RSNA Press Release

Deep Learning Model Estimates Cancer Risk of Lung Nodules

Released: September 16, 2025

OAK BROOK, Ill. — An artificial intelligence (AI) deep learning tool that estimates the malignancy risk of lung nodules achieved high cancer detection rates while significantly reducing false-positive results. Results of the study, which used data from large, multi-site lung cancer screening trials, were published today in *Radiology*, a journal of the Radiological Society of North America (RSNA).

Lung cancer remains a significant global health issue, causing the most cancer-related deaths worldwide. Screening high-risk individuals with low-dose chest CT has been shown to reduce lung cancer mortality. However, early screening trials have reported high false-positive rates, leading to unnecessary follow-up procedures, increased patient anxiety and health care costs.

Pulmonary nodules—small round or oval growths in the lungs—are common, and identifying which are malignant is challenging in lung cancer screening.

"Deep learning offers promising solutions, but robust validation is essential," said Noa Antonissen, M.D., lead researcher and Ph.D. candidate at Radboud University Medical Center, Nijmegen, the Netherlands. "AI accounts for factors that we might not even see on the CT scan to further assess a nodule as likely to be malignant."

download photo



Noa Antonissen, M.D.

Most current lung cancer screening protocols rely on nodule size, type and growth to estimate malignancy risk. The Pan-Canadian Early Detection of Lung Cancer (PanCan) model, which estimates nodule malignancy risk through a combination of patient and nodule characteristics, illustrates how probability-based tools can refine risk assessment. Such probability-based risk thresholds are increasingly used to guide management protocols. Deep learning offers a promising alternative by enabling fully data-driven predictions, but more evidence is needed before it can be adopted in clinical practice.

In the retrospective study, the researchers trained their in-house developed deep learning algorithm to estimate the risk for malignancy for lung nodules using data from the National Lung Screening Trial which included 16,077 nodules (1,249 malignant).

External testing was conducted using baseline CT scans from the Danish Lung Cancer Screening Trial, the Multicentric Italian Lung Detection trial and the Dutch–Belgian NELSON trial. The pooled cohort included 4,146 participants (median age 58 years, 78% male, median smoking history 38 pack-years) with 7,614 benign and 180 malignant nodules.

The researchers assessed the algorithm's performance for the pooled cohort and two subsets: indeterminate nodules (5-15 mm) and malignant nodules that were size-matched to benign nodules.

"We selected nodules sized 5–15 mm, due to their diagnostic challenges and frequent need for short-term follow-up," Dr. Antonissen said. "Accurate risk classification of these nodules could reduce unnecessary procedures."

For comparison, the algorithm's performance was evaluated against the PanCan model at nodule and participant levels using the area under the receiver operating characteristic curve (AUC), among other parameters. AUC sums up how well a model can produce relative scores to discriminate between positive or negative instances across all classification thresholds.

In the pooled cohort, the deep learning model achieved AUCs of 0.98, 0.96, and 0.94 for cancers diagnosed within one year, two years, and throughout screening, respectively, compared to PanCan at 0.98, 0.94, and 0.93.

For indeterminate nodules (129 malignant, 2,086 benign), the deep learning model significantly outperformed PanCan across all timeframes with AUCs of 0.95, 0.94, 0.90 vs. 0.91, 0.88, 0.86. For the cancers size-matched to benign nodules, (180 malignant, 360 benign), the deep learning model's AUC was 0.79 versus PanCan at 0.60.

At 100% sensitivity for cancers diagnosed within 1 year, the deep learning model classified 68.1% of benign cases as low risk compared to 47.4% using the PanCan model, representing a 39.4% relative reduction in false positives.

"Deep learning algorithms can assist radiologists in deciding whether follow-up imaging is needed, but prospective validation is required to determine the clinical applicability of these tools and to guide their implementation in practice," Dr. Antonissen said. "Reducing false positive results will make lung cancer screening more feasible."

"External Test of a Deep Learning Algorithm for Pulmonary Nodule Malignancy Risk Stratification Using European Screening Data." Collaborating with Dr. Antonissen, were Kiran Vaidhya Venkadesh, Ph.D., Renate Dinnessen, M.Sc., Ernst Th. Scholten, M.D., Ph.D., Zaigham Saghir, M.D., Ph.D., Mario Silva, M.D., Ph.D., Ugo Pastorino, M.D., Grigory Sidorenkov, Ph.D., Marjolein A. Heuvelmans, M.D., Ph.D., Geertruida H. de Bock, Ph.D., Firdaus A. A. Mohamed Hoesein, M.D., Ph.D., Pim A. de Jong, M.D., Ph.D., Harry J. M. Groen, M.D., Ph.D., Rozemarijn Vliegenthart, M.D., Ph.D., Hester A Gietema, M.D., Ph.D., Mathias Prokop, M.D., Ph.D., Cornelia Schaefer-Prokop, M.D., Ph.D., and Colin Jacobs, Ph.D., for the NELSON-POP consortium. This study received funding from the Dutch Cancer Society (KWF Kankerbestrijding, project number 9037) and Siemens Healthineers.

Copyright ©2026 Radiological Society of North America (RSNA)

Radiology is edited by co-interim editors Vicky Goh, M.B.B.Ch., King's College London, U.K., and Kathryn Fowler, M.D., University of California San Diego, California, and owned and published by the Radiological Society of North America, Inc. (https://pubs.rsna.org/journal/radiology)

RSNA is an association of radiologists, radiation oncologists, medical physicists and related scientists promoting excellence in patient care and health care delivery through education, research and technologic innovation. The Society is based in Oak Brook, Illinois. (RSNA.org)

For patient-friendly information on lung cancer screening, visit *RadiologyInfo.org*.

Images (JPG, TIF):

Copyright ©2026 Radiological Society of North America (RSNA)

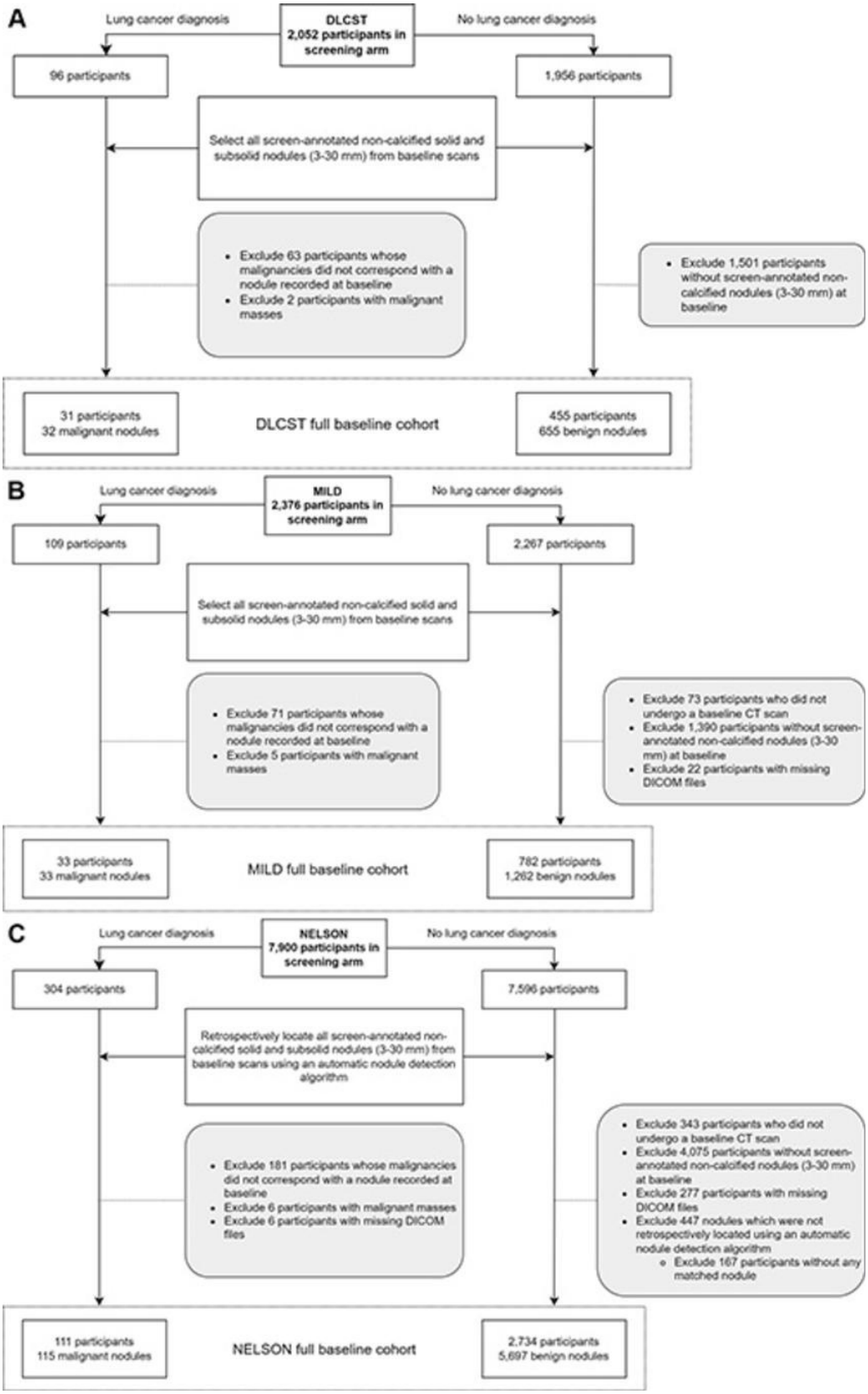


Figure 1. Flowcharts show inclusion and exclusion criteria for participants and pulmonary nodules from **(A)** the Danish Lung Cancer Screening Trial (DLCST), **(B)** the Multicentric Italian Lung Detection (MILD) trial, and **(C)** the Dutch-Belgian Lung Cancer Screening Trial (NELSON). DICOM = Digital Imaging and Communications in Medicine.
High-res (TIF) version

Copyright ©2026 Radiological Society of North America (RSNA)

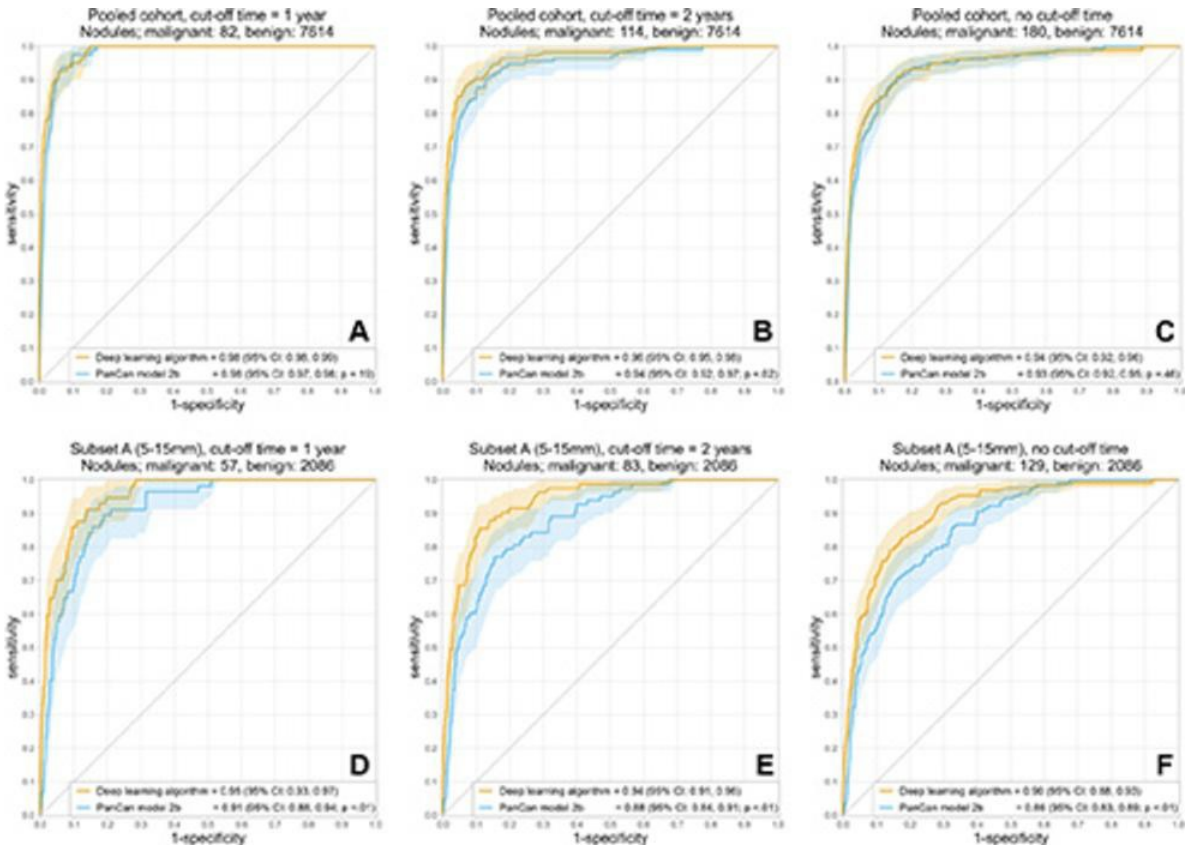


Figure 2. Receiver operating characteristic curves with 95% CIs show a comparison of the deep learning algorithm and the Pan-Canadian Early Detection of Lung Cancer (PanCan) model in distinguishing malignant from benign nodules. Receiver operating characteristic curve results for the pooled cohort including **(A)** cancers diagnosed within 1 year from baseline, **(B)** cancers diagnosed within 2 years from baseline, and **(C)** all cancers. Receiver operating characteristic curve results for subset A (5–15-mm nodules) including results for subset A (5–15 mm nodules), including **(D)** cancers diagnosed within 1 year from baseline, **(E)** cancers diagnosed within 2 years from baseline, and **(F)** all cancers. Areas under the receiver operating characteristic curve with 95% CIs are on each legend.
High-res (TIF) version

Copyright ©2026 Radiological Society of North America (RSNA)

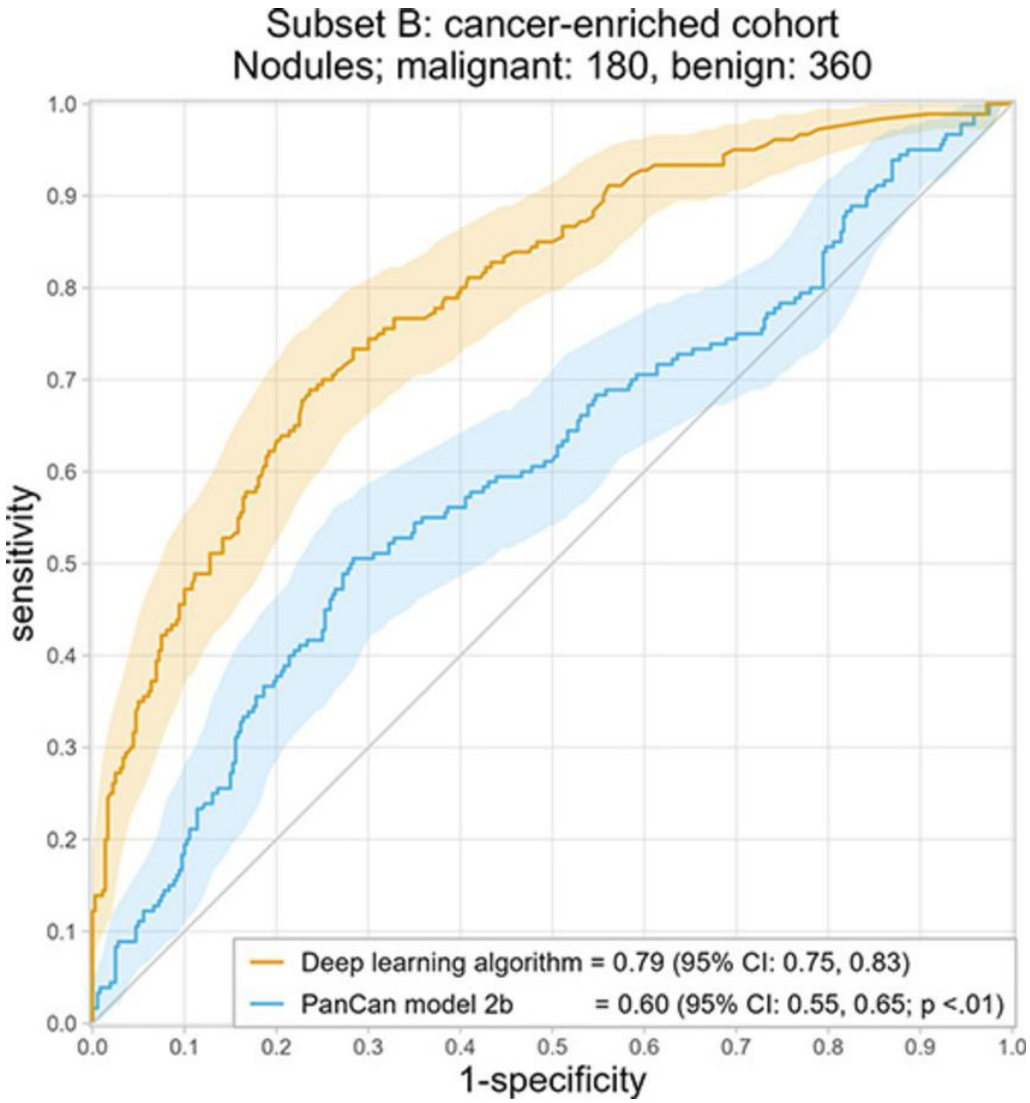


Figure 3. Receiver operating characteristic curves with 95% CIs show a comparison of the performance of the deep learning algorithm and the Pan-Canadian Early Detection of Lung Cancer (PanCan) model in distinguishing malignant from benign nodules in subset B, which includes all cancers with size-matched benign nodules. The areas under the receiver operating characteristic curve with associated 95% CIs in parentheses are on each legend.
High-res (TIF) version

Copyright ©2026 Radiological Society of North America (RSNA)

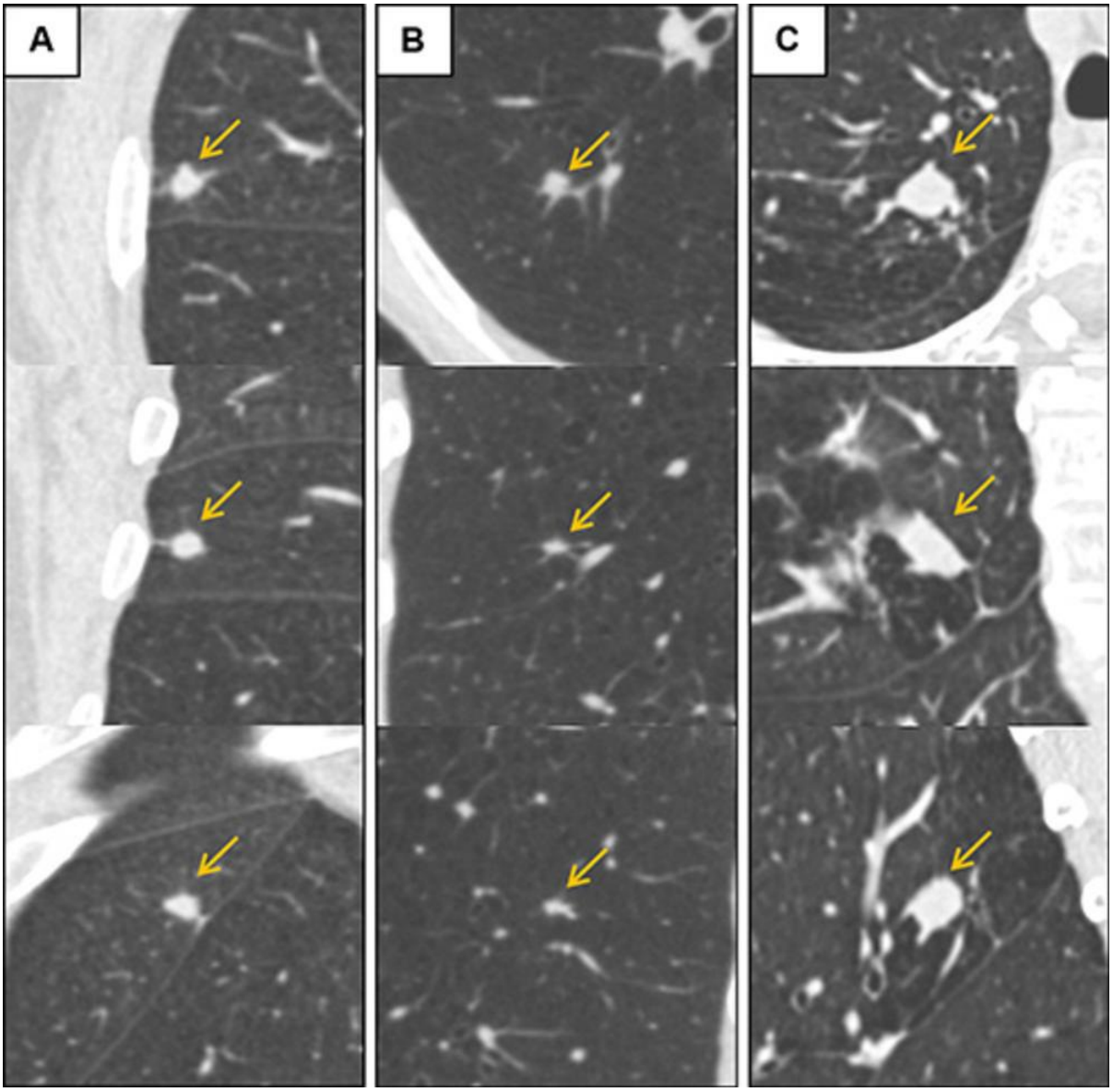


Figure 4. Low-dose CT images show examples of screen-detected pulmonary nodules (arrows) where the deep learning algorithm provides a more accurate malignancy risk estimation than the Pan-Canadian Early Detection of Lung Cancer (PanCan) model on axial (top), coronal (middle), and sagittal (bottom) planes. **(A)** Image shows a 9.7-mm malignant nodule (arrows) with a high deep learning risk score (32.3%) and low PanCan risk score (3.2%) in a 74-year-old male participant diagnosed with squamous cell carcinoma. **(B)** Image shows a 6.8-mm malignant nodule (arrows) with a high deep learning risk score (15.9%) and low PanCan risk score (1.2%) in a 71-year-old male participant diagnosed with adenocarcinoma. **(C)** Image shows a 19- mm benign nodule (arrows) with a low deep learning risk score (4.7%) and high PanCan risk score (32.7%) in a 50-year-old female participant. Additional PanCan input features used in the model were retrieved from original trial records, as follows: **(A)** negative for family history of lung cancer, negative for emphysema, negative for spiculation, negative for upper lobe location, nodule count: four; **(B)** negative for family history of lung cancer, negative for emphysema, negative for spiculation, negative for upper lobe location, nodule count: two; **(C)** negative for family history of lung cancer, positive for emphysema, negative for spiculation, positive for upper lobe location, nodule count: one.
High-res (TIF) version

Resources:

Study abstract

Copyright ©2026 Radiological Society of North America (RSNA)