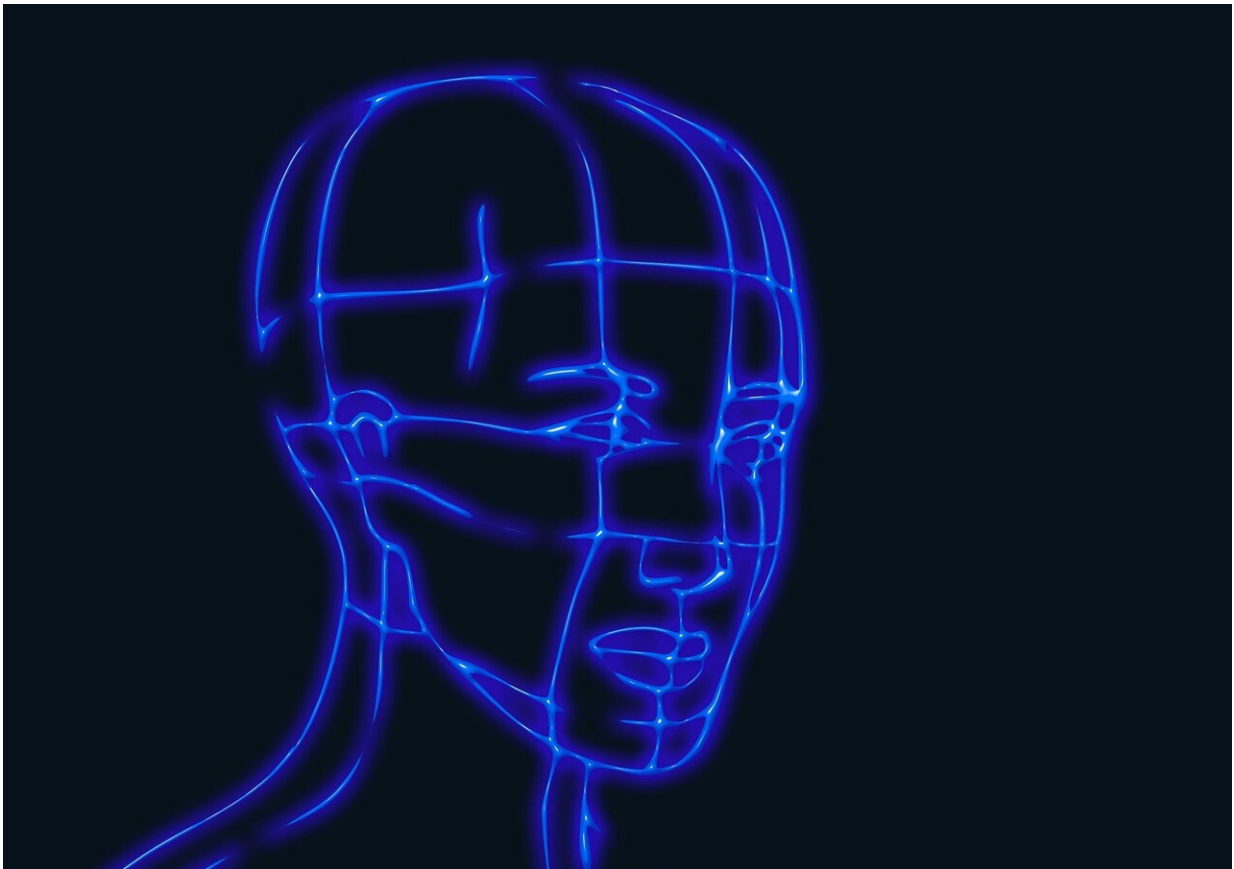


# Machine learning model uses clinical and genomic data to predict immune checkpoint blockade effectiveness

November 4 2021

---



Credit: CC0 Public Domain

A computer model developed by Cleveland Clinic oncologist Timothy

Chan, MD, Ph.D., and colleagues accurately predicts whether immune checkpoint blockade (ICB) will be effective in patients diagnosed with a wide variety of cancers.

The forecasting tool, developed using [machine learning](#), assesses multiple biological and clinical variables in an individual patient's condition to predict the degree of response to immune checkpoint inhibitors and survival outcomes. It markedly outperforms individual biomarkers or other combinations of variables developed so far.

With further validation, the tool eventually should help oncologists better identify patients likely to benefit from ICB for which biomarkers are needed. Discerning, prior to treatment, patients for whom ICB would be ineffective could reduce unnecessary expense and exposure to potential side effects. It could also indicate the need to pursue alternate strategies, such as combination therapies.

"It's important to know which treatment modalities patients are most suited for," says Dr. Chan, Director of Cleveland Clinic's Center for Immunotherapy and Precision Immuno-Oncology. "Moreover, understanding who responds and who doesn't allows you to know what to target next, because those are factors that are impeding response. Our model provides a comprehensive understanding of the diversity of responses among patients to [immune checkpoint blockade](#). It's the first to assemble such a large-scale set of clinical and genomic variables that have predictive value for immunotherapy across numerous cancer types."

## **The complexities of immunotherapy response**

Immune checkpoint pathways are inhibitory cell surface signaling proteins such as programmed cell death receptor/ligand 1 (PD-1/PD-L1) and cytotoxic T lymphocyte-associated molecule 4 and its ligand

(CTLA-4/B7-1/B7-2) that work in concert to downregulate T cell-mediated immunogenicity, thus maintaining self-tolerance and protecting against collateral tissue damage.

Cancer cells have evolved multiple mechanisms to avoid immune attack, including the upregulation of negative regulatory pathways to exploit immune checkpoints by suppressing T cell functioning in the [tumor microenvironment](#).

The recent advent of ICB as a tactic to revive antitumor immune surveillance has been a significant advance in cancer treatment. Antibodies targeting CTLA-4 or PD-1/PD-L1—the most common checkpoint objectives—have induced durable responses in patients with some advanced-stage cancers.

However, ICB is not effective in all cancer types, and even in responsive cancers, efficacy rates do not top 50 percent, meaning half or more of patients do not derive clinical benefit. These patients experience disease progression while also incurring substantial costs; the list price of the anti-PD-1 monoclonal antibody pembrolizumab, for example, exceeds \$10,000 per course.

Previous research has identified some biomarkers and genomic features associated with ICB efficacy. But no single factor can be considered an optimal predictor of treatment outcomes.

"There has been a big push to try to understand what is driving response to immunotherapy," says Dr. Chan, whose lab at Memorial Sloan Kettering, prior to his arrival at Cleveland Clinic, made [foundational discoveries](#) in this area, including the [finding](#) that immune checkpoint inhibitors ultimately target somatic mutations that develop in tumors.

"That discovery sparked major activity around the world to study these

neoantigens," he says. "But it turns out that mutational load is only part of the story. Our latest research was an unbiased global analysis searching for all of the different factors that may be affecting response to immune checkpoint blockade."

## Applying machine learning

The machine learning method has been shown to produce reliable outcome predictions derived from multiple, seemingly unrelated variables. Dr. Chan and his colleagues decided to apply it to the problem of predicting immune checkpoint blockade efficacy.

Machine learning is a way of programming a computer to execute a complex task driven by statistics and comparison with known occurrences. The programming algorithm guides the computer's review of a large, diverse dataset with the goal of identifying patterns and using them to predict outcomes or reach conclusions.

Initially, the computer program (known as a classifier) learns using a training dataset. It extracts and classifies information. Through iterative trial-and-error experience, comparing its results to examples of correct outcomes, the classifier infers how to consistently derive accurate answers, thus improving its predictive capability without explicit instruction from programmers. It can then apply this learned experience to new, unstructured datasets.

Dr. Chan and his colleagues began by assembling a dataset containing clinical, [tumor](#) and genetic sequencing information from 1,479 patients with 16 different cancer types: non-small cell lung cancer (36 percent), melanoma (13 percent), renal (6 percent), bladder (6 percent), head and neck (5 percent), sarcoma (5 percent), endometrial (4 percent), gastric (4 percent), hepatobiliary (4 percent), small cell lung cancer (3 percent), colorectal (3 percent), esophageal (3 percent), pancreatic (2 percent),

mesothelioma (2 percent), ovarian (2 percent) and breast (2 percent). The patients were treated with PD-1/PD-L1 inhibitors, CTLA-4 blockade, or a combination of both. A total of 409 patients (28 percent) responded to ICB either partially or completely; 1,070 (72 percent) were nonresponsive, meaning they experienced either stable or progressive disease.

The researchers then applied an algorithm known as random forest, an approach that is composed of multiple individual decision trees that operate together to improve the program's predictive accuracy.

Their random forest classifier incorporated 16 genomic, molecular, clinical and demographic variables, some of which have been shown to be associated with ICB response. The variables were tumor mutational burden, fraction of copy-number alteration, human leukocyte antigen class I (HLA-I) evolutionary divergence, loss of heterozygosity status in HLA-I, microsatellite instability status, blood neutrophil-to-lymphocyte ratio, tumor stage at the start of ICB treatment, type of ICB drug, body mass index, gender, age at the start of ICB treatment, cancer type, whether the patient received chemotherapy before ICB, and blood levels of albumin, platelets and hemoglobin.

The researchers refined their classifier by applying it to a randomized training subsample of the original dataset, then tested its predictive capability on a second subsample.

The trained classifier can provide a cancer-specific prediction of an individual patient's probability of response to ICB, based on the aggregated predictive power of the 16 selected clinical, molecular, demographic and genomic factors. It can also quantify how much each of those factors contributes to variation in response among patients.

The classifier revealed that the variable exerting the greatest influence

on ICB response is tumor mutational burden, followed closely by a patient's chemotherapy history. Surprisingly, the three selected blood markers included in the classifier—albumin, platelet and hemoglobin levels, which are indicative of a patient's overall health—also had strong predictive value, not only for forecasting a patient's overall survival (as some previous studies had established), but the actual radiographic response to ICB treatment itself.

"We did not expect that some of these factors were actually important for tumor shrinkage," Dr. Chan says. "To find albumin levels at No. 3 is surprising. How these variables all work together is really the key here. This model shows that, rather than a single predictive biomarker, we're headed toward a multifactor nomogram for clinical use."

## Judging the model's performance

To gauge how well their model performed, Dr. Chan and his colleagues compared its predictions with those of two other forecasting tools:

- Tumor mutational burden, which the FDA approved in 2020 as a biomarker to predict anti-PD-1 ICB efficacy in solid tumors.
- A second random forest classifier the researchers created that retained 11 ICB response-associated variables from the original model (tumor mutational burden, fraction of copy-number alteration, HLA-I evolutionary divergence, loss of heterozygosity status in HLA-I, microsatellite instability status, neutrophil-to-lymphocyte ratio, BMI, gender, age, tumor stage, and ICB drug class) but eliminated five clinical variables (cancer type, chemotherapy history, and levels of albumin, hemoglobin and platelets)

The original, fully integrated model proved to be highly accurate, significantly outperforming both tumor mutation burden and the reduced-

variable model in predicting ICB responders and non-responders across cancer types. The fully integrated model's predictions of progression-free and overall survival were significantly more accurate than those of either tumor mutational burden or the reduced-variable model.

When tested alone, none of the individual variables in the original model could match the predictive power of the fully integrated model, which indicates to the researchers that those factors are combining in a nonlinear way to achieve their accuracy.

"The model works well, despite what type of cancer is being assessed, which shows that these commonalities are what's important," Dr. Chan says. "These are primary factors that affect ICB response. The factors may be weighted a little bit differently from cancer to cancer, but it's almost like a common language" for response prediction.

Compared to tumor mutational burden alone, the fully integrated model consistently performed better as measured by sensitivity, specificity, accuracy, and positive and negative predictive value.

The model's predictive superiority to tumor mutational burden could be particularly important in making treatment decisions involving patients with low mutation-burden tumors. "There are certain disease types, like sarcomas or bladder cancer or rarer tumors, where physicians don't really have the ability to detect which patients are likely to be exceptional immunotherapy responders," Dr. Chan says. "This model extends upon the [predictive value](#) of mutational load. So we might be able to find groups of patients who today would not be treated with immunotherapies but might actually be able to avail themselves and have some success."

## **The path to clinical use and improved prediction**



Taken together, the positive results support moving forward to test the model in a clinical trial with a large, diverse cohort of cancer patients, Dr. Chan says. That should provide a more accurate assessment of its performance in a real-world setting.

"We're in talks with genomics diagnostic companies to explore developing this into a product," he says. "One could make the predictive model a companion diagnostic in a clinical trial of an immunotherapy agent," as was done with tumor mutational burden as the companion diagnostic to identify patients with unresectable or metastatic solid tumors who might benefit from treatment with pembrolizumab. "If the model is predictive in a prospective clinical trial, the next step is to file for FDA approval."

Meanwhile, as knowledge of the factors affecting ICB response advances, Dr. Chan says the model's predictive accuracy could be improved by using machine learning to assess the combinatorial power of additional potential predictors. Those could include molecular features of the tumor microenvironment, composition of the microbiome, the diversity of the T-cell receptor repertoire, specific tumor genomic alterations or mutations associated with resistance to ICB, and transcriptomic data.

**More information:** Diego Chowell et al, Improved prediction of immune checkpoint blockade efficacy across multiple cancer types, *Nature Biotechnology* (2021). DOI: 10.1038/s41587-021-01070-8

Provided by Cleveland Clinic

Citation: Machine learning model uses clinical and genomic data to predict immune checkpoint blockade effectiveness (2021, November 4) retrieved 24 May 2024 from



<https://medicalxpress.com/news/2021-11-machine-clinical-genomic-immune-checkpoint.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.