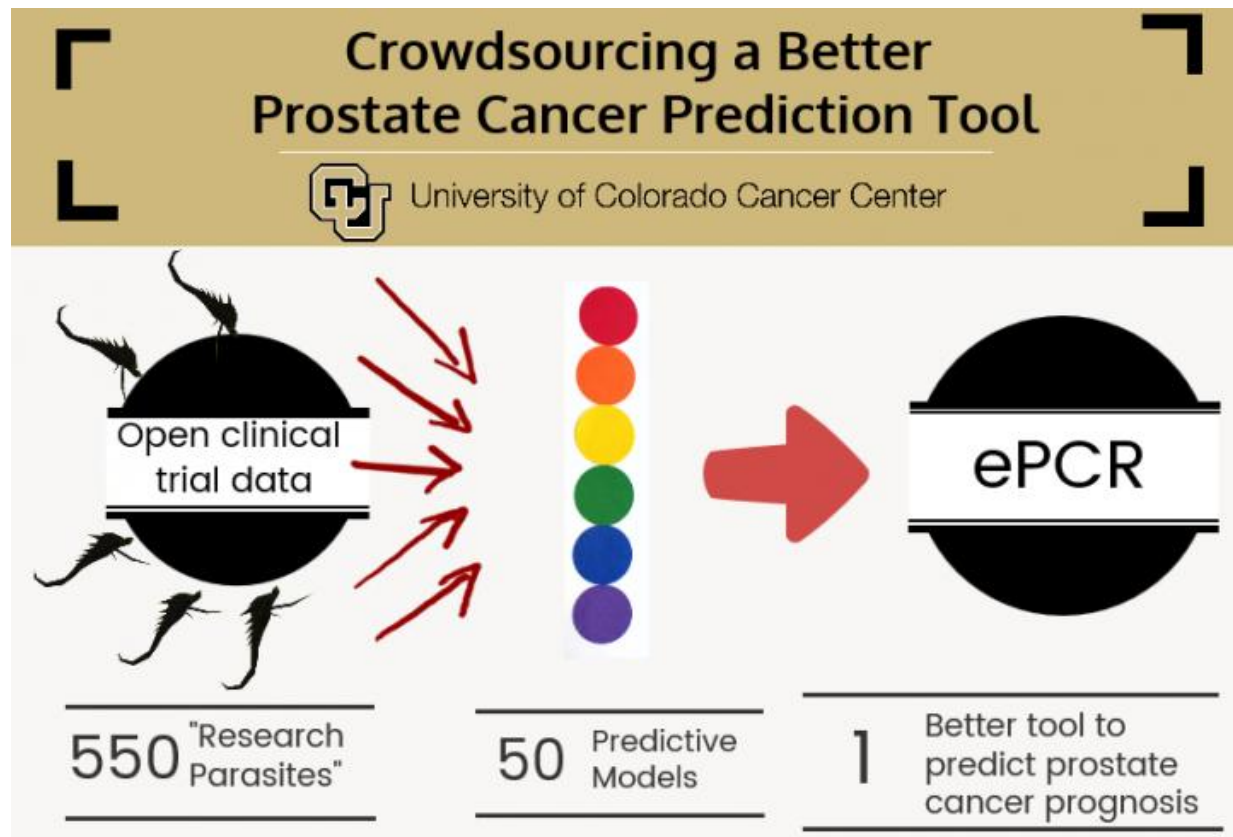# Crowdsourcing a better prostate cancer prediction tool

November 16 2016

Crowdsourced challenge with open clinical trial data leads to international collaboration on prostate cancer prognosis tool. Credit: University of Colorado Cancer Center

Knowing the likely course of cancer can influence treatment decisions.

Now a new prediction model published today in *Lancet Oncology* offers a more accurate prognosis for a patient's metastatic castration-resistant prostate cancer. The approach was as novel as the result - while researchers commonly work in small groups, intentionally isolating their data, the current study embraces the call in Joe Biden's "Cancer Moonshot" to open their question and their data, collecting previously published clinical trial data and calling for worldwide collaboration to evaluate its predictive power. That is, researchers crowdsourced the question of prostate cancer prognosis, eventually involving over 550 international researchers and resulting in 50 computational models from 50 different teams. The approach was intentionally controversial.

"Scientists like me who mine open data have been called 'research parasites'. While not the most flattering name, the idea of leveraging existing data to gain new insights is a very important part of modern biomedical research. This project shows the power of the parasites," says James Costello, PhD, senior author of the paper, investigator at the University of Colorado Cancer Center, assistant professor in the Department of Pharmacology at the CU School of Medicine, and director of Computational and Systems Biology Challenges within the Sage Bionetworks/DREAM organization.

The project was overseen as a collaborative effort between 16 institutions, led by academic research institutions including CU Cancer Center, open-data initiatives including Project Data Sphere, Sage Bionetworks, and the National Cancer Institute's DREAM Challenges, and industry and research partners including Sanofi, AstraZeneca, and the Prostate Cancer Foundation. Challenge organizers made available the results from five completed clinical trials. Teams were challenged to connect a deep set of clinical measurements to overall patient survival, organizing their insights into novel computational models to better predict patient survival based on clinical data.

"The idea is that if a patient comes into the clinic and has these measurements and test results, can we put this data in a model to say if this patient will progress slowly or quickly. If we know the features of patients at the greatest risk, we can know who should receive standard treatment and who might benefit more from a clinical trial," Costello says.

The most successful of the 50 models was submitted by a team led by Tero Aittokallio, PhD, from the Institute for Molecular Medicine Finland, FIMM, at University of Helsinki, and professor in the Department of Mathematics and Statistics at University of Turku, Finland.

"My group has a long-term expertise in developing multivariate machine learning models for various biomedical applications, but this Challenge provided the unique opportunity to work on clinical trial data, with the eventual aim to help patients with metastatic castration-resistant prostate cancer," Aittokallio says.

Basically, the model depended on not only groups of single patient measurements to predict outcomes, but on exploring which interactions between measurements were most predictive - for example, data describing a patient's blood system composition and immune function were only weakly predictive of survival on their own, but when combined became an important part of the winning model. The model used a computational learning strategy technically referred to as an ensemble of penalized Cox regression models, hence the model's name ePCR. This model then competed with 49 other entries, submitted by other teams working independently around the world.

"Having 50 independent models allowed us to do two very important things. First when a single clinical feature known to be predictive of patient survival is picked out by 40 of the 50 teams, this greatly

strengthens our overall confidence. Second, we were able to discover important clinical features we hadn't fully appreciated before," Costello says.

In this case, many models found that in addition to factors like prostate-specific antigen (PSA) and lactate dehydrogenase (LDH) that have long been known to predict prostate cancer performance, blood levels of an enzyme called asparate aminotransferease (AST) is an important predictor of patient survival. This AST is an indirect measure of liver function and the fact that disturbed levels of AST are associated with poor patient performance implies that studies could evaluate the role of AST in prostate cancer.

"The benefits of a DREAM Challenge are the ability to attract talented individuals and teams from around the world, and a rigorous framework for the assessment of methods. These two ingredients came together for our Challenge, leading to a new benchmark in metastatic prostate cancer ," says paper first author, Justin Guinney, PhD, director of Computational Oncology for Sage Bionetworks located at Fred Hutchinson Cancer Research Center.

"A goal of the Project Data Sphere initiative is to spark innovation - to unlock the potential of valuable data by generating new insights and opening up a new world of research possibilities. Prostate Cancer DREAM Challenge did just that. To witness cancer clinical trial data from Project Data Sphere be used in research collaboration and ultimately help improve patient care in the future is extremely rewarding!" says Liz Zhou, MD, MS, director of Global Health Outcome Research at Sanofi.

The goal now is to make the ePCR model publicly accessible through an online tool with an eye towards clinical application. In fact, the National Cancer Institute (NCI) has contracted the winning team to do exactly

this. Soon, when patients face difficult decisions about the best treatment for metastatic castration-resistant prostate cancer, ePCR tool could be an important piece of the decision-making process.

Provided by CU Anschutz Medical Campus

Citation: Crowdsourcing a better prostate cancer prediction tool (2016, November 16) retrieved 24 May 2024 from https://medicalxpress.com/news/2016-11-crowdsourcing-prostate-cancer-tool.html