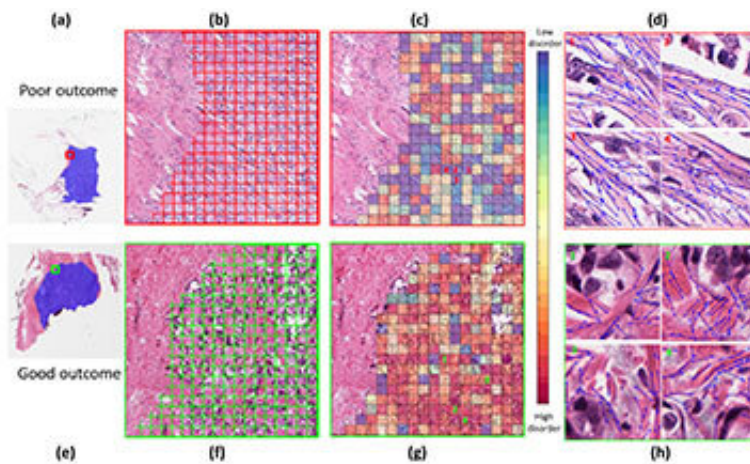# Artificial Intelligence aids in discovery of new prognostic biomarkers for breast cancer

August 6 2021

Images of the collagen network in breast tissue samples. Credit: Case Western Reserve University

Scientists at Case Western Reserve University have used Artificial Intelligence (AI) to identify new biomarkers for breast cancer that can predict whether the cancer will return after treatment—and which can be identified from routinely acquired tissue biopsy samples of early-stage breast cancer.

The key to that initial determination is collagen, a common protein found throughout the body, including in breast tissue. Previous research had suggested that the collagen network, or arrangement of the fibers, relates strongly to breast cancer aggressiveness. But this work by Case

Western Reserve researchers definitively demonstrated collagen's critical role—using only standard tissue biopsy slides and AI.

The researchers, using machine-learning technology to analyze a dataset of digitized tissue samples from breast cancer patients, were able to prove that a well-ordered arrangement of collagen is a key prognostic biomarker for an aggressive tumor and a likely recurrence.

Conversely, they showed that a disordered or broken-down collagen infrastructure not only indicates a better outcome, but actually promotes one. They also found that the disordered collagen network prevents an otherwise aggressive tumor from migrating out of the breast tissue and helps prevent its return after various cancer treatments like chemotherapy.

"It sounds counter-intuitive, but the collagen fibers play a role in tumor migration," said Anant Madabhushi, the Donnell Institute Professor of Biomedical Engineering at Case Western Reserve and head of the Center for Computational Imaging and Personalized Diagnostics (CCIPD). "One way to understand it is to say that if the collagen 'highway' is in terrible shape, it's more difficult for the tumor to migrate, but if it's smooth and organized, it makes it easier for the tumor to hitch a ride."

Doctoral student Haojia Li led the research, which was published in the journal *npj Breast Cancer*. Other authors included Pingfu Fu, professor of Population and Quantitative Health Sciences at the Case Western Reserve School of Medicine, and others from several institutions.

Simple tissue slides, complex computing

Li said the project was important because:

- It validates findings from other published research that suggested highly organized collagen indicates a worse prognosis.
- It was accomplished with digitized images of those simple tissue slides, suggesting this method could become part of a pathologist's routine.

Current methods for examining and investigating the collagen architecture require an expensive and less common electron microscope. "Our method would make predicting outcomes much more available to more doctors and in hospitals which don't have the resources to have an advanced imaging microscope," Li said. "That's why this is so exciting—because it can give the physician the information he or she needs to guide how aggressively to treat the cancer."

The computational work was done in 2020, based on a dataset of routine tissue samples, known as H&E (hematoxylin and eosin) stain slides, taken from patients diagnosed with early stage Estrogen Receptor Positive (ER+) breast cancer.

Breast cancer is the second leading cause of cancer death among women in the United States, with approximately 80% of these cancers being ER+ and 64% being early stage, Li said.

Madabhushi said that because the models built by his team were validated on a completed clinical trial data set, it would "provide a higher level of evidence with regard to the validity of the Collagen signature" and that it would also function as a "natural segue into prospective clinical trial validation."

Madabhushi established the CCIPD at Case Western Reserve in 2012. The lab now includes over 70 researchers and is a global leader in the detection, diagnosis and characterization of various cancers and other diseases, including breast cancer, by meshing medical imaging, machine

learning and AI.

Some of the lab's most recent work, in collaboration with New York University and Yale University, has used AI to predict which lung cancer patients would benefit from adjuvant chemotherapy based on tissue-slide images. That advancement was named by Prevention Magazine as one of the top 10 medical breakthroughs of 2018.

  More information: Haojia Li et al, Collagen fiber orientation disorder from H&E images is prognostic for early stage breast cancer: clinical trial validation, *npj Breast Cancer* (2021). DOI: 10.1038/s41523-021-00310-z

Provided by Case Western Reserve University