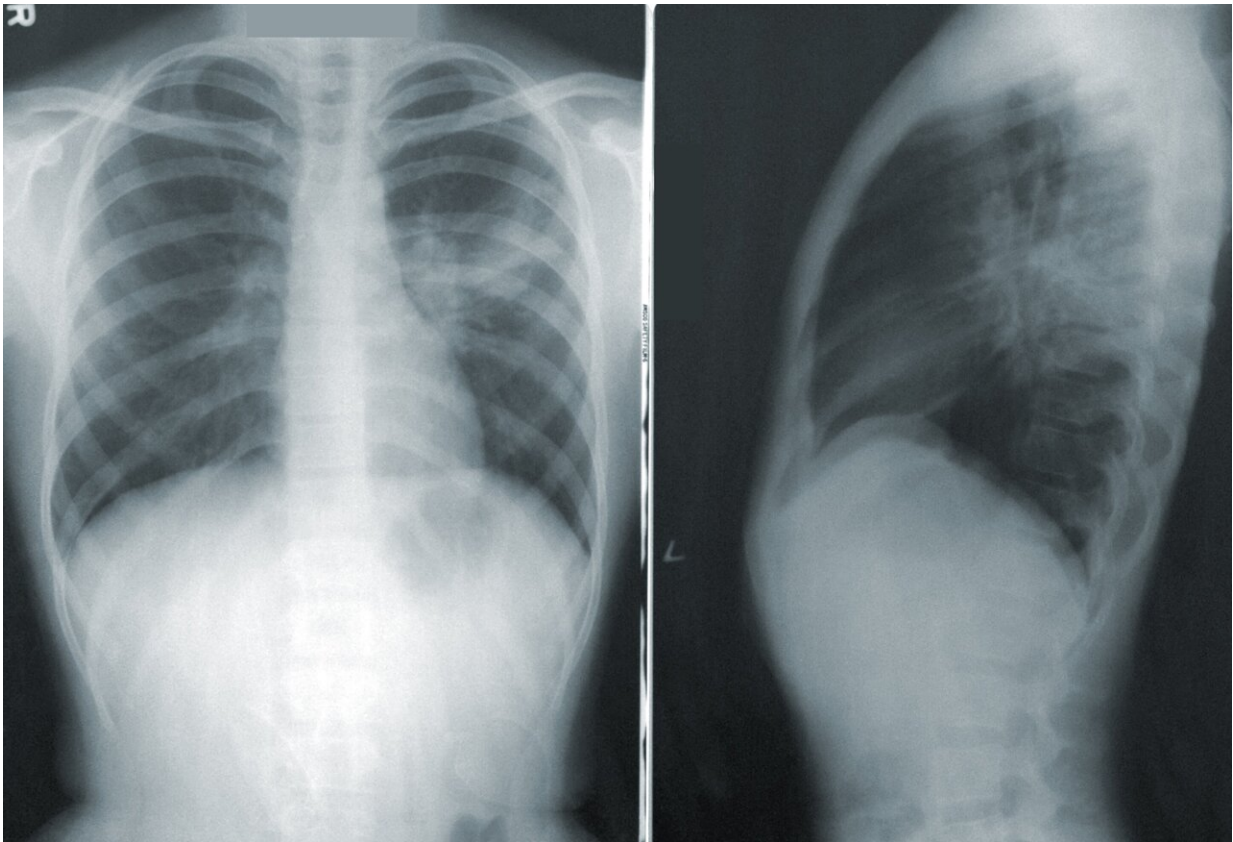


'Self-taught' AI tool helps to diagnose and predict severity of common lung cancer

June 11 2024



Credit: Unsplash/CC0 Public Domain

A computer program based on data from nearly a half-million tissue images and powered by artificial intelligence can accurately diagnose cases of adenocarcinoma, the most common form of lung cancer, a new

study shows.

Researchers at NYU Langone Health's Perlmutter Cancer Center and the University of Glasgow developed and tested the program. They say that because it incorporates structural features of tumors from 452 adenocarcinoma patients, who are among the more than 11,000 patients in the United States National Cancer Institute's Cancer Genome Atlas, the program offers an unbiased, detailed, and reliable second opinion for patients and oncologists about the presence of the cancer and the likelihood and timing of its return (prognosis).

The research team also points out that the program is independent and "self-taught," meaning that it determined on its own which structural features were statistically most significant to gauging the severity of disease and had the greatest impact on tumor recurrence.

Publishing in the journal *Nature Communications*, the study program, also called an algorithm—or specifically, histomorphological phenotype learning (HPL)—was found to accurately distinguish between similar lung cancers, adenocarcinoma and [squamous cell cancers](#), 99% of the time. The HPL program was also found to be 72% accurate at predicting the likelihood and timing of cancer's return after therapy, improving upon the 64% accuracy in the predictions made by pathologists who directly examined the same patients' tumor images, researchers say.

"Our new histomorphological phenotype learning program has the potential to offer cancer specialists and their patients a quick and unbiased diagnostic tool for lung adenocarcinoma, that—once further testing is complete—can also be used to help validate and even guide their treatment decisions," said study lead investigator Nicolas Coudray, Ph.D., a bioinformatics programmer at NYU Grossman School of Medicine and Perlmutter Cancer Center.

"Patients, physicians, and researchers know they can rely on this predictive modeling because it is self-taught, provides explainable decisions, and is based only on the knowledge drawn specifically from each patient's tissue, including such features as its proportion of dying cells, tumor-fighting immune cells, and how densely packed the tumor cells are, among other features," said Coudray.

"Lung tissue samples can now be analyzed in minutes by our computer program to provide fairly accurate predictions of whether their cancer will return, predictions that are better than current standards of care for making a prognosis in lung adenocarcinoma," said study co-senior investigator Aristotelis Tsirigos, Ph.D. Tsirigos is a professor in the Departments of Pathology and Medicine at NYU Grossman School of Medicine and Perlmutter Cancer Center, where he also serves as co-director of precision medicine and director of its Applied Bioinformatics Laboratories.

Tsirigos says that thanks to such tools and other advances in lung cancer biology, pathologists will be examining tissue scans on their computer screens, and fewer and fewer on microscopes, and then using their AI program to analyze the image and produce its own image of the scan.

The new image, or "landscape," they add, will offer a detailed breakdown of the tissue's content. It might note, for example, that there is 5% necrosis and 10% tumor infiltration and what that means in terms of survival. That reading may statistically equate to an 80% chance of remaining cancer-free for two years or more, based on information from all the patient data in the program.

To develop the HPL program, the researchers first analyzed lung adenocarcinoma tissue slides from the Cancer Genome Atlas. Adenocarcinoma was chosen for the test model because the disease is known for characteristic features. As an example, they note that its

tumor cells tend to group in so-called acinar, or saclike patterns and spread predictably along the surface lining of lung cells.

From their analysis of the slides, whose visual images were digitally scanned and broken into 432,231 small quadrants or tiles, researchers found 46 key characteristics, what they term histomorphological phenotype clusters, from both normal and diseased tissue, a subset of which were statistically linked to either cancer's early return or to long-term survival. The findings were then confirmed by further and separate testing on tissue images from 276 men and women who were treated for adenocarcinoma at NYU Langone from 2006 to 2021.

Researchers say their goal is to use the HPL algorithm to assign to each patient a score between 0 and 1 that reflects their statistical chance of survival and tumor recurrence for up to five years. Because the program is self-learning, they stress HPL will become increasingly more accurate as more data is added over time. To build public trust, researchers have posted their programming code [online](#) and have plans to make the new HPL tool freely available upon completion of further testing.

Characteristics linked to tumors recurring included high tile percentages of dead cancer cells and tumor-fighting [immune cells](#) called lymphocytes, and the dense clustering of [tumor cells](#) in the outer linings of the lungs. Features tied to increased likelihood for survival were high percentages of unchanged or preserved lung sac tissue, and lack of or mild presence of inflammatory cells.

Tsirigos says the team next plans to look at developing HPL-like programs for other cancers, such as breast, ovarian, and colorectal, that are similarly based on distinctive and key morphological features and additional molecular data. The team also has plans to expand and improve the accuracy of the current adenocarcinoma HPL program by including other data from hospital electronic health records about other

illnesses and diseases, or even income and home ZIP code.

Besides Tsirigos and Coudray, other NYU Langone researchers involved in this study are Anna Yeaton, Bojing Liu, Hortense Le, Luis Chiriboga, Afreen Karimkhan, Navneet Natula, Christopher Park, Harvey Pass, and Andre Moreira. Study co-lead investigator Adalberto Claudio Quiros, study co-investigators Xinyu Yang and John Le Quesne, and study co-senior investigator Ke Yuan are all at the University of Glasgow, UK. Study co-investigator David Moore is at University College London, UK.

More information: Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unlabeled, unannotated pathology slides, *Nature Communications* (2024). [DOI: 10.1038/s41467-024-48666-7](https://doi.org/10.1038/s41467-024-48666-7)

Provided by NYU Langone Health

Citation: 'Self-taught' AI tool helps to diagnose and predict severity of common lung cancer (2024, June 11) retrieved 18 June 2024 from <https://medicalxpress.com/news/2024-06-taught-ai-tool-severity-common.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.