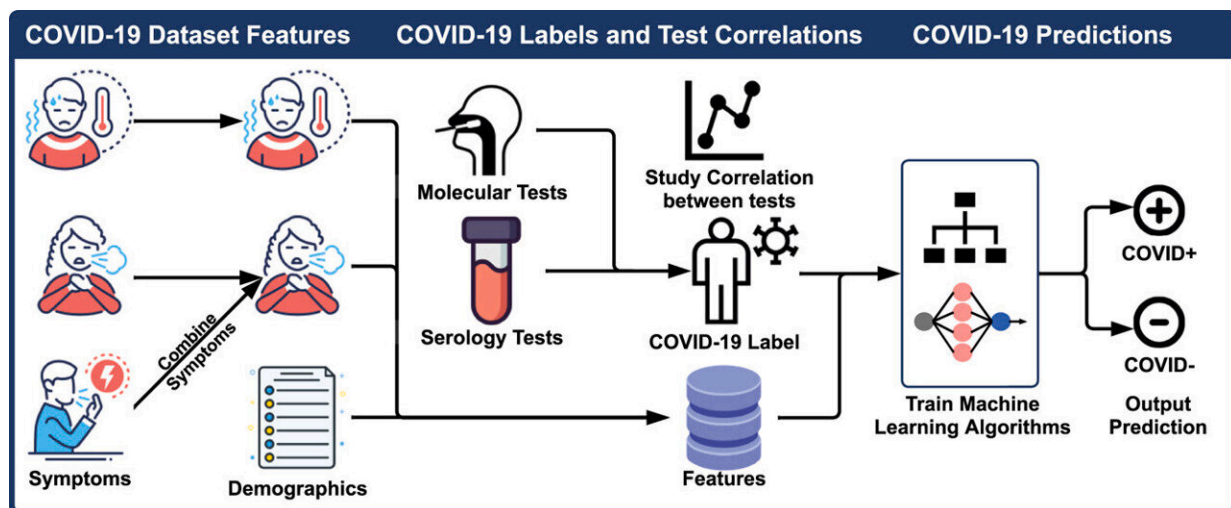


AI model proactively predicts if a COVID-19 test might be positive or not

December 13 2022



Graphical abstract. Credit: *Smart Health* (2022). DOI: 10.1016/j.smhl.2022.100331

COVID-19 and its latest omicron strains continue to cause infections across the country as well as globally. Serology (blood) and molecular tests are the two most commonly used methods for rapid COVID-19 testing. Because COVID-19 tests use different mechanisms, they vary significantly. Molecular tests measure the presence of viral SARS-CoV-2 RNA while serology tests detect the presence of antibodies triggered by the SARS-CoV-2 virus.

Currently, there is no existing study on the correlation between serology and [molecular tests](#) and which COVID-19 symptoms play a key role in producing a positive test result. A study from Florida Atlantic University's College of Engineering and Computer Science using machine learning provides important new evidence in understanding how molecular tests versus serology tests are correlated, and what features are the most useful in distinguishing between COVID-19 positive versus test outcomes.

Researchers from the College of Engineering and Computer Science trained five classification algorithms to predict COVID-19 test results. They created an accurate predictive model using easy-to-obtain symptom features, along with demographic features such as number of days post-symptom onset, fever, temperature, age and gender.

The study demonstrates that [machine-learning models](#), trained using simple symptom and demographic features, can help predict COVID-19 infections. Results, published in the journal *Smart Health*, identify the key symptom features associated with COVID-19 infection and provide a way for rapid screening and cost effective infection detection.

Findings reveal that number of days experiencing symptoms such as fever and difficulty breathing play a large role in COVID-19 test results. Findings also show that molecular tests have much narrower post-symptom onset days (between three to eight days), compared to post-symptom onset days of serology tests (between five to 38 days). As a result, the molecular test has the lowest positive rate because it measures current infection.

Furthermore, COVID-19 tests vary significantly, partially because donors' immune response and viral load—the target of different test methods—continuously change. Even for the same donor, it might be possible to observe different positive/negative results from two types of

tests.

"Molecular tests depend on viral load and serology tests depend on seroconversion, which is the period during which the body starts producing detectable levels of antibodies. Both of these tests are time dependent," said Xingquan "Hill" Zhu, Ph.D., senior author and a professor in FAU's Department of Electrical Engineering and Computer Science. "Our results suggest that the number of days post symptomatic are highly important for a positive COVID-19 test and should be under careful consideration when screening patients."

For the study, researchers used test results from 2,467 donors, each tested using one or multiple types of COVID-19 tests, which were collected as the testbed. They combined symptoms and demographic information to design a set of features for predictive modeling using the five types of machine-learning models.

By cross checking test types and results, they examined the correlation between serology and molecular tests. For test outcome prediction, they labeled the 2,467 donors as positive or negative by using their serology or molecular test results, and created symptom features to represent each donor for machine learning.

"Because COVID-19 produces a wide range of symptoms and the data collection process is essentially error prone, we grouped similar symptoms into bins," said Zhu. "Without a standardization of symptom reporting, the symptom feature space greatly increases. To combat this, we utilized this binning approach, which was able to decrease symptom feature space while keeping sample feature information."

By using created bin features, combined with the five machine-learning algorithms, these predictive models achieved more than 81 percent AUC scores (Area under the ROC Curve, which provides an aggregate

measure of performance across all possible classification thresholds), and more than 76 percent classification accuracy.

"One unique feature of our testbed is that some donors may have multiple test results, which allowed us to analyze the relationship between serology tests versus molecular tests, and also understand consistency within each type of test," said Zhu.

The five [machine learning](#) models used by the researchers are Random Forest, XGBoost, Logistic Regression, Support Vector Machine (SVM) and Neural Network. They compared performance by using three performance metrics: Accuracy, F1-score and AUC.

"Predictive modeling is complicated by many puzzling questions unanswered by research. The testbed created by our researchers is indeed novel and clearly shows correlation between different types of COVID-19 tests," said Stella Batalama, Ph.D., dean, FAU College of Engineering and Computer Science.

"Our researchers have designed a new way to narrow down noisy symptom features for clinical interpretation and predictive modeling. Such AI based predictive modeling approaches are becoming increasingly powerful to combat infectious diseases and many other aspects of health issues."

More information: Magdalyn E. Elkin et al, A machine learning study of COVID-19 serology and molecular tests and predictions, *Smart Health* (2022). [DOI: 10.1016/j.smhl.2022.100331](https://doi.org/10.1016/j.smhl.2022.100331)

Provided by Florida Atlantic University

Citation: AI model proactively predicts if a COVID-19 test might be positive or not (2022, December 13) retrieved 21 July 2023 from <https://medicalxpress.com/news/2022-12-ai-proactively-covid-positive.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.