Radiology: Imaging Cancer

ORIGINAL RESEARCH

A Self-Supervised Foundation Model Based on Three-Dimensional Chest CT Scans for Lung Cancer Diagnosis and Prognosis Prediction

Junxian Li, PhD[1] • *Yuchen Xing, BS*[2] • *Ximin Gao, BS*[2] • *Zhaoxiang Ye, PhD*[3] • *Meng Wang, PhD*[4] • *Fengju Song, PhD*[2]

Author affiliations, funding, and conflicts of interest are listed at the end of this article.

Radiology: Imaging Cancer 2026; 8(2):e250360 • https://doi.org/10.1148/rycan.250360 • Content codes: OI CH CT

Purpose: To develop a self-supervised chest CT foundation model and evaluate its performance in lung cancer clinical tasks.

Materials and Methods: In this retrospective multicenter study, the authors developed the Unified CT-Based Lung Cancer Imaging Foundation (UCLIF) model using self-supervised learning on 33 901 three-dimensional chest CT scans acquired between June 1958 and February 2019. The model was pretrained with a contrastive masked image modeling task and then fine-tuned for lung cancer histologic subtype classification, cancer staging, survival, and recurrence prediction using multicenter patient datasets. Histopathology, TNM stage, and follow-up outcomes served as reference standards. UCLIF was compared with mainstream deep learning and machine learning algorithms, and model performance was assessed by accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC); superiority was tested using the DeLong test.

Results: A total of 656 patients were included for downstream evaluation (mean age, 68.55 years ± 10.01 [SD]; 450 males). Compared with self-supervised pretraining on natural images or single tumor regions, UCLIF achieved superior performance (DeLong test, $P < .001$) and provided high AUCs for histologic subtype (AUC, 0.96 [95% CI: 0.88, 1.00]; AUC, 0.82 [95% CI: 0.60, 0.98]; and AUC, 0.93 [95% CI: 0.80, 0.99] for adenocarcinoma, large cell lung cancer, and squamous cell carcinoma, respectively), cancer staging (AUC, 0.95 [95% CI: 0.79, 1.00]; AUC, 0.99 [95% CI: 0.96, 1.00]; AUC, 0.92 [95% CI: 0.74, 1.00]; and AUC, 0.91 [95% CI: 0.78, 1.00] for stages I–IV, respectively), survival (AUC, 0.97 [95% CI: 0.92, 1.00]; AUC, 0.90 [95% CI: 0.72, 0.98]; and AUC, 0.90 [95% CI: 0.77, 1.00] for 1-, 3-, and 5-year survival, respectively), and recurrence (AUC, 0.95; 95% CI: 0.88, 0.99).

Conclusion: The UCLIF model accurately predicted lung cancer histologic subtype, stage, survival, and recurrence.

Supplemental material is available for this article.

© RSNA, 2026

Lung cancer is the leading cause of cancer-related mortality worldwide (1). Non–small cell lung cancer (NSCLC) accounts for approximately 85% of all lung cancer cases (2). In recent years, large-scale lung cancer screening trials, such as the National Lung Screening Trial and the Nederlands-Leuvens Longkanker Screenings Onderzoek, show that repeated low-dose CT screening (approximately 1–2 mSv per scan) of high-risk populations (3–6) can considerably reduce lung cancer mortality (7). Early-stage lung cancer (stages I and II) has a more favorable prognosis, underscoring the importance of early detection. However, the shortage of radiologists and the lack of obvious clinical symptoms in early disease cause delayed diagnosis and missed lesions, further exacerbating the high mortality. Improving early diagnostic and staging capability is therefore crucial for better patient outcomes.

However, important challenges remain in the clinical diagnosis and treatment of lung cancer. The traditional TNM staging system is a key indicator for prognosis prediction and treatment planning (8) but has limitations: even within the same stage, patients can show substantial differences in recurrence and survival (9). Some patients with stage I cancer experience rapid recurrence or metastasis shortly after surgery, whereas other patients with advanced-stage cancer achieve long-term survival (10). Pathology classification from biopsy or surgery, although essential for decisions on surgical strategies and adjuvant therapy (11), requires invasive sampling and may not capture intra- and peritumoral heterogeneity, limiting treatment precision.

CT is central to lung cancer screening, diagnosis, staging, and treatment evaluation. In the preoperative setting, quantitative analysis of CT scans offers a noninvasive way to assess tumor pathology, staging, and prognosis. Radiomics-based studies show that quantitative image features can predict lymph node metastasis (12) and prognosis (13) through predictive models, thereby supporting personalized treatment strategies. Nevertheless, several factors hinder the widespread clinical application of artificial intelligence models. For radiomics methods, low reproducibility and limited generalizability are barriers due to variations in acquisition parameters, reconstruction algorithms, scanner vendors, patient populations, and feature extraction and validation protocols. Moreover, the limited availability and high cost of expert-annotated datasets, inconsistent imaging quality, and the single-task focus of most deep learning (DL) models limit their ability to handle the complex, multidimensional aspects of lung cancer diagnosis and treatment planning.

To address these challenges, self-supervised learning (SSL) has emerged as a powerful approach, enabling models to learn meaningful feature representations from vast amounts of unlabeled data (14). SSL-based models have demonstrated comparable or better performance than traditional supervised learning models, reducing reliance on labeled datasets. The advent of foundation models has further propelled artificial intelligence research, achieving remarkable performance across various domains, including natural language processing, computer vision, and medical imaging (14). These models have been successfully

This copy is for personal use only. To order printed copies, contact *reprints@rsna.org*

Abbreviations

AUC = area under the receiver operating characteristic curve, DL = deep learning, ML = machine learning, NSCLC = non–small cell lung cancer, SSL = self-supervised learning, TCGA-LUSC = The Cancer Genome Atlas Lung Squamous Cell Carcinoma program, TCIA = The Cancer Imaging Archive, UCLIF = Unified CT-Based Lung Cancer Imaging Foundation

Summary

The self-supervised Unified CT-Based Lung Cancer Imaging Foundation model, trained on 33 901 chest scans, surpassed state-of-the-art algorithms in predicting lung cancer histologic subtype, stage, survival, and recurrence while reducing annotation demands.

Key Points

- The Unified CT-Based Lung Cancer Imaging Foundation (UCLIF) model performed well in predicting lung cancer histologic subtype classification, staging, survival, and recurrence (area under the receiver operating characteristic curve: 0.96 for adenocarcinoma, 0.95 for stage I tumors, and 0.97 for 1-year survival).
- The UCLIF model outperformed traditional self-supervised pretraining algorithms by extracting more comprehensive and meaningful features from three-dimensional chest CT scans, leading to better diagnostic and prognostic predictions ($P < .001$).
- The model showcased strong scalability and interpretability across multiple clinical tasks, reducing the need for extensive annotated data.

Keywords

Lung Cancer, CT, Foundation Model, Diagnosis, Classification

applied to tasks such as automatic radiology report generation (15), fluorescence microscopy image restoration (16), and multimodal medical imaging analysis (17). SSL-based foundation models can reach competitive performance using only a small fraction of labeled data; models such as soma and neurite density imaging (18) and tree-guided convolutional neural networks (19) have matched or outperformed fully supervised models while using about 1% of annotated data and achieving an area under the receiver operating characteristic curve (AUC) of up to 1.0 in cancer prediction tasks. Despite advancements, foundation models still struggle with interpretable three-dimensional CT-based lung cancer diagnosis and prognosis due to challenges related to data management and model architecture.

Accordingly, the aim of our study was to develop an SSL foundation model based on chest CT scans and evaluate its performance in various clinical tasks.

Materials and Methods

In this study, we proposed the Unified CT-Based Lung Cancer Imaging Foundation (UCLIF) model, which adopts a DL pretraining and fine-tuning framework. UCLIF first learns generalizable chest CT representations from unlabeled data and is then fine-tuned for specific downstream predictive tasks. The model attends to lesions, whole-lung structure, and perilesional context to extract features relevant for lung cancer diagnosis, staging, and prognosis. Using the UCLIF model, we performed predictive analyses for lung cancer histologic subtype classification, cancer staging, and patient survival and recurrence and compared performance with other advanced DL and traditional machine learning (ML) methods. This unified framework may support multiple stages of lung cancer care and provide more precise, multidimensional decision support for clinicians. The retrospective secondary analysis complied with the Declaration of Helsinki and was approved by the ethics committee of Tianjin Medical University Cancer Institute & Hospital (approval no. EK20250156). All datasets, collected from June 1958 to February 2019, were obtained from publicly available repositories and were de-identified in accordance with each repository's data-use policies. CT scans from The Cancer Imaging Archive (TCIA) are fully de-identified under Health Insurance Portability and Accountability Act standards, and no protected health information was available to the investigators. Because only de-identified public data were used, the ethics committee waived informed consent. Before analysis, we verified that Digital Imaging and Communications in Medicine headers contained no protected health information and removed any residual private tags to ensure confidentiality. This study did not overlap with prior publications on these datasets.

Datasets for Stage I Pretraining

In this retrospective study, we pretrained UCLIF using 33 901 chest CT scans from 27 092 adult patients drawn from 15 publicly available TCIA datasets. These datasets comprise multimodal CT scans, including low-dose CT, spiral CT, four-dimensional fan-beam CT, and four-dimensional cone-beam CT, spanning diverse imaging modalities and clinical information. Advanced multimodal approaches were used in these datasets, integrating imaging with genomics, proteomics, and pathology to improve diagnostic accuracy and treatment response assessment. The datasets feature a range of imaging parameters and are designed for specific applications, such as the evaluation of lung function and monitoring of tumor motion during respiration. The research objectives within these datasets are extensive, ranging from lung cancer screening and diagnosis to the evaluation of treatment efficacy and immune responses. For this study, all adult patients with available chest CT imaging were included, and we extracted CT scans from these datasets to serve as pretraining input data. In accordance with the 2024 Checklist for Artificial Intelligence in Medical Imaging, this TCIA cohort was used solely for stage I pretraining and was not used for model evaluation. Our primary patients were those with NSCLC, and all chest CT scans used were anonymized to ensure patient confidentiality. Detailed information regarding these datasets, including their names, data characteristics, manufacturer, and data source, is provided in Tables S1 and S2.

Datasets for Stage II Downstream Task Fine-Tuning

To comprehensively assess the performance of our foundation model in lung cancer analysis, we used diverse datasets and conducted evaluations across multiple downstream tasks, including histopathology classification, tumor staging, and survival and recurrence prediction. To prevent data leakage, we excluded all datasets used for pretraining the foundation model. Moreover, to protect patient privacy, all Digital Imaging and Communications in Medicine images were anonymized, with patient information removed. The detailed information for each dataset used in downstream task fine-tuning is shown in Appendix S1 and Tables S2 and S3.

Histologic subtype classification

In this task, we classified the CT scans of patients into three distinct histologic lung cancer subtypes: adenocarcinoma, large cell lung cancer, and squamous cell carcinoma. The NSCLC Radiogenomics dataset (20), including 196 CT scans with histologic subtype annotations, was used as the training set in adenocarcinoma and squamous cell carcinoma classification to optimize our classification model. For large cell lung cancer classification, because of the scarcity of patients with this type in the NSCLC Radiogenomics dataset, we used the NSCLC-Radiomics (21) dataset, with 317 CT scans serving as the training set. Additionally, we extracted 19 CT scans from the NSCLC-Radiomics-Interobserver1 (22) dataset as the test set for all three subtypes.

Stage classification

To comprehensively evaluate model performance, we categorized patient CT scans into four clinical stages—stages I–IV—according to the TNM classification system. We used the NSCLC Radiogenomics dataset, with 146 CT scans, as the training set and the NSCLC-Radiomics-Genomics dataset (23), with 87 CT scans, as the test set.

Patient survival prediction

We categorized patient survival time into three intervals: 1 year, 3 years, and 5 years. We extracted CT scans with survival data from the NSCLC Radiogenomics dataset to serve as the training set, and we included 32 CT scans with detailed survival information from The Cancer Genome Atlas Lung Squamous Cell Carcinoma (TCGA-LUSC) dataset (24) as the test set.

Patient recurrence prediction

In this task, we used 217 CT scans from the NSCLC Radiogenomics dataset as the training set and 86 CT scans from the TCGA-LUSC dataset as the test set. The final outcome was determined according to whether lung cancer had recurred.

Data Preprocessing

The CT data we used in this study were obtained from publicly available datasets and approved by the hospital ethics committee. Patient privacy protection complied with relevant regulatory requirements. The original scans were stored in Digital Imaging and Communications in Medicine format with a matrix size of 512 × 512, ensuring continuity across sections without missing or fragmented sections. To mitigate differences in resolution and section thickness caused by variations in scanning equipment and imaging parameters, we first applied cubic B-spline interpolation to resample the CT volumetric data to an isotropic voxel size of 1 mm × 1 mm × 1 mm. Subsequently, we performed intensity normalization within a window width of −1400 HU to 400 HU using min-max normalization, mapping CT values to the range [0,1]. During the region of interest construction, two thoracic radiologists (Z.Y. and M.W., each with 10 years' experience) delineated lesion boundaries in consensus; discrepancies were adjudicated by a senior thoracic oncologic radiologist (20 years' experience). Surrounding tissue information was evaluated within an 8-mm peritumoral ring around the lesion boundary.

UCLIF Model Architecture and Training Details

To enhance SSL for extracting visual features, we introduced an autoencoder rooted in the Vision Transformer framework. This innovative model, designated the UCLIF model, is specifically designed to tackle three-dimensional medical imaging challenges. The architecture is composed of two principal modules: a Vision Transformer encoder and a complementary Transformer decoder. During pretraining, the encoder receives image patches with selective masking and generates advanced feature representations that encapsulate the intricate visual characteristics of the original medical scans. Concurrently, the decoder uses masked tokens as placeholders, drawing on the full set of encoder outputs for contextual guidance. The primary function of the decoder is to reconstruct the obscured portions of the scan, effectively approximating the original content through a self-supervised image completion strategy. This approach enables the model to learn robust visual representations without the need for explicit labels.

We conducted performance comparisons between the UCLIF model and two pretrained baseline methods, namely SSL-ImageNet and SSL-Lung. Although each method adopted a distinct pretraining strategy, they maintained an identical architecture and followed a uniform fine-tuning protocol for downstream tasks. The flow of the UCLIF architecture throughout both pretraining and fine-tuning stages is illustrated in Figure S1. Furthermore, the UCLIF model incorporated SSL-ImageNet weights as its initialization baseline and extended the pretraining phase to include lung scans, effectively using a sequential two-stage SSL process that first leverages natural images and then lung scans. An overview of these pretraining strategies is provided in Figure S2 and Appendix S2.

The complete workflow of UCLIF model development is depicted in Figure 1A.

Development of Classifier and Comparison Algorithms

To harness the pretrained ViTAutoEnc for subsequent classification applications, we integrated a dedicated classification module comprising an initial feature selection layer followed by several fully connected layers. In detail, the pretrained encoder generated high-dimensional latent representations, where each scan produced 216 patches represented by 2048-dimensional vectors. We subsequently compressed these high-dimensional features to a lower-dimensional space—256 dimensions per patch—via principal component analysis. Following this reduction, the resulting compact representations (with dimensions of 216 × 256) were flattened into a single feature vector that served as input for a cascade of fully connected layers, sequentially configured with 128, 64, and ultimately two neurons to yield classification predictions. Incorporating principal component analysis not only streamlined the extraction of pertinent features but also enhanced the discriminative power of the classification module by retaining only the most informative elements. This strategy markedly improved both the efficiency and accuracy of the classifier (Fig 1B).

To perform an extensive performance evaluation, we used both DL and ML methods. Initially, we trained three distinct convolutional neural networks—ResNet34 (25), DenseNet121 (26), and ShuffleNet (27)—which had been pretrained on the ImageNet1k dataset and accessed via the timm (PyTorch Image Models) library.

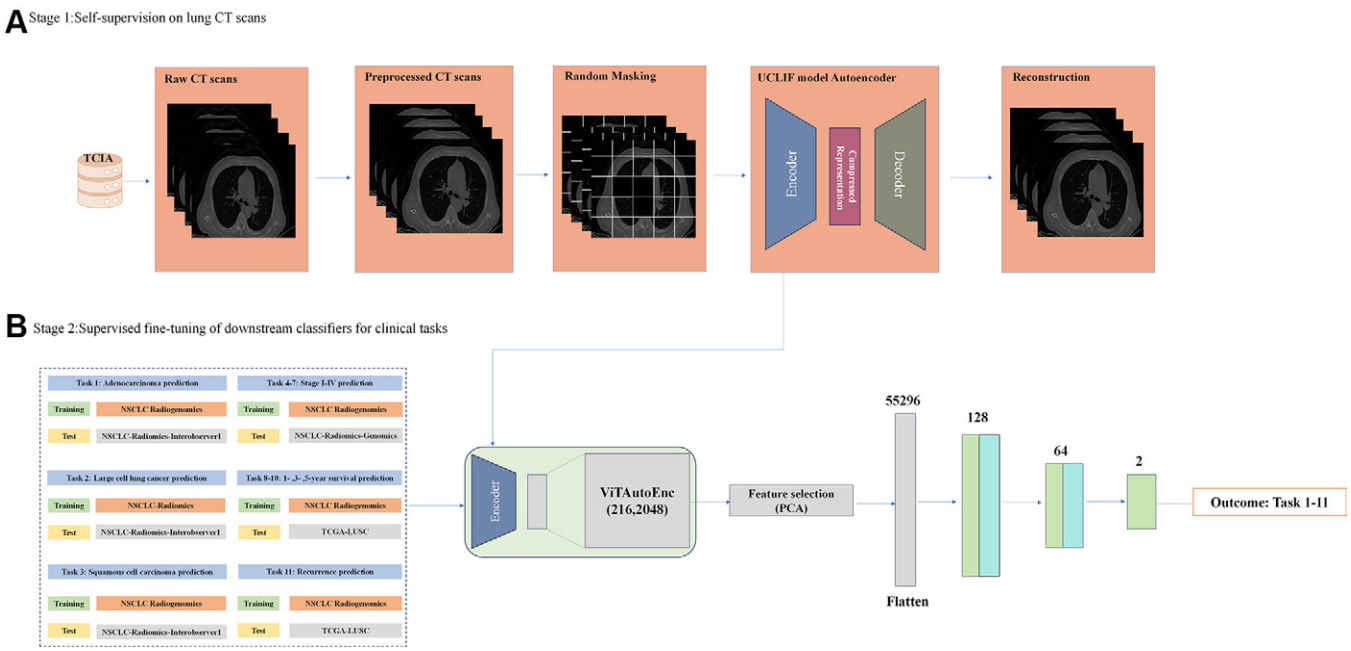

Figure 1: Flowchart of stages I and II. For the development and assessment of the Unified CT-Based Lung Cancer Imaging Foundation (UCLIF) model, a structured approach was followed. **(A)** In stage I, the UCLIF model was built through self-supervised learning, incorporating CT data sourced from The Cancer Imaging Archive (TCIA). **(B)** Stage II involved refining the UCLIF model for specific downstream tasks, and supervised learning was used to facilitate both internal test and external performance evaluation. PCA = principal component analysis, NSCLC = non–small cell lung cancer, TCGA-LUSC = The Cancer Genome Atlas Lung Squamous Cell Carcinoma dataset, ViT = Vision Transformer.

Subsequently, we used features extracted from these convolutional neural networks to construct and assess various conventional ML classifiers. Specifically, we implemented classifiers based on logistic regression, support vector machines, random forest, and eXtreme gradient boosting. We systematically compared the performance of these classifiers using the extracted features, providing a comprehensive evaluation of the effectiveness of each model for specific tasks. We executed all experiments using two Nvidia RTX 3090 graphics processing units. Detailed information regarding the convolutional neural network and ML algorithms we used in this study is available in Appendixes S3 and S4.

For the classification task, we trained the model on the designated training set using a cross-entropy loss function in conjunction with the AdamW optimizer. We set an initial learning rate of $1e^{-5}$, with a cosine decay schedule governing the learning rate throughout the training process. We performed training with a batch size of 36, extended for up to 300 epochs. At the end of each epoch, we evaluated model performance on the test set, and the weights that achieved the highest AUC were preserved as checkpoints for both internal and external test sets.

To evaluate the scalability of the pretrained UCLIF model across diverse challenges in pulmonary medicine, we designed several binary classification tasks: histologic subtype classification, tumor staging, survival prediction, and recurrence prediction.

Statistical Analysis

Because this was a retrospective study using fixed public imaging cohorts, no formal a priori sample size calculation was performed; instead, we included all eligible patients to maximize statistical power and model stability. All statistical analyses were performed by J.L. and Y.X. To develop a robust and generalizable model, we used 10-fold cross-validation to evaluate the performance of each DL and ML model on the test subset. For each fold, we computed the true-positive, true-negative, false-positive, and false-negative results. According to these values, we constructed confusion matrices and calculated key performance metrics, including accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. Additionally, we estimated 95% CIs for the cross-validated AUC of each predictive model using bootstrapping with 1000 iterations.

For time-to-event end points (overall survival and recurrence), we used multivariable Cox proportional hazards models; we right-censored patients without events at the last follow-up, and we assessed the proportional hazards assumption using Schoenfeld residuals. We adjusted for baseline covariates (age, sex, smoking status) by concatenating them to the UCLIF representation and optimizing a Cox partial-likelihood loss in a deep survival model and then including them alongside the UCLIF risk score in a multivariable Cox model (age centered and standardized; sex as a binary indicator; smoking as a categorical variable—never, former, or current—via indicator coding). The median risk value was used as the reference to categorize patients into low- and high-risk groups. Adjusted hazard ratios and performance metrics are reported alongside unadjusted results. For the analytical strategy, we developed models on the basis of the NSCLC Radiogenomics dataset and evaluated them on the TCGA-LUSC dataset as an external validation without additional training or fine-tuning; we estimated all preprocessing parameters and the risk-score threshold within the training folds, then fixed them for external application to avoid information leakage, under the assumptions of noninformative censoring and proportional hazards. Furthermore, to evaluate the model's ability in predicting patients' survival, we also explored a deep survival learning approach (DeepSurv) under the same data splits and tuning protocol (Appendix S4).

Table 1: Comprehensive Count of CT Scans Used in Each Task

Task Description	Training Set	Test Set	Total
Stage I: SSL on chest CT scans	27 121	6780	33 901
Stage II: Supervised fine-tuning for clinical tasks			
Task 1: Adenocarcinoma prediction			
Adenocarcinoma	166	10	176
Other type	30	9	39
Task 2: Large cell lung cancer prediction			
Large cell lung cancer	114	3	117
Other type	203	16	219
Task 3: Squamous cell carcinoma prediction			
Squamous cell carcinoma	30	6	36
Other type	166	13	179
Task 4: Stage I prediction			
Stage I	94	15	109
Other stage	52	72	124
Task 5: Stage II prediction			
Stage II	26	54	80
Other stage	120	33	153
Task 6: Stage III prediction			
Stage III	22	13	35
Other stage	124	74	198
Task 7: Stage IV prediction			
Stage IV	4	5	9
Other stage	142	82	224
Task 8: 1-year survival prediction			
Deceased	22	22	44
Alive	38	10	48
Task 9: 3-year survival prediction			
Deceased	45	28	73
Alive	15	4	19
Task 10: 5-year survival prediction			
Deceased	57	28	85
Alive	3	4	7
Task 11: Recurrence prediction			
No	168	77	245
Yes	49	9	58

Note.—Data are numbers of instances. SSL = self-supervised learning.

In addition to survival analyses, we performed the DeLong test to assess the significance of differences in AUCs between classification models. We conducted this analysis using the pROC package in R statistical software (version 4.3.2; R Foundation), with $P < .05$ considered statistically significant.

Code Availability

The code used for model development and data analysis can be made available from the corresponding author upon reasonable request for academic, noncommercial use.

Results

Dataset Characteristics

In this study, we included 656 patients (median age, 68.55 years ± 10.01 [SD]; 450 male) for downstream tasks to evaluate the model. A detailed breakdown of the CT scans used in this two-stage study is provided in Table 1 and Figure S3. In stage I, we used 33 901 chest CT scans for SSL, comprising 27 121 scans for training and 6780 for testing. Following this pretraining phase, stage II involved supervised fine-tuning for multiple clinical end points, each supported by a distinct combination of training and test sets. We trained and tested the model for histologic subtype classification tasks using the NSCLC Radiogenomics and NSCLC-Radiomics datasets and the NSCLC-Radiomics-Interobserver 1 dataset, respectively. Specifically, the NSCLC Radiogenomics dataset contained 166 adenocarcinoma and 30 squamous cell carcinoma CT scans; the NSCLC-Radiomics dataset contained 114 large cell lung cancer and 203 other-type CT scans; and the NSCLC-Radiomics-Interobserver 1 dataset contained 10 adenocarcinoma, three large cell lung cancer, and six squamous cell carcinoma CT scans. We trained the model for stage I–IV prediction tasks on the NSCLC Radiogenomics dataset and then tested the model on the NSCLC-Radiomics-Genomics dataset, with different distributions of stage-specific and other-stage cancer CT scans. Stage I training included 94 stage I and 52 other-stage cancer CT scans, and validation included 15 stage I and 72 other-stage cancer CT scans. Stage II training included 26 stage II and 120 other-stage cancer CT scans, and validation included 54 stage II and 33 other-stage cancer CT scans. Stage III training included 22 stage III and 124 other-stage cancer CT scans, and validation included 13 stage III and 74 other-stage cancer CT scans. Stage IV training included four stage IV and 142 other-stage cancer CT scans, and validation included five stage IV and 82 other-stage cancer CT scans. We trained the model for survival predictions on the NSCLC Radiogenomics dataset (CT scans for 22 deceased and 38 alive, 45 deceased and 15 alive, and 57 deceased and three alive patients for 1-, 3-, and 5-year survival, respectively) and then tested the model on the TCGA-LUSC dataset with corresponding outcome distributions (CT scans for 22 deceased and 10 alive, 28 deceased and four alive, and 28 deceased and four alive patients, respectively). We trained the model for recurrence prediction (no recurrence vs recurrence) on the NSCLC Radiogenomics dataset (CT scans from 168 patients without and 49 with recurrence) and then tested the model on the TCGA-LUSC dataset (CT scans from 77 patients without and nine with recurrence).

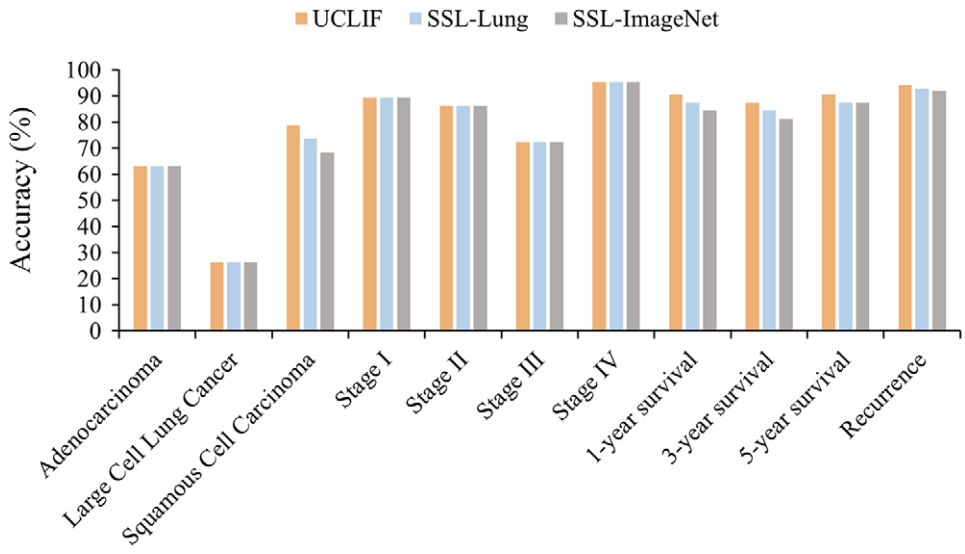

Figure 2: Bar graph of performance comparison of the Unified CT-Based Lung Cancer Imaging Foundation (UCLIF) model and self-supervised learning (SSL)–based methods across classification tasks. Across all tasks, the UCLIF model consistently outperformed the SSL-based methods, achieving higher accuracy in all categories.

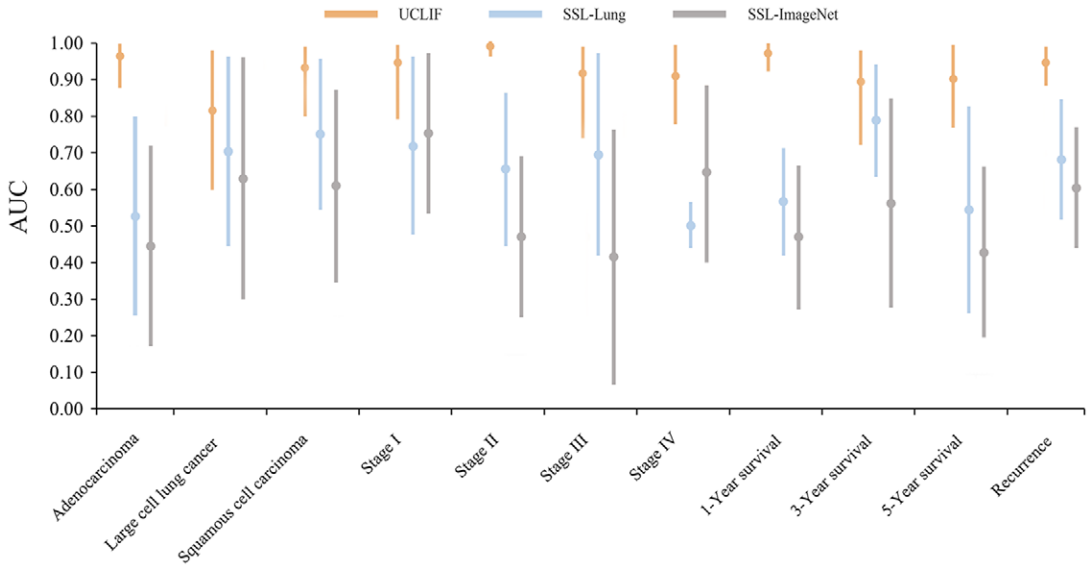

Figure 3: Dot graph of comparison of performance between the Unified CT-Based Lung Cancer Imaging Foundation (UCLIF) and self-supervised learning (SSL)–based methods. The figure presents the area under the receiver operating characteristic curve (AUC; dots) along with 95% CIs (lines) for the UCLIF model, SSL-Lung, and SSL-ImageNet across multiple classification tasks. The UCLIF model consistently achieved higher AUC values than the SSL-based approaches, indicating superior predictive performance. Pairwise comparisons of AUCs between UCLIF and SSL-based methods were performed using the DeLong test and were statistically significant (all $P < .001$). The CIs for the UCLIF model are also narrower for several tasks, suggesting greater model stability.

Assessment of UCLIF Model Performance for Histologic Subtype Classification and Tumor Staging

We evaluated the histologic subtype classification performance of the proposed UCLIF model in comparison with that of SSL-Lung and SSL-ImageNet. Across different histologic classifications and staging tasks, the UCLIF model achieved equal or superior accuracy compared with the SSL baselines, with the degree of improvement varying across specific subtypes and stages (large cell lung cancer: 26% [five of 19]; adenocarcinoma: 63% [12 of 19]; squamous cell carcinoma: 79% [15 of 19]) (Fig 2). Additionally, we observed higher classification confidence for the UCLIF model, with AUCs and 95% CIs consistently exceeding those of SSL-based methods (Fig 3). For instance, the model achieved AUCs of 0.96 (95% CI: 0.88, 1.00) for adenocarcinoma, 0.82 (95% CI: 0.60, 0.98) for large cell lung cancer, and 0.93 (95% CI: 0.80, 0.99) for squamous cell carcinoma, outperforming the SSL baselines. From a clinical evaluation perspective, the UCLIF model also achieved superior performance, with the model exhibiting a strong specificity (92%; 12 of 13) and positive predictive value (75%; three of four) for squamous cell carcinoma (Tables 2, 3).

In the tumor staging task, the UCLIF model achieved accuracies of 89% (17 of 19), 86% (25 of 29), 72% (21 of 29),

Table 2: Comparison of the Performance of UCLIF Model on Clinical Tasks of Varying Difficulty in Test Datasets

	UCLIF Model				
Outcome	Accuracy	Sensitivity	Specificity	NPV	PPV
Adenocarcinoma	63 (12/19)	56 (5/9)	70 (7/10)	64 (7/11)	63 (5/8)
Large cell lung cancer	26 (5/19)	19 (3/16)	67 (2/3)	13 (2/15)	75 (3/4)
Squamous cell carcinoma	79 (15/19)	50 (3/6)	92 (12/13)	80 (12/15)	75 (3/4)
Stage I	89 (17/19)	67 (2/3)	94 (15/16)	94 (15/16)	67 (2/3)
Stage II	86 (75/87)	60 (12/20)	94 (63/67)	89 (63/71)	75 (12/16)
Stage III	72 (63/87)	72 (51/71)	75 (12/16)	38 (12/32)	93 (51/55)
Stage IV	95 (83/87)	60 (3/5)	98 (80/82)	98 (80/82)	60 (3/5)
1-Year survival	91 (29/32)	96 (22/23)	78 (7/9)	88 (7/8)	92 (22/24)
3-Year survival	88 (7/8)	50 (3/6)	96 (25/26)	89 (25/28)	75 (3/4)
5-Year survival	91 (29/32)	75 (3/4)	93 (26/28)	96 (26/27)	60 (3/5)
Recurrence	94 (81/86)	56 (5/9)	99 (76/77)	95 (76/80)	83 (5/6)

Note.—Data are percentages with numerators/denominators in parentheses. NPV = negative predictive value, PPV = positive predictive value, UCLIF = Unified CT-Based Lung Cancer Imaging Foundation.

Table 3: Comparison of the Performance of Different Algorithms on Clinical Tasks of Varying Difficulty in Test Datasets

	SSL-Lung					SSL-ImageNet				
Outcome	Accuracy	Sensitivity	Specificity	NPV	PPV	Accuracy	Sensitivity	Specificity	NPV	PPV
Adenocarcinoma	63 (12/19)	55 (5/9)	70 (7/10)	64 (7/11)	63 (5/8)	63 (12/19)	56 (5/9)	70 (7/10)	64 (7/11)	63 (5/8)
Large cell lung cancer	26 (5/19)	19 (3/16)	67 (2/3)	13 (2/15)	75 (3/4)	26 (5/19)	19 (3/16)	67 (2/3)	13 (2/15)	75 (3/4)
Squamous cell carcinoma	74 (14/19)	50 (3/6)	85 (11/13)	79 (11/14)	60 (3/5)	68 (13/19)	50 (3/6)	77 (10/13)	77 (10/13)	50 (3/6)
Stage I	89 (17/19)	67 (2/3)	94 (15/16)	94 (15/16)	67 (2/3)	89 (17/19)	67 (2/3)	94 (15/16)	94 (15/16)	67 (2/3)
Stage II	86 (75/87)	60 (12/20)	94 (63/67)	89 (63/71)	75 (12/16)	86 (75/87)	60 (12/20)	94 (63/67)	89 (63/71)	75 (12/16)
Stage III	72 (63/87)	72 (51/71)	75 (12/16)	38 (12/32)	93 (51/55)	72 (63/87)	72 (51/71)	75 (12/16)	38 (12/32)	93 (51/55)
Stage IV	95 (83/87)	60 (3/5)	98 (80/82)	98 (80/82)	60 (3/5)	95 (83/87)	60 (3/5)	98 (80/82)	98 (80/82)	60 (3/5)
1-Year survival	88 (28/32)	91 (21/23)	78 (7/9)	78 (7/9)	91 (21/23)	84 (27/32)	87 (20/23)	78 (7/9)	70 (7/10)	91 (20/22)
3-Year survival	84 (27/32)	50 (3/6)	92 (24/26)	89 (24/27)	60 (3/5)	81 (26/32)	50 (3/6)	89 (23/26)	89 (23/26)	50 (3/6)
5-Year survival	88 (28/32)	50 (2/4)	93 (26/28)	93 (26/28)	50 (2/4)	88 (28/32)	50 (2/4)	93 (26/28)	93 (26/28)	50 (2/4)
Recurrence	93 (80/86)	56 (5/9)	97 (75/77)	95 (75/79)	71 (5/7)	92 (79/86)	56 (5/9)	96 (74/77)	95 (74/78)	63 (5/8)

Note.—Data are percentages with numerators/denominators in parentheses. SSL = self-supervised learning, NPV = negative predictive value, PPV = positive predictive value.

and 95% (83 of 87) for stage I–IV lung cancer, respectively, significantly outperforming SSL-Lung and SSL-ImageNet ($P < .001$; Fig 2). The model also demonstrated higher AUCs than SSL-Lung and SSL-ImageNet, with AUCs of 0.95 (95% CI: 0.79, 1.00), 0.99 (95% CI: 0.96, 1.00), 0.92 (95% CI: 0.74, 1.00), and 0.91 (95% CI: 0.78, 1.00) for stage I–IV lung cancer, respectively, confirming its robust feature extraction capability in tumor staging classification (Fig 3).

Assessment of UCLIF Model Performance for Survival and Recurrence Prediction

The proposed UCLIF model was further evaluated for its performance in predicting survival time and recurrence, demonstrating superior classification performance across multiple datasets. In the survival time prediction task, the UCLIF model achieved accuracies of 91% (29 of 32), 88% (seven of eight), and 91% (29 of 32) for 1-, 3-, and 5-year survival, respectively (Fig 2). AUC analysis revealed even stronger performance, with the UCLIF model achieving AUCs of 0.97 (95% CI: 0.92, 1.00), 0.90

(95% CI: 0.72, 0.98), and 0.90 (95% CI: 0.77, 1.00) for 1-, 3-, and 5-year survival prediction, respectively (Fig 3). These results exceeded those of SSL-Lung and SSL-ImageNet ($P < .001$), reinforcing the superior discriminative capability of the model.

For recurrence prediction, the UCLIF model achieved an accuracy of 94% (81 of 86), consistently outperforming SSL-based methods (Fig 2). We observed high classification confidence for the UCLIF model (AUC, 0.95; 95% CI: 0.88, 0.99) (Fig 3). Furthermore, the UCLIF model achieved a specificity of 99% (76 of 77), a negative predictive value of 95% (76 of 80), and a robust positive predictive value of 83% (five of six), thereby underscoring the significant advantage of the model in balancing detection rate and predictive accuracy. In contrast, SSL-Lung achieved a sensitivity of only 56% (five of nine), a specificity of 97% (75 of 77), a negative predictive value of 95% (75 of 79), and a positive predictive value of 71% (five of seven), while SSL-ImageNet performance further lagged, with a specificity of 96% (74 of 77) (Tables 2, 3).

Comparison of DL and ML Model Performance

Building on the insights gained from previous analyses, we evaluated how both ML and DL approaches perform in various classification scenarios (Tables S4, S5). Conventional ML algorithms, including eXtreme gradient boosting and support vector machines, generated different results depending on the dataset under consideration, whereas state-of-the-art DL architectures (eg, DenseNet121, ResNet34, and ShuffleNet) each demonstrated distinctive advantages. Across numerous datasets—particularly for molecular biomarker classification and treatment outcome prediction—the UCLIF model consistently attained higher accuracy and AUC than other methods ($P < .001$).

Performance of Multivariable Cox Proportional Hazards Models and DeepSurv Models

Multivariable Cox proportional hazards models revealed that after integrating baseline covariates and risk scores, the UCLIF model could accurately predict patients' risks of death and recurrence with patients at high risk facing significantly greater mortality and recurrence risk than those at low risk in the long-term analyses, whereas the 1-year multivariate result was not statistically significant (Table S6). This lack of significance prompted a brief failure analysis, which suggested that most incorrect 1-year predictions were driven by peritreatment and other unmeasured short-term factors not fully captured by the baseline covariates of the UCLIF model. Moreover, the results of the DeepSurv models indicated that, under the same data splits and outcome definitions, DeepSurv provided discrimination and calibration that were comparable but slightly inferior to those of the Cox-based UCLIF pipeline for both death and recurrence end points (Table S7, Fig S4).

Discussion

In this study, we addressed the increasing workload and annotation burden in CT-based lung cancer care by developing the UCLIF model, an SSL framework that reduces dependence on labeled data. Trained on 33 901 chest CT scans from multiple institutions, the UCLIF model learned cross-contrast context representations and was fine-tuned on clinically relevant tasks. Across multicenter test sets, it achieved AUCs of 0.96, 0.82, and 0.93 for adenocarcinoma, large cell lung cancer, and squamous cell carcinoma, respectively; AUCs of 0.95, 0.99, 0.92, and 0.91 for stages I–IV, respectively; AUCs of 0.97, 0.90, and 0.90 for 1-, 3-, and 5-year survival, respectively; and an AUC of 0.95 for recurrence prediction. In all tasks, the UCLIF model outperformed convolutional neural networks, ML models, and the SSL-Lung and SSL-ImageNet baselines, indicating that domain-specific SSL pretraining on chest CT yields more informative and transferable representations for thoracic oncology (see the Results section for detailed statistics and *P* values).

Chest CT is the cornerstone imaging modality for lung cancer diagnosis and staging and is widely used because of the high incidence of lung cancer (12.4%) (28). However, the increase in CT examinations has placed mounting pressure on radiology services. DL-based artificial intelligence systems have been proposed to alleviate workload and improve diagnostic consistency, with prior studies showing high performance for benign-malignant nodule classification (accuracy, 92.8% [29]) and 5-year recurrence prediction (AUC, 0.817 [30]). Nonetheless, most existing approaches rely on supervised learning and labor-intensive region-of-interest annotations, limiting scalability and increasing overfitting risk with small labeled datasets. By leveraging large-scale unlabeled CT data and modest labeled datasets for fine-tuning, our model addresses these limitations and enables accurate early nodule classification and prognosis prediction with reduced annotation burden.

For histologic subtype classification, the relatively lower performance for large cell lung cancer likely reflects limited samples and class imbalance, whereas other subtypes showed consistently strong metrics, indicating that the UCLIF model can manage complex multiclass tasks with imbalanced data. Han et al (31) used a VGG16-based model to differentiate NSCLC subtypes and achieved an AUC of 0.903 in a single-center cohort, while Ding et al (32) demonstrated the advantages of SSL for histologic subtype prediction using whole-slide images. By pretraining on multi-institutional CT data and evaluating on distinct downstream test sets, our work extends these studies and provides more generalizable evidence for SSL-based lung cancer subtype classification, supporting its use as a complementary tool in routine workflows to accelerate subtype assessment and guide treatment.

In the tumor staging task, the UCLIF model captured radiographic patterns relevant to TNM staging, delivering accurate stage I–IV discrimination. Trabelsi et al (33) previously used a ResNet-50-based model to distinguish T categories within the TNM system, focusing on a single component of tumor staging. In contrast, our model produces a comprehensive staging output within the same framework that also supports multiple prognostic tasks, enhancing potential clinical utility. Its superior performance compared with SSL-Lung and SSL-ImageNet underscores the limitations of generic SSL models trained on natural images in capturing tumor-specific structural and textural features. Consistent with Tayebi Arasteh et al (34), who showed that SSL pretraining on medical-specific data substantially improves disease classification accuracy versus pretraining on natural images, our results support domain-specific SSL as particularly advantageous for thoracic oncology.

The UCLIF model also showed strong prognostic capabilities for survival and recurrence. Existing survival prediction

models using neural networks or semisupervised learning with CT achieved AUCs or accuracies around 0.79–0.89 (35,36), showing room for improvement. In our study, SSL pretraining reduced the need for extensive high-quality labels while enabling simultaneous prediction of short- and long-term survival, which is relevant in research settings where outcome labels are costly. For recurrence prediction, Na et al (37) reported an AUC of 0.860 using a single-hospital prognostic model, and Wang et al (38) reported an AUC of 0.817 for early recurrence in stage I lung cancer using ResNet18. Compared with these models, our model maintained high prognostic performance across multiple external datasets (recurrence AUC, 0.95), supporting more effective formulation and adjustment of treatment strategies and personalized clinical management. These retrospective multicenter results motivate large-scale prospective multicenter studies in broader and more diverse patient populations to further validate generalizability, calibration, and clinical utility before routine deployment. Building on these promising retrospective multicenter results, future large-scale prospective multicenter studies in broader and more diverse patient populations are warranted to further validate the generalizability, calibration, and clinical utility of the UCLIF model before its widespread integration into routine clinical decision support systems.

Our findings are aligned with a growing body of work demonstrating the value of SSL-based foundation models in medical image analysis. Zhou et al (39) trained an SSL foundation model on 1.6 million retinal images. Pai et al (40) applied contrastive learning in pathology, and Bluethgen et al (41) developed a chest radiograph visual-language model, all achieving strong performance across diverse tasks. Extending this concept to three-dimensional chest CT, the UCLIF model functions as a CT-based foundation model for thoracic oncology and demonstrated substantial robustness to domain shifts. It was validated on downstream datasets that differed in acquisition protocols, scanner vendors, reconstruction parameters, and patient populations, whereas conventional supervised models trained from scratch or initialized on natural images showed performance deterioration and variability across external test sets. The consistently superior and stable performance of our model suggests that SSL pretraining on large unlabeled CT cohorts not only reduces dependence on labeled data but also enhances robustness to real-world imaging heterogeneity.

This study had limitations. First, our model relied solely on CT images and did not incorporate complementary clinical, pathologic, or molecular data, which might further improve predictive performance and decision support. Second, like many DL models, the current UCLIF model functions as a "black box" with limited interpretability, potentially hindering clinician trust and adoption. Although we reported subgroup performance across institutions, scanner vendors, and demographic strata to assess generalizability and fairness, residual biases from retrospective site-specific referral patterns and imaging protocols cannot be excluded. Third, all datasets were retrospective, and the model has not been prospectively validated or embedded into routine clinical workflows, so its real-world effectiveness, calibration, and clinical impact remain to be established.

In conclusion, we developed the UCLIF model, an SSL framework for CT-based lung cancer analysis using a two-phase training strategy on natural images and three-dimensional chest CT scans. By reducing reliance on annotated datasets and effectively capturing CT spatial characteristics, our model substantially improved histopathology subtype classification, tumor staging and survival, and recurrence prediction compared with traditional supervised approaches and generic SSL baselines, particularly in small-sample scenarios. Future work should integrate multimodal data sources (including clinical, pathologic, and genomic information), incorporate explainability methods (such as attention-based visualizations, integrated gradients, and perturbation or occlusion tests) to enhance interpretability, and conduct large-scale prospective multicenter studies with consecutive patient enrollment, prespecified operating thresholds, and workflow-embedded deployment. These efforts will be crucial to validate the robustness, fairness, and clinical utility of the UCLIF model; facilitate its integration into triage pipelines, presurgical planning, and multidisciplinary team discussions; and ultimately support personalized, data-driven management of patients with lung cancer.

Author affiliations:

[1] Department of Blood Transfusion, Key Laboratory of Cancer Prevention and Therapy, Tianjin, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin Medical University, Tianjin, China

[2] Department of Cancer Epidemiology and Biostatistics, Key Laboratory of Molecular Cancer Epidemiology, Tianjin, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin Medical University, Huanhu Xi Road, Tiyuan Bei, Hexi District, Tianjin 300060, China

[3] Department of Radiology, National Clinical Research Centre for Cancer; Key Laboratory of Cancer Prevention and Therapy, Tianjin Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin Medical University, Tianjin, China

[4] Department of Lung Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin; Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin Medical University, Tianjin, China

Received July 1, 2025; revision requested August 4; revision received November 27; accepted January 8, 2026.

Address correspondence to: F.S. (email: songfengju@163.com).

Supplemental material: Supplemental material is available at *Radiology: Imaging Cancer* online.

Funding: This study was supported by the Chinese National Key Research and Development Project (grant no. 2021YFC2500404), the introduction of talents and doctoral start-up fund of Tianjin Medical University Cancer Institute & Hospital (grant no. B2317), and Tianjin Key Medical Discipline Construction Project (grant no. TJYXZDXK-3-003A).

Acknowledgments: This work was supported by the high-performance computing platform of Tianjin Medical University.

Author contributions: Junxian Li: Conceptualization, Data curation, Formal analysis, Methodology, Writing - original draft; Yuchen Xing: Data curation, Formal analysis, Writing - original draft; Ximin Gao: Formal analysis, Writing - original draft; Zhao Xiang Ye: Data curation; and Meng Wang: Data curation, Resources; Fengju Song: Writing - review & editing.

Disclosures of conflicts of interest: Please see ICMJE form(s) for author conflicts of interest. These have been provided as supplemental materials.

References

1. American Lung Association: Lung Cancer Trends Brief. https://www.lung.org/research/trends-in-lung-disease/lung-cancertrends-brief. Accessed March 18, 2025.
2. Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. Chin Med J (Engl) 2022;135(5):584–590.
3. Wood DE, Kazerooni EA, Baum SL, et al. Lung Cancer Screening, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2018;16(4):412–441.
4. Tanoue LT, Tanner NT, Gould MK, Silvestri GA. Lung cancer screening. Am J Respir Crit Care Med 2015;191(1):19–33.

5. Jonas DE, Reuland DS, Reddy SM, et al. Screening for Lung Cancer With Low-Dose Computed Tomography: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 2021;325(10):971–987.
6. Mulshine JL, D'Amico TA. Issues with implementing a high-quality lung cancer screening program. CA Cancer J Clin 2014;64(5):352–363.
7. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. N Engl J Med 2020;382(6):503–513.
8. Woodard GA, Jones KD, Jablons DM. Lung Cancer Staging and Prognosis. Cancer Treat Res 2016;170:47–75.
9. Lu C, Bera K, Wang X, et al. A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: a multicentre, retrospective study. Lancet Digit Health 2020;2(11):e594–e606.
10. Garon EB, Hellmann MD, Rizvi NA, et al. Five-Year Overall Survival for Patients With Advanced Non–Small-Cell Lung Cancer Treated With Pembrolizumab: Results From the Phase I KEYNOTE-001 Study. J Clin Oncol 2019;37(28):2518–2527.
11. Ettinger DS, Wood DE, Aisner DL, et al. NCCN Guidelines Insights: Non-Small Cell Lung Cancer, Version 2.2021. J Natl Compr Canc Netw 2021;19(3):254–266.
12. Zhong Y, Cai C, Chen T, et al. PET/CT based cross-modal deep learning signature to predict occult nodal metastasis in lung cancer. Nat Commun 2023;14(1):7513.
13. Huang B, Sollee J, Luo YH, et al. Prediction of lung malignancy progression and survival with machine learning based on pre-treatment FDG-PET/CT. EBioMedicine 2022;82:104127.
14. Bommasani R, Hudson DA, Adeli E, et al. On the Opportunities and Risks of Foundation Models. arXiv 2021. Preprint posted online August 16, 2021; doi:10.48550/arXiv.2108.07258.
15. Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. Nature 2023;616(7956):259–265.
16. Ma C, Tan W, He R, Yan B. Pretraining a foundation model for generalizable fluorescence microscopy-based image restoration. Nat Methods 2024;21(8):1558–1567.
17. Zhao T, Gu Y, Yang J, et al. A foundation model for joint segmentation, detection and recognition of biomedical objects across nine modalities. Nat Methods 2025;22(1):166–176.
18. Zhang H, AbdulJabbar K, Grunewald T, et al. Self-supervised deep learning for highly efficient spatial immunophenotyping. EBioMedicine 2023;95:104769.
19. Sun J, Pi P, Tang C, Wang SH, Zhang YD. TSRNet: Diagnosis of COVID-19 based on self-supervised learning and hybrid ensemble model. Comput Biol Med 2022;146:105531.
20. Bakr S, Gevaert O, Echegaray S, et al. A radiogenomic dataset of non-small cell lung cancer. Sci Data 2018;5:180202.
21. Aerts HJWL, Wee L, Rios Velazquez E, et al. Data From NSCLC-Radiomics. The Cancer Imaging Archive 2014. doi:10.7937/K9/TCIA.2015.PF0M9REI.
22. Wee L, Aerts HJL, Kalendralis P, Dekker A. Data from NSCLC-Radiomics-Interobserver1. The Cancer Imaging Archive 2019. doi:10.7937/tcia.2019.cwvlpd26.
23. Aerts HJWL, Leijenaar RTH, Parmar C, et al. Data From NSCLC-Radiomics-Genomics. The Cancer Imaging Archive 2015. doi:10.7937/K9/TCIA.2015.L4FRET6Z.
24. Kirk S, Lee Y, Kumar P, et al. The Cancer Genome Atlas Lung Squamous Cell Carcinoma Collection (TCGA-LUSC). The Cancer Imaging Archive 2016. doi:10.7937/K9/TCIA.2016.TYGKKFMQ.
25. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016; 770–778. doi:10.1109/CVPR.2016.90.
26. Huang G, Liu Z, Maaten LVD, Weinberger KQ. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017; 2261–2269. doi:10.1109/CVPR.2017.243.
27. Zhang X, Zhou X, Lin M, Sun J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018; 6848–6856. doi:10.1109/CVPR.2018.00716.
28. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229–263.
29. Lin CY, Guo SM, Lien JJJ, et al. Combined model integrating deep learning, radiomics, and clinical data to classify lung nodules at chest CT. Radiol Med 2024;129(1):56–69.
30. Huang P, Illei PB, Franklin W, et al. Lung Cancer Recurrence Risk Prediction through Integrated Deep Learning Evaluation. Cancers (Basel) 2022;14(17):4150.
31. Han Y, Ma Y, Wu Z, et al. Histologic subtype classification of non-small cell lung cancer using PET/CT images. Eur J Nucl Med Mol Imaging 2021;48(2):350–360.
32. Ding R, Yadav A, Rodriguez E, Araujo Lemos da Silva AC, Hsu W. Tailoring pretext tasks to improve self-supervised learning in histopathologic subtype classification of lung adenocarcinomas. Comput Biol Med 2023;166:107484.
33. Trabelsi M, Romdhane H, Ben Salem L, Ben-Sellem D. Advanced artificial intelligence framework for T classification of TNM lung cancer in(18)FDG-PET/CT imaging. Biomed Phys Eng Express 2024;10(6):10.1088/2057–11976/ad81ff.
34. Tayebi Arasteh S, Misera L, Kather JN, Truhn D, Nebelung S. Enhancing diagnostic deep learning via self-supervised pretraining on large-scale, unlabeled non-medical images. Eur Radiol Exp 2024;8(1):10.
35. Hsu JC, Nguyen PA, Phuc PT, et al. Development and Validation of Novel Deep-Learning Models Using Multiple Data Types for Lung Cancer Survival. Cancers (Basel) 2022;14(22):5562.
36. Salmanpour MR, Gorji A, Mousavi A, et al. Enhanced Lung Cancer Survival Prediction Using Semi-Supervised Pseudo-Labeling and Learning from Diverse PET/CT Datasets. Cancers (Basel) 2025;17(2):285.
37. Na KJ, Kim YT, Goo JM, Kim H. Clinical Utility of a CT-based AI Prognostic Model for Segmentectomy in Non-Small Cell Lung Cancer. Radiology 2024;311(1):e231793.
38. Wang Y, Ding Y, Liu X, et al. Preoperative CT-based radiomics combined with tumour spread through air spaces can accurately predict early recurrence of stage I lung adenocarcinoma: a multicentre retrospective cohort study. Cancer Imaging 2023;23(1):83.
39. Zhou Y, Chia MA, Wagner SK, et al; UK Biobank Eye & Vision Consortium. A foundation model for generalizable disease detection from retinal images. Nature 2023;622(7981):156–163.
40. Pai S, Bontempi D, Hadzic I, et al. Foundation model for cancer imaging biomarkers. Nat Mach Intell 2024;6(3):354–367.
41. Bluethgen C, Chambon P, Delbrouck JB, et al. A vision-language foundation model for the generation of realistic chest X-ray images. Nat Biomed Eng 2025;9(4):494–506.