope would be to provide this kind of analysis for any new cancer patient who has this kind of RNA analysis done," he added. "We could have a consensus of dozens of the smartest people in oncology and make them available for each patient's unique situation."

The tool is available on the [open source](#) Github repository for download and use. Hospitals and cancer clinics may install the software and use it without sharing their results, but the researchers hope organizations using the software will help the system improve.

"The accuracy of machine learning will improve not only as the amount of training data increases, but also as the diversity within that data increases," said Evan Clayton, a Ph.D. student in the Georgia Tech School of Biological Sciences. "There's potential for improvement by including DNA data, demographic information and patient histories. The model will incorporate any information if it helps predict the success of specific drugs."

More information: Cai Huang et al. Machine learning predicts individual cancer patient responses to therapeutic drugs with high accuracy, *Scientific Reports* (2018). [DOI: 10.1038/s41598-018-34753-5](https://doi.org/10.1038/s41598-018-34753-5)

Provided by Georgia Institute of Technology

Citation: Open source machine learning tool could help choose cancer drugs (2018, November 7) retrieved 4 May 2024 from <https://medicalxpress.com/news/2018-11-source-machine-tool-cancer-drugs.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.