A prognostic model may predict survival in African-American women with breast cancer

November 2 2018

A prognostic model developed using a machine learning approach could identify African-American breast cancer patients with increased risk of death, according to results of a study presented at the 11th AACR Conference on The Science of Cancer Health Disparities in Racial/Ethnic Minorities and the Medically Underserved, held here Nov. 2-5.

"Using gene expression data, we have developed a machine learning pattern to accurately stratify African-American breast cancer patients with high and low risks of death, which could help inform clinical decision making," said Shristi Bhattarai, Ph.D. candidate in the lab of Ritu Aneja, Ph.D., at the Department of Biology, Georgia State University. "As African-American women tend to have worse breast cancer outcomes, this study will help us to identify race-based differences in this cohort, which could potentially lead to specific therapeutic regimens for African-American women with breast cancer."

While the incidence of breast cancer is similar between European-American and African- American women in the U.S., the age-adjusted mortality rates are 40 percent higher in African- American women with breast cancer, said Bhattarai. "The etiology of this startling outcome disparity is multifactorial, arising from the combination of socioeconomic inequality with inherently more aggressive tumor biology in women of African ancestry," she noted. "We wanted to identify a fingerprint that could stratify African-American breast cancer patients with different prognostic risks."

Utilizing data from The Cancer Proteome Atlas (TCPA), Bhattarai and colleagues analyzed protein expression levels of 224 proteins from 754 breast cancer patients. Of these patients, 620 were of European descent, and 134 were African-American. The algorithm they developed enabled the researchers to identify significant protein combinations that were associated with breast cancer survival, the authors explained.

The deep learning algorithm identified a combination of four proteins for optimal prognostic prediction: Bcl2-like protein (BAX), inositol polyphosphate-4-phosphatase, type II (INPP4B), X-ray repair cross-complementing protein 1 (XRCC1), and Cleaved Poly (ADP-ribose) polymerase (c-PARP). This combination of proteins could stratify high-risk African-American breast cancer patients with 86 percent accuracy.

"Interestingly, these proteins did not have a significant prognostic value individually," said co-author Sergey Klimov, Ph.D. candidate in the lab of Ritu Aneja in the Department of Biology at Georgia State University. "However, their combined effect within the machine-learning model could identify an African-American cohort that had five times increased risk of death."

After controlling for clinicopathological variables including patients' age and cancer stage, the model could identify African-American women that had nearly 11 times increased risk of death.

The researchers were not able to stratify European-American breast cancer patients into low- and high-risk populations using this specific model, suggesting that this model is only prognostic for African-American breast cancer patients.

"We are moving toward the phase of clinical research where we can identify very specific patterns for understudied demographic groups to find high-risk patients so that they can be recruited for additional

therapies," said Aneja. "We are excited that our model has the potential to inform clinicians to prioritize African-American breast cancer patients for appropriate clinical trials and also help patients make decisions about enrolling in specific clinical trials."

Limitations of this study include a lack of validation in other cohorts. "We will need to validate this model in different groups of African-American breast cancer patients," Aneja noted. "We want to make sure that this model is generalizable to different methodologies."

Provided by American Association for Cancer Research