A Machine Learning Model May Enable Liver Cancer Risk Prediction With Routine Clinical Information

Many individuals with undiagnosed cirrhosis or other risk factors might benefit from liver cancer screening.

March 31, 2026 By American Association for Cancer Research

Philadelphia – A machine learning model that analyzes patient demographics, electronic health record data, and routine blood test results predicted a patient's risk of hepatocellular carcinoma (HCC), the most common type of liver cancer, with high accuracy, according to a study published in Cancer Discovery, a journal of the American Association for Cancer Research (AACR).

Individuals who are considered to have an elevated risk for HCC may be eligible for imaging-based and blood-based cancer screening to enable early detection of the disease; however, current guidelines focus on a narrow, high-risk population and miss many at-risk individuals, explained co-senior and corresponding author Carolin Schneider, MD, an assistant professor at RWTH Aachen University in Germany, who co-led the study with Jakob Kather, MD, MSc, a professor of clinical artificial intelligence at the Technical University of Dresden, Germany.

"Screening is typically recommended for patients with confirmed liver cirrhosis or severe liver disease, since many cases of HCC occur in these patients, but there are many individuals with undiagnosed cirrhosis or other risk factors who might also benefit from liver cancer screening," she said.

Additional factors that increase a patient's risk for developing HCC include being male, smoking, and heavy alcohol consumption, among others, added Jan Clusmann, MD, the first author of the study and a clinician-scientist at the Technical University of Dresden.

"With so many factors impacting risk, there is an urgent need for effective tools to help clinicians identify high-risk patients," he said. "Machine learning tools that can simultaneously work with different types of clinical data could be particularly useful for this major clinical challenge."

In this study, Clusmann, Kather, Schneider, and colleagues used data from the UK Biobank to develop machine learning models that analyzes different types of clinical data to assess HCC risk. The UK Biobank contained data from more than 500,000 individuals in the United Kingdom and

included 538 HCC cases, 69% of which occurred in patients without prior diagnoses of liver cirrhosis, viral hepatitis, or other chronic liver diseases.

The researchers trained their models on 80% of the data from the UK Biobank and performed an initial validation on the remaining 20%. External validation was performed using the All of Us registry, which included data from more than 400,000 individuals in the United States, including substantial representation of populations that have been historically underrepresented in medical research, the authors noted. The registry included 445 cases of HCC.

The models developed by the authors used a "random forest architecture," a method that combines hundreds of decision trees. Each tree makes a series of simple yes-or-no decisions based on a series of variables from patient data, and the final prediction is determined by aggregating the results across all trees, making the model more robust, reliable, and interpretable, Clusmann explained.

A separate random forest model was trained for each of five different types of clinical data, as well as for stepwise combinations of data in ascending order of clinical availability: patient demographics, electronic health record data, blood test results, genomics, and metabolomics. The performance of these models was assessed by calculating the area under the receiver operating characteristic (AUROC), which describes the algorithm's ability to distinguish between two groups (in this case, patients in the validation cohort with HCC vs. those without), with 1 being a perfect score.

The researchers found that a model combining demographics, electronic health records, and blood tests (Model C) resulted in the best performance, with an AUROC of 0.88. Adding genomics and/or metabolomics data did not substantially increase performance.

"This showed that we can predict HCC risk using simple, readily available data without the need for complex and expensive genetic sequencing," said Schneider, noting that this feature increases the model's potential for widespread use, particularly in resource-limited settings.

The researchers then compared the performance of their models with that of previously reported liver cancer risk prediction models. They included the clinically available FIB-4, APRI, and NFS scores, which are commonly used to determine a patient's likelihood of liver fibrosis (a known risk factor for liver cancer), as well as the aMAP score, which uses clinical factors such as age, sex, albumin levels, bilirubin levels, and platelet count to predict the risk of liver cancer in patients with chronic liver disease.

The authors found that their model was better than existing scores at finding true cases of HCC while producing fewer false positives. To make Model C more practical for a clinical setting, the researchers then reduced the number of clinical features it examined in a so-called "ablation experiment." The result, a simplified model version that examined as few as 15 routinely collected clinical features, still outperformed existing risk prediction models.

"Our study highlights the potential of a simple, easily utilized machine learning model to improve

risk stratification for HCC using only routinely collected clinical data," said Schneider. "If validated in additional populations, our model would enable primary care physicians to efficiently identify at-risk patients and refer them to liver cancer screening. This could enable earlier detection and improved outcomes for patients with this aggressive disease."

Clusmann added that the final model demonstrated strong generalizability: Although trained predominantly on data from white participants in the UK Biobank, it maintained robust performance when evaluated specifically in the non-white subgroup of the more ethnically diverse All of Us cohort, suggesting broad applicability across populations.

Limitations of the study include its retrospective design and low fraction of patients with viral hepatitis, a known risk factor for liver cancer, in the training and validation cohorts. Further validation is needed to evaluate the performance of the machine learning model in different populations, the authors noted.

The study was supported by the German Cancer Aid; the German Federal Ministry of Research, Technology and Space; the German Research Foundation; the German Academic Exchange Service; the German Federal Joint Committee; the European Union Horizon Europe Research and Innovation Programme; the Breast Cancer Research Foundation; the National Institute for Health and Care Research; the German Federal Ministry of Education and Research; the German Federal Ministry of Health; the Interdisciplinary Centre for Clinical Research at RWTH Aachen University; the Junior Principal Investigator Fellowship Program of RWTH Aachen Excellence Strategy; the NRW Rueckkehr Programme of the Ministry of Culture and Science of the German State of North Rhine-Westphalia; and the National Institutes of Health. Clusmann has received honoraria from Johnson & Johnson. Kather declares ongoing consulting services for AstraZeneca and Bioptimus; holds shares in StratifAI, Synagen, and Spira Labs; and has received institutional research grants from GSK and AstraZeneca and honoraria from AstraZeneca, Bayer, Daiichi Sankyo, Eisai, Janssen, Merck, MSD, Bristol Myers Squibb, Roche, Pfizer, and Fresenius. Schneider declares no conflicts of interest.

This news release was published by the American Association for Cancer Research on March 26, 2026.
