

<https://doi.org/10.1038/s41746-025-01917-3>

Emphysema progression risk in COPD using a localized foundational model of density evolution

Check for updates

Ariel H. Curiale¹✉, Carrie Pistenmaa², Rubén San José Estépar¹, Alejandro A. Diaz², George Washko² & Raúl San José Estépar¹✉

Emphysema progression in chronic obstructive pulmonary disease (COPD) presents a notable challenge due to its significant variability among individuals and the current lack of reliable prognostic markers. Given the limited therapeutic options available for emphysema, there is a critical need for early detection and intervention strategies. Identifying individuals at risk of rapid progression is essential to effectively halt or slow the disease's advancement. This study introduces an innovative approach employing a localized foundational model of density evolution to pinpoint local radiographic embeddings indicative of emphysema progression. Central to our methodology is the Local Emphysema Progression (LEP) score, a novel metric derived from our model that aggregates localized lung tissue activations into a comprehensive, subject-level risk assessment tool. The model's performance was tested on 3728 COPDGene participants, comparing baseline to 5-year, and 1421 scans taken from the 5-year to 10-year interval period. Additionally, our findings were replicated in 1058 ECLIPSE participants. The model effectively identifies lung regions with emphysema progression, achieving an AUC of 0.88. The LEP risk score shows good correlation with the change in the percentage of low attenuation areas below -950 Hounsfield Units ($\Delta\%LAA-950$), with correlation values of 0.50 in the COPDGene cohort and 0.40 in the ECLIPSE cohort among subjects with emphysema progression ($\Delta\%LAA-950 > 0$). Furthermore, LEP risk score associates with mortality and several COPD outcomes, underscoring its potential as a valuable tool in clinical prognosis and management of emphysema progression in COPD patients.

Emphysema, a lung disease characterized by abnormal and permanent enlargement of the distal airspaces¹, exhibits varied extent and distribution within the lung, leading to heterogeneous spatial emphysema distribution^{2,3}. This destructive process of the lung tissue is thought to afflict over two million people in the US and is the pathobiological basis for chronic obstructive pulmonary disease (COPD), which is one of the leading causes of death worldwide^{4,5}. While most commonly due to chronic tobacco smoke exposure, emphysema is increasingly detected in never smokers without clear risks. Further, numerous studies including those reflecting the broader population have demonstrated that people with emphysema are at increased risk of death⁶⁻⁹. There is, therefore, great interest in developing therapies that target the underlying pathology of emphysema rather than just addressing its physiological

symptoms¹⁰. However, aside from augmentation therapy for the small subset of people with Alpha-1 antitrypsin deficiency (AATD), these efforts have had limited success. Image-based predictions can be a powerful tool to assess lobe-specific emphysema progression and provide lobe-specific treatments such as lung volume reduction.

The tool most commonly used to detect and quantify emphysema is computed tomography (CT) of the chest and extensive investigation has identified several image-based metrics of the lung tissue used in research and clinical care¹¹⁻¹³. These range from measures of the density of the parenchyma^{14,15} to more advanced machine learning algorithms that can recognize different forms of disease and its distribution¹⁶⁻²². Clinical studies of emphysema are, however, challenging to conduct because the evolution of the disease is slow, and the regional heterogeneity of this process limits

¹Applied Chest Imaging Laboratory, Department of Radiology and Medicine, Brigham and Women's Hospital, Harvard Medical School, 399 Revolution Drive, Somerville, 02145 MA, USA. ²Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.

✉ e-mail: acuriale@bwh.harvard.edu; rsanjose@bwh.harvard.edu

our ability to identify where in the lung this pathologic remodeling is most active^{23,24}. Moreover, while we generally believe that people with emphysema will develop more emphysema, a therapy to heal destroyed lung tissue is likely more challenging to create than one that prevents its onset. The ability to predict where emphysema will develop in the lung may fill a critical gap that is currently preventing therapeutic development and optimal clinical management of patients at high risk of developing emphysema.

We hypothesized that a CT-based deep learning model could predict where emphysema would develop after learning the features that characterize its radiologic progression. To test this hypothesis, we leveraged longitudinal data from over 5000 people enrolled in one of two observational studies to identify lung-specific “embeddings” that described local emphysema progression (LEP), associated with the local onset of emphysema over a five-year period. A detailed description of the local emphysema progression can be found in the Methods and the Supplementary Fig. 5. These “embeddings” are mathematical representations of patterns embedded in the image, which are not evident on subjective review but are able to identify tissue that is most likely to become emphysematous. Once the embeddings of disease progression were identified, we then sought to test their ability to predict airspace dilation in a new set of longitudinal images that were not used for training.

Our deep learning approach leveraged a local foundational model (FM) of density evolution with an attention and scale mechanism where the model focuses on important details mimicking a cognitive attention mechanism^{25,26}. The training involved: (1) image co-registration²⁷, local densitometric assessment of the parenchyma (percentage of low attenuation areas below -950 Hounsfield Units, %LAA-950), and z-score image normalization. This was followed by (2) a conditional training strategy that balanced encoding emphysema progression information into the embeddings of the local FM, proper local reconstruction of the co-registered follow-up CT scan, and emphysema progression based on the change of %LAA-950 (Fig. 1). Once trained, the algorithm was applied to de-novo baseline CT scans using the following stepwise approach: (1) z-score image normalization; (2) inference of the local emphysema progression; (3) aggregation of the those spatial likelihoods

of emphysema progression across the entire lung to create the Local Emphysema Progression Activity Map (Fig. 1), and finally (4) generation of a local emphysema progression score by aggregating local activations. Details are provided in the Methods section. We predicted the amount of emphysema progression each person would experience as well as estimated their clinical impairment due to that progression. We then compared those predicted metrics of progression to the same outcomes ascertained from serial observation.

Results

Study participants

A total of 10,198 smokers with and without COPD were enrolled in the COPDGene²⁷ study at phase 1 baseline (P1), and over 6000 participants came back for a second visit (P2) 5 years later. Our training and testing data consisted of 4712 COPDGene²⁷ participants with complete data at baseline and the 5-year follow-up. This cohort included 1882 participants with Global Initiative for Obstructive Lung Disease (GOLD) COPD grades I-IV expiratory airflow obstruction²⁸ and 2805 smokers with normal lung function and preserved ratio impaired spirometry (PRISm). For this study, we utilized the primary COPDGene spirometry data, which employed race-adjusted equations based on post-bronchodilator spirometry values, using NHANES race-specific equations and the 2023 GOLD criteria²⁸. The local FM and LEP models were developed using 6.888 million lung patches of 32 by 32 pixels (0.64 mm²) from 984 randomly selected COPDGene participants at P1P2 with complete data (only 7000 patches by participant). Subjects were selected to keep the same GOLD severity distribution observed in our entire cohort (see Table 1). The LEP model’s performance was locally tested on 981.5 million patches from 795 COPDGene participants randomly selected from the test data at P1P2 (Supplementary Fig. 1—Local Reconstruction). In addition, LEP reproducibility was tested on a subset of COPDGene participants at phase 2 with a high- and low-dose protocol ($n = 1421$ and $n = 41$, Table 1) and complete data at 10-year later (P3), and 1058 ECLIPSE²⁹ participants with complete data at baseline and 3-year follow-up. Details of these populations can be found in Table 1 and in the Supplementary Fig. 1.

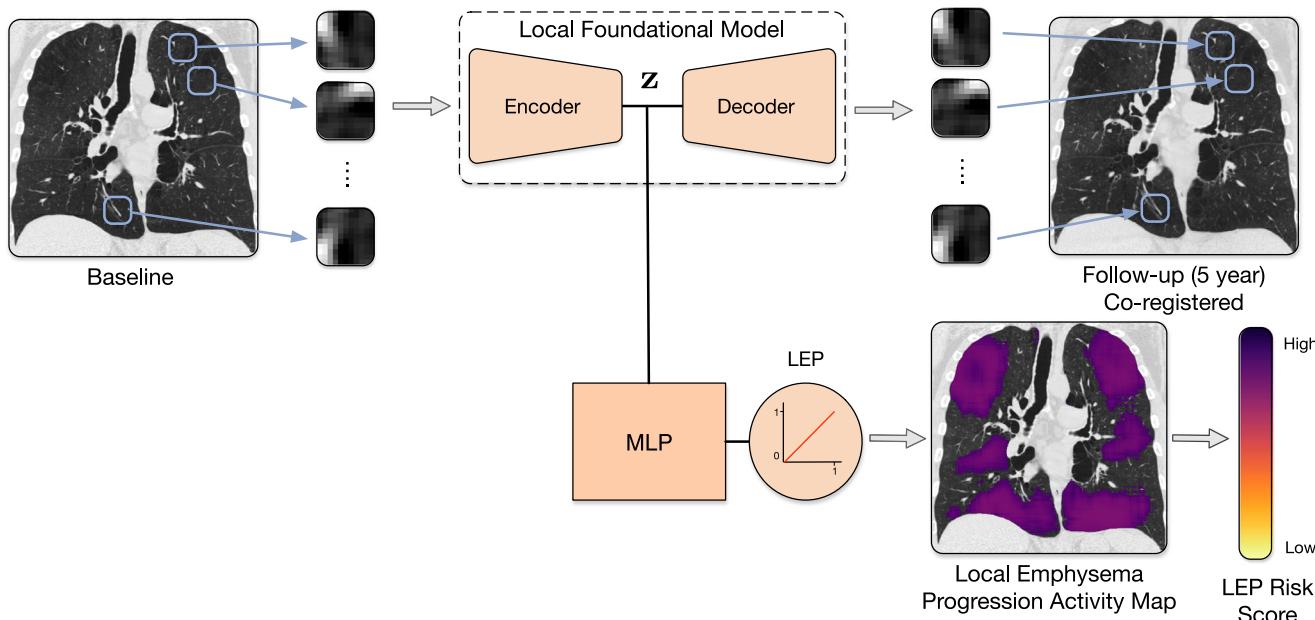


Fig. 1 | Schematic of the deep learning workflow for prognostication of emphysema progression. This workflow diagram illustrates the process of local emphysema progression prediction using a local foundational model of density evolution. Starting with high-resolution CT scans, regions of interest (ROIs) are identified and fed into the FM. The encoder part of the autoencoder compresses the ROI into a lower-dimensional representation z , which is then reconstructed back to

the original data space by the decoder. The feature vector z is subsequently processed through a multilayer perceptron (MLP) to estimate the progression risk score. The risk score is depicted as the local emphysema progression. The final output includes the localized visualization of emphysema progression on the CT scan, highlighted for clinical assessment.

Table 1 | Clinical characteristics of the populations used for training and testing

Baseline Characteristics	COPDGene						ECLIPSE	
	Phase 1–2				Phase 2–3			
	All	Training	LEP Score Test	Local Model Test	HD Test	LD Test		
<i>N</i>	4712	984	3728	795	1421	41	1058	
Race	Non-Hispanic White	3341 (70.9%)	727 (73.9%)	2614 (70.1%)	574 (72.2%)	1013 (71.3%)	23 (56.1%)	
	Non-Hispanic Black	1371 (29.1%)	257 (26.1%)	1114 (29.9%)	221 (27.8%)	408 (28.7%)	28 (43.9%)	
Females	2315 (49.1%)	476 (48.4%)	1839 (49.3%)	381 (47.9%)	722 (50.8%)	24 (58.5%)	416 (39.4%)	
Age [y]	59.8 (8.7)	60.0 (8.5)	59.7 (8.7)	59.8 (8.3)	59.2 (8.3)	55.1 (7.6)	61.5 (7.9)	
BMI	29.2 (6.1)	28.9 (5.8)	29.1 (6.2)	29.1 (6)	28.9 (6)	29.1 (6.5)	26.5 (5.2)	
Height [cm]	170.1 (9.6)	170 (9.6)	170.1 (9.5)	170 (9.5)	169.9 (9.4)	169.1 (8.9)	169.7 (9.0)	
Pack-Years	42.4 (23.5)	43.2 (23.9)	42.2 (23.4)	43.0 (24.0)	42.1 (22.8)	40.1 (29.9)	44.8 (28.4)	
FEV1 [liter]	2.4 (0.8)	2.3 (0.8)	2.4 (0.8)	2.4 (0.8)	2.4 (0.8)	2.6 (0.9)	1.71 (0.94)	
FEV1/FVC	0.7 (0.1)	0.7 (0.1)	0.7 (0.1)	0.7 (0.1)	0.7 (0.1)	0.7 (0.1)	0.51 (0.18)	
No COPD	2805 (60%)	554 (57%)	2251 (60%)	489 (62%)	860 (61%)	29 (71%)	218 (20.6%)	
GOLD status	1	427 (9%)	87 (9%)	340 (9%)	65 (8%)	162 (11%)	5 (12%)	
	2	907 (19%)	199 (20%)	708 (19%)	152 (19%)	256 (18%)	5 (12%)	
	3	446 (10%)	124 (12%)	322 (9%)	78 (10%)	111 (8%)	2 (5%)	
	4	102 (2%)	20 (2%)	82 (2%)	11 (1%)	19 (1%)	0 (0%)	
%LAA-950	5.2 (7.9)	5.5 (8.1)	5.1 (7.9)	5.0 (7.3)	5.2 (7.7)	3.6 (6.2)	10.7 (12.2)	
Smoking Status	CS	2439 (51.8%)	507 (51.5%)	1932 (51.8%)	419 (52.7%)	781 (55%)	18 (43.9%)	
	FS	2273 (48.2%)	477 (48.5%)	1796 (48.2%)	376 (47.3%)	640 (45%)	23 (56.1%)	

Continuous variables are presented as mean (SD) and categorical variables as *N* (%).

HD High Dose, LD Low Dose, CS Current Smoker, FS Former Smoker.

Local prognostic model performance and reproducibility

We first studied the mean absolute error (std) of the LEP model predictions on 981.5 million patches from 795 randomly selected COPDGene participants from phase 1 to phase 2 (P1P2) not included in the training dataset (around 1.23 million patches by participant). In general, the LEP prediction shows a small MAE = 4.4% (7.7%). This error increased along with the COPD status with values of 1.6% (4.7), 2.9% (6.5%), 6.1% (8.8%), 5.8% (8.3%), 9.1% (8.8%) and 10.7% (8.8%) for PRISm, controls, GOLD 1 to 4, respectively.

Then, we explored the performance of the LEP model to determine whether the model could effectively discriminate between regions of lung tissue stability and those showing emphysema progression (regions with > 5% LEP voxels, i.e. voxels than transition from tissue to air). Analyzing the patches described before (981.5 million patches from 795 COPDGene participants P1P2), our model demonstrated a robust Area Under the Curve (AUC) of 0.88. At an operating point that maximizes the Youden's index on the training dataset (probability th = 0.158), the model's sensitivity (Sens.), positive predictive value (PPV), accuracy (ACC), F1-score, and Cohen's Kappa coefficient (K) were Sens. = 0.55, PPV = 0.60, ACC = 0.85, F1 = 0.58, and K = 0.49, respectively.

When the model was applied to COPDGene participants from phase 2 to phase 3 (P2P3) under a high-dose protocol (*n* = 2464 million patches, 1.3 million patches by participant), the model's performance metrics remained consistent (AUC = 0.88, Sens.=0.64, PPV = 0.53, ACC = 0.85, F1 = 0.59, and K = 0.50). However, in the subset evaluated with a low-dose protocol at P2P3, a slight reduction in performance was observed (AUC = 0.84, Sens. = 0.61, PPV = 0.45, ACC = 0.83, F1 = 0.52, and K = 0.42) with decrements of 4.5%, 4.7%, 15%, 2.3%, 11.8% and 16% in AUC, sensitivity, PPV, ACC, F1-score, and K, respectively.

Additionally, we observed that the performance of our model was generally consistent across different baseline disease severities, except in the case of severe emphysema. We categorized the subjects into four groups based on baseline emphysema severity (no emphysema 0-5% (*n* = 563),

mild 5-10% (*n* = 109), moderate 10-20% (*n* = 70), and severe ≥ 20% (*n* = 53)) and generated Receiver Operating Characteristic (ROC) curves for each group by aggregating the per-subject ROC curves within the same emphysema severity group measured at baseline. These curves demonstrated an AUC ≥ 0.78 for the milder groups (no emphysema, mild, and moderate) in high-dose CT scans with a slight drop in performance for subjects with advanced emphysema at baseline (Fig. 2a, b), and an AUC of ≥ 0.73 for low dose scans (Fig. 2c).

To enhance the interpretability of our model, we conducted a visual evaluation of the local FM predictions, which reconstructed the 5-year CT patches from the baseline scans. The results for various patches, representing subjects with differing disease burdens and progression rates, are presented in Supplementary Fig. 2.

Local emphysema progression activity maps

We computed subject-level maps of emphysema progression activity by sequentially running the LEP model across the lung parenchyma and aggregating the probability outputs of local progression. Figure 3a shows the LEP activity maps for four subjects with different degrees of severity. Our model is able to detect incipient regions of the parenchyma most likely to progress to emphysema in a subject without emphysema at baseline and without change in their clinical staging by GOLD guidelines (GOLD stage 2) during the 5-year follow-up (Fig. 3 (a-No Emphysema)). Figure 3 (a-Mild) shows a strong spatial correlation between the model's LEP prediction and the measurement at 5-year, where the model effectively identifies specific regions in the lung where the progression activity is concentrated. Additionally, Fig. 3a shows that the model is not merely identifying local regions with emphysema at baseline (Mild and Severe). A visual inspection of several patches confirms the same local behavior (see Supplementary Note and Supplementary Fig. 2). In cases with more advanced disease stages (moderate to severe emphysema and GOLD stages 2 and 3), the LEP activity maps were also able to capture the heterogeneous pattern of emphysema progression (Fig. 3 (a -Moderate and Severe)). This visual demonstration

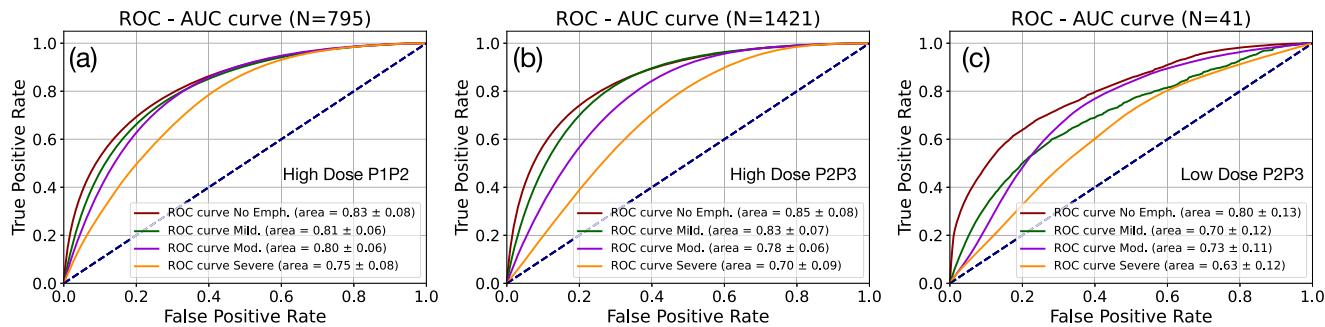


Fig. 2 | Performance of the LEP Model at the local patch level. The LEP model was evaluated on test subjects from the COPDGene cohort with high-dose CT scans from phase 1 to phase 2 (P1P2) and phase 2 to phase 3 (P2P3). We also evaluated the model robustness in 41 participants with low-dose CT scan from Phase 2 to Phase 3.

ROC curves for the detection of local emphysema progression. Local emphysema progression was defined as more than 5% of the pixels within a local patch transitioning from tissue (CT density > -950 HU) to air (CT density ≤ -950 HU). LEP Local Emphysema progression, E Emphysema measured at baseline.

underscores the model's precision in predicting emphysema progression, highlighting its potential clinical utility.

Association of LEP risk score and emphysema progression

The LEP activity maps serve as the foundation for creating a subject-level emphysema progression risk metric known as the LEP score. The score is calculated as the mean binary local activation at the threshold that maximizes Youden's index (probability $th = 0.158$) across the entire lung parenchyma. A comprehensive explanation of the LEP risk score methodology is detailed in the Methods section.

We examined the relationship of the LEP risk score and emphysema progression, quantified as the change of %LAA-950 ($\Delta\%LAA-950$), in 3728 COPDGene participants at P1P2, and in 1058 participants of our replication cohort, ECLIPSE. A positive change of the %LAA-950 (indicative of progression) signals an increase of the emphysema by P2, whereas a negative change corresponds to an increase in voxels above the -950 HU threshold at P2. While emphysema is understood as an irreversible process, lung density can increase due to technical and physiological factors like scanner parameters, inflation levels, smoking resumption and underlying inflammatory processes during disease activity that could potentially lead to emphysema. On the other hand, higher LEP scores indicate a greater risk of emphysema progression.

The correlation between the LEP score and $\Delta\%LAA-950$ among participants with emphysema progression ($\Delta\%LAA-950 > 0$) was 0.59, and -0.57 among those with an increase in lung density (i.e., $\Delta\%LAA-950 \leq 0$) (see Supplementary Fig. 3). In the ECLIPSE replication cohort, the corresponding correlations were 0.40 and -0.44 , respectively, further supporting our findings in COPDGene. This pattern suggests that the LEP score captures not only emphysema progression but also potential inflammatory or reparative processes that may precede or counteract tissue destruction. The variability in model correlation observed between the COPDGene and ECLIPSE cohorts (-27.5%) is primarily attributed to differences in CT acquisition protocols—including noise characteristics—and scanner manufacturers. To further investigate this, we assessed the model's performance across different scanner brands within the COPDGene cohort. The correlation between the LEP score and $\Delta\%LAA-950$ varied by scanner type, with values of -8.6% for Siemens, 10.4% for GE, and 14.3% for Philips scanners, relative to the correlation computed across all COPDGene participants (see Supplementary Table 2). This comparison provides useful insights into the potential influence of scanner-related variability on model performance; however, because the same subjects were not scanned across different scanner manufacturers, definitive conclusions about the effect of scanner type on variability cannot be drawn.

We then stratified the correlation between the LEP score and $\Delta\%LAA-950$ by COPD severity (GOLD 0–IV) among COPDGene participants. For those with emphysema progression ($\Delta\%LAA-950 > 0$), the correlations were 0.33 for GOLD 0, 0.33 for GOLD I, 0.49 for GOLD II, 0.34 for GOLD

III, and 0.27 for GOLD IV. Among participants with stable or increasing lung density ($\Delta\%LAA-950 \leq 0$), the corresponding correlations were -0.65 , -0.45 , -0.54 , -0.47 , and -0.09 for GOLD stages 0 through IV, respectively.

Additionally, when stratifying participants based on $\Delta\%LAA-950$ quartiles in the COPDGene dataset, the correlation values were -0.48 for $\Delta\%LAA-950 < -0.87\%$ (25th percentile), -0.26 for values between -0.87% and 0.00% (25th to 50th percentile), 0.26 for values between 0.00% and 1.07% , and 0.47 for values $\geq 1.07\%$ (75th percentile). In the ECLIPSE replication cohort, the corresponding correlations were -0.48 , -0.26 , 0.25 , and 0.47 , respectively.

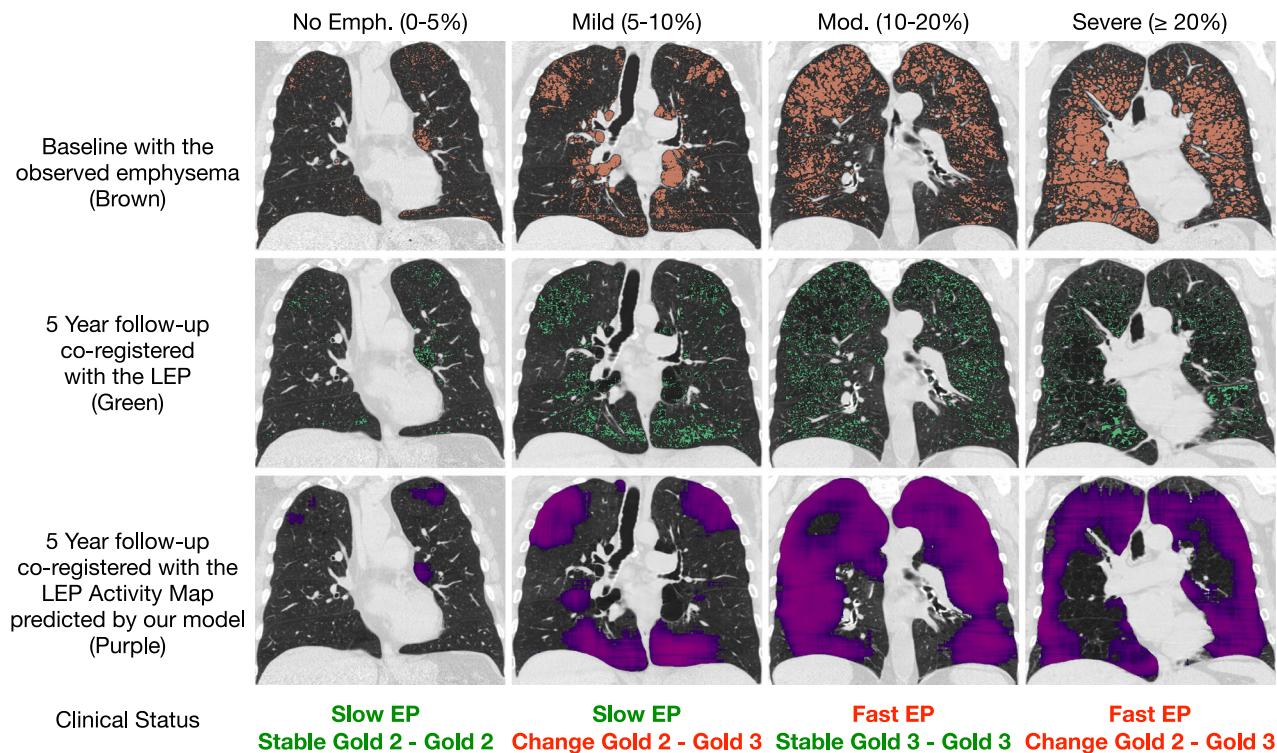
Association of local emphysema progression risk score with prospective outcomes

In the COPDGene LEP Score test group (P1P2, $n = 3728$), we examined the association of LEP risk scores and $\Delta\%LAA-950$ with various COPD outcomes by fitting a multivariate linear model adjusted by the following covariate variables measured at baseline: age, gender, race, emphysema, BMI, TLC, smoking status, and packs per year (see Data and statistical analyses section on Online Methods). Emphysema progression was classified as “fast” or “slow” based on %LAA-950 changes above 0.66%. Emphysema progression risk was categorized as “high-risk” or “low-risk” using an LEP score threshold above 0.15. We established these thresholds at the 70th percentile of the participants in the “LEP score” test.

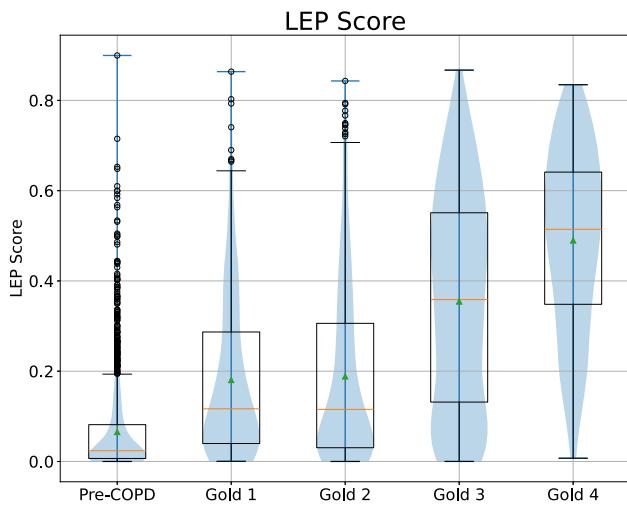
LEP scores were significantly correlated with COPD severity, increasing from 0.065 in smoking controls to 0.49 in GOLD 4. As expected, LEP risk scores also rose with emphysema progression from 0.099 (slow) to 0.236 (fast) (see Fig. 3b, c). The mean (std) LEP values for slow and fast progressors according to $\Delta\%LAA-950$ were 0.06 (0.10) and 0.08 (0.11) for smoking controls, 0.16 (0.17) and 0.21 (0.19) for GOLD 1, 0.14 (0.17) and 0.24 (0.21) for GOLD 2, 0.30 (0.23) and 0.40 (0.23) for GOLD 3, and 0.45 (0.20) and 0.50 (0.20) for GOLD 4.

Table 2 summarizes the associations between the LEP risk score and $\Delta\%LAA-950$ across various COPD outcomes. Statistically significant differences were observed in the 5-year FEV1/FVC ratio and DLCO between low and high-risk groups for the LEP score and slow vs fast progressors for $\Delta\%LAA-950$ ($p < 0.001$). Patients classified as high risk based on their LEP score showed lower FEV1/FVC ratios and DLCO values. Additionally, the high-risk group, as per LEP scores, exhibited a greater decline in FEV₁ (approximately 12 ml/y; -46.51 vs. -34.80 , $p < 0.001$) and a larger decrease in exercise capacity (about 40 feet; -168 vs. -127.4 , $p < 0.001$) than low-risk subjects, indicating a worsening of their disease condition. In contrast, $\Delta\%LAA-950$ showed a less pronounced impact on those outcomes between slow and fast progressors (approximately 0.31 ml/y in the difference of FEV₁ change, $p = 0.03$; and about 26 feet in the exercise capacity, $p = 0.83$). Additionally, Table 3 shows the outcome distribution of four risk groups

(a)



(b)



(c)

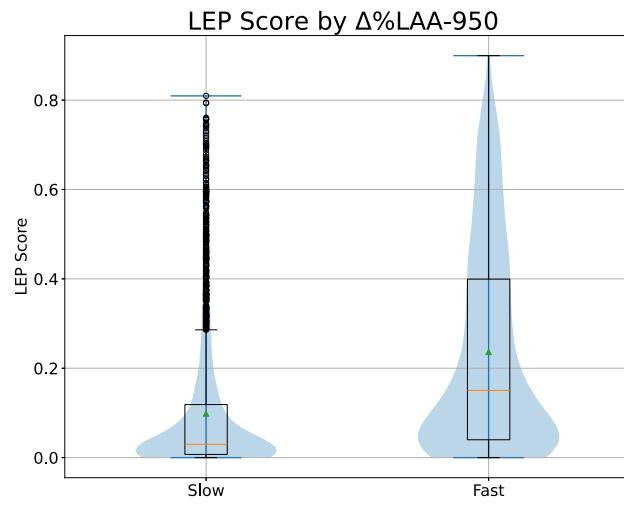


Fig. 3 | Illustration of the proposed LEP Activity Maps and LEP Risk Score distributions. **a** LEP Activity Maps for four subjects with different degrees of severity. The model's prediction (i.e., LEP activity map) is plotted over the co-registered CT scan in purple (bottom row), while the observed emphysema at baseline is depicted in brown in the first row. The observed emphysema progression

based on the change in CT density of %LAA-950 between baseline and follow-up is depicted in green in the second row. LEP: Local Emphysema Progression; EP: Emphysema Progression. **b** Distribution of the LEP risk score on subjects at P1P2 stratify by Pre-COPD (excluding PRISM) and GOLD stages; and **c** by Low/High % LAA-950.

defined according to the quartiles Q1 = 0%–25%, Q2 = 25%–50%, Q3 = 50%–75%, and Q4 = 75%–100%.

Survival analysis

Lastly, we evaluated the prognostic utility of the LEP risk score by conducting a survival analysis. Kaplan-Meier estimator was utilized to delineate survival trajectories, and the Cox proportional hazard model assessed the

significance of the risk stratification. Adjustments in the Cox model accounted for an array of covariates measured at baseline including age, gender, race, emphysema, body mass index (BMI), total lung capacity (TLC), smoking status, and pack-years of smoking. A second survival model included those covariates and FEV1.

As depicted in Fig. 4a, our analysis within the COPDGene test cohort (P1P2, $n = 3728$) revealed distinct survival trajectories that paralleled the

Table 2 | Association between LEP Score, $\Delta\%$ LAA-950, and various COPD outcomes over a 5-year period

Outcomes (n = 3728)	LEP Score				$\Delta\%$ LAA-950			
	Low Risk mean (std)	High Risk mean (std)	Diff.	p value	Slow Prog. mean (std)	Fast Prog. mean (std)	Diff.	p value
FEV1/FVC @ P2	0.74 (0.04)	0.55 (0.10)	-0.06	<0.001	0.72 (0.08)	0.59 (0.12)	-0.07	<0.001
Change of FEV1 [ml/yr]	-34.80 (8.52)	-46.51 (9.30)	-8.15	<0.001	-38.24 (9.85)	-38.55 (10.5)	3.74	0.03
Dlco @ P2	19.33 (4.12)	15.28 (5.98)	-2.07	<0.001	18.93 (4.60)	16.13 (5.74)	-1.62	<0.001
Change dist. walked [feet]	-127.4 (48.3)	-168.0 (65.7)	31.32	0.06	-131.6 (52.6)	-158.3 (59.6)	-4.46	0.83
Change MMRC	0.01 (0.11)	0.19 (0.12)	0.04	0.49	0.01 (0.12)	0.17 (0.14)	0.10	0.04

This table summarizes the associations and contrasts between low and high-risk groups based on the LEP score, as well as between slow and fast emphysema progressors as defined by changes in measured %LAA-950, a quantitative metric for emphysema progression. The association was measured by fitting a linear model adjusted by the following covariate variables measured at baseline: age, gender, race, %LAA950, BMI, TLC, smoking status, and packs per year. Key COPD outcomes examined include the FEV1/FVC ratio, annual change in FEV1, DLCO measurements, changes in exercise capacity (distance walked), and modifications in the MMRC score. Adjusted covariance means, and the ordinary least square difference (Diff.) are shown for each group. Statistically significant results (p-value ≤ 0.05) are in bold.

Table 3 | Association between LEP Score, $\Delta\%$ LAA-950, and various COPD Outcomes over a 5-year Period according to the risk groups defined by the quartiles Q1 = 0%–25%, Q2 = 25%–50%, Q3 = 50%–75%, and Q4 = 75%–100%

Outcomes (n = 3728)	LEP Score, mean (std)					$\Delta\%$ LAA-950, mean (std)				
	Q1	Q2	Q3	Q4	p-val	Q1	Q2	Q3	Q4	p-val
FEV1/FVC @ P2	0.8 (0.1)	0.7 (0.1)	0.7 (0.1)	0.5 (0.2)	<0.001	0.7 (0.1)	0.7 (0.1)	0.7 (0.1)	0.6 (0.2)	<0.001
Change of FEV1 [ml/yr]	-30.3 (46.0)	-35.8 (48.6)	-40.9 (47.7)	-46.2 (48.7)	<0.001	-41.5 (48.0)	-39.2 (49.1)	-33.5 (47.1)	-39.0 (48.0)	0.004
Dlco @ P2	18.8 (5.2)	19.7 (5.6)	19.7 (6.2)	15.0 (7.0)	<0.001	19.1 (6.3)	19.1 (5.7)	19.1 (6.0)	15.9 (6.8)	<0.001
Change dist. walked [feet]	-102.7 (343.4)	-139.5 (360.2)	-141.2 (356.3)	-173.2 (370.7)	0.001	-138.9 (347.2)	-118.8 (360.6)	-139.1 (335.9)	-160.1 (387.8)	0.111
Change MMRC	-0.1 (1.4)	0.0 (1.2)	0.1 (1.1)	0.2 (1.2)	<0.001	0.0 (1.0)	-0.0 (1.3)	0.0 (1.2)	0.2 (1.3)	<0.001

This table summarizes the associations and contrasts between low and high-risk groups based on the LEP score, as well as between slow and fast emphysema progressors as defined by changes in measured %LAA-950.

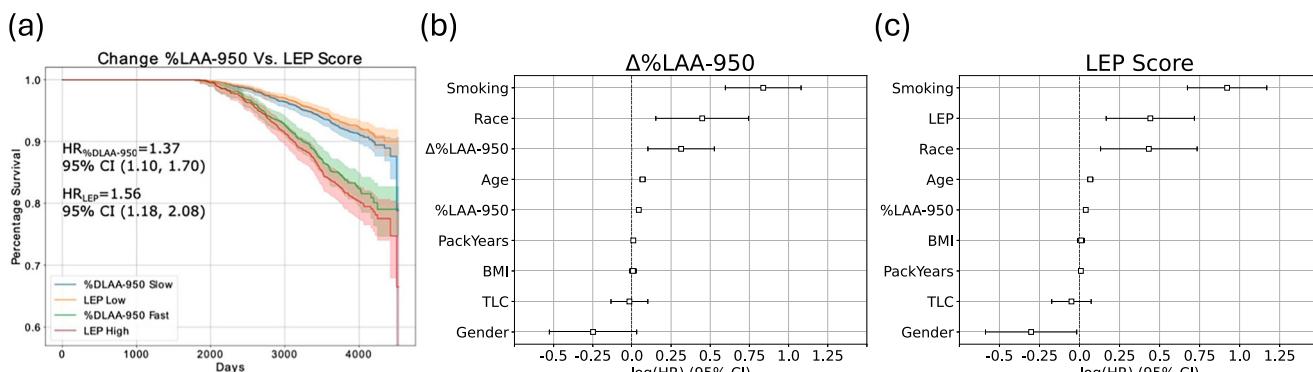


Fig. 4 | Kaplan–Meier and Cox proportional-hazards analysis for high/low LEP risk and fast/slow emphysema progression. a Kaplan–Meier survival curves demonstrating the differential impact of $\Delta\%$ LAA-950 and LEP score categories on the survival of patients in the COPDGene study (n = 3728). Patients were stratified into two groups based on $\Delta\%$ LAA-950 (slow and fast emphysema progression) and into two groups based on LEP scores (low and high risk). Forest plot of the Cox

model for $\Delta\%$ LAA-950 (b) and LEP score (c) adjusted for baseline covariates, showing the log(HR) for mortality risk associated with fast progressors and high-risk after adjustment for clinical and demographic factors. Both models underscore the enhanced risk of mortality associated with higher LEP scores and faster emphysema progression, independent of other risk factors.

classification of patients with fast/slow $\Delta\%$ LAA-950 progression and high/low LEP risk. Notably, individuals presenting with high LEP scores or those exhibiting a rapid increase in $\Delta\%$ LAA-950 were associated with an elevated risk of mortality when compared to their respective reference cohorts. Hazard ratios were notably higher in these groups, with a hazard ratio (HR) of 1.56 for high LEP scores and 1.37 for fast $\Delta\%$ LAA-950 progression, underscoring the predictive relevance of these biomarkers in survival outcomes (refer to Fig. 4b and c). While smoking status exhibits a consistent

influence across both models, the LEP score emerges as a particularly significant determinant in the context of emphysema progression and its impact on mortality. Additionally, when the Cox model is adjusted for spirometric measures, such as FEV1, the significance of LEP and $\Delta\%$ LAA-950 is lost (LEP score p value = 0.06 and $\Delta\%$ LAA-950 p value = 0.22) because the impact of spirometric measures on mortality is greater than that of emphysema—particularly due to PRISm participants. However, in subjects with high emphysema percentage, for example, $\%LAA-950 > 10\%$, the

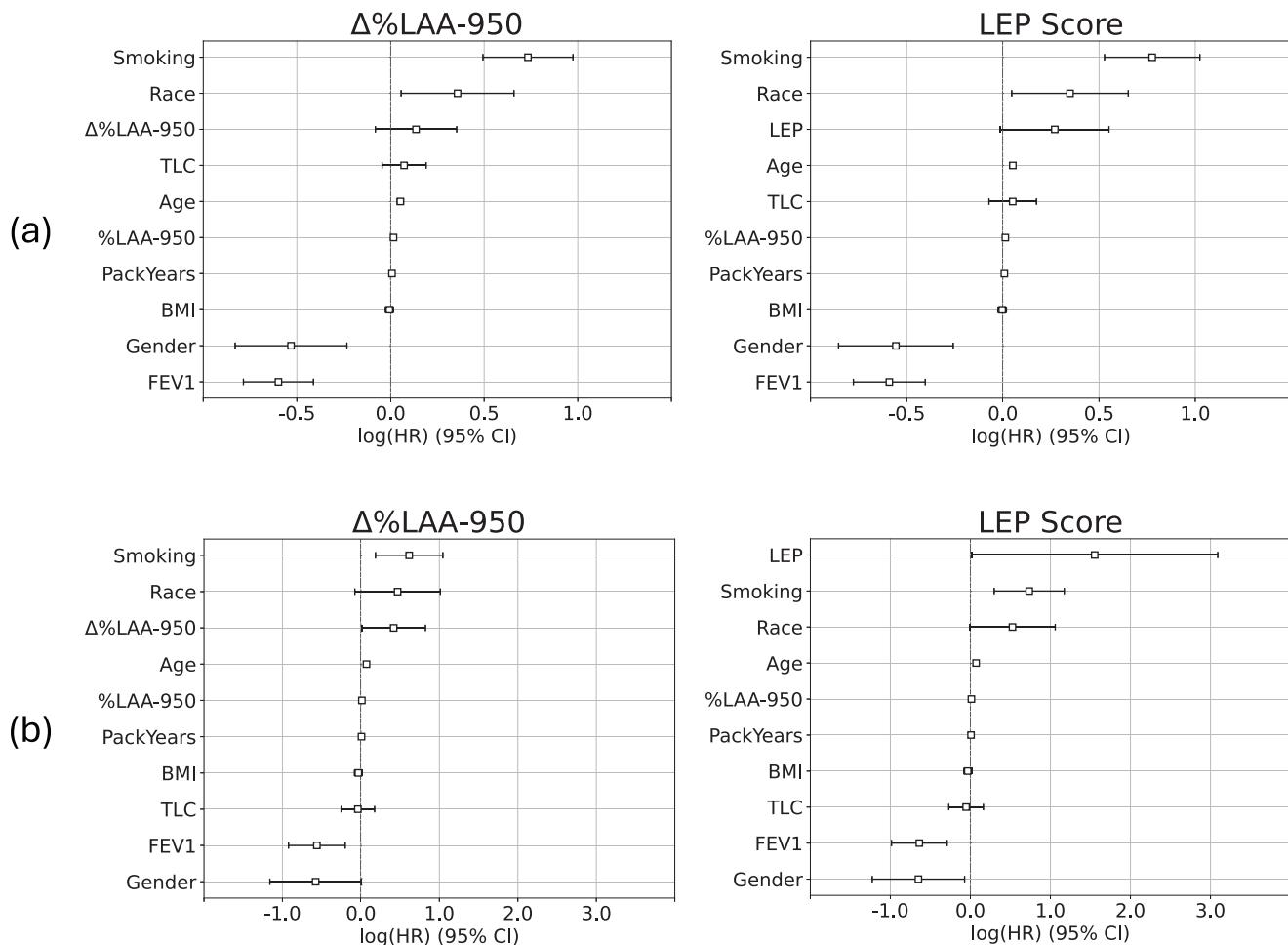


Fig. 5 | Cox analysis of LEP risk and emphysema progression adjusted for covariates including FEV₁. Cox Proportional-Hazards analysis for high/low LEP risk and fast/slow emphysema progression in (a) the COPDGene study ($n = 3728$) and (b) on COPDGene participants with emphysema at baseline higher than 10%

($n = 576$). The Cox model is adjusted for baseline covariates including FEV₁, showing the log(HR) for mortality risk associated with fast progressors and high-risk after adjustment for clinical and demographic factors.

impact of emphysema on mortality increases. This is illustrated in Fig. 5b where the LEP score emerges as a particularly significant determinant of mortality impact when stratified by emphysema severity. Additionally, Supplementary Fig. 4 presents a survival analysis restricted to the pre-COPD group (excluding PRISM) and GOLD 1–2 stages, in which the LEP score remains significantly associated with mortality even after adjusting the cox model for FEV₁. This may appear to contrast with earlier analyses suggesting that the prognostic performance of our model declines as emphysema severity increases. Nevertheless, this can be explained by considering both the structural limitations in advanced disease and the shifting contributions of different risk factors to mortality. As baseline emphysema severity increases, the amount of preserved lung tissue available for prediction decreases. Severely emphysematous regions are less likely to undergo further progression, meaning they contain less information about future structural decline. As demonstrated in the local prediction analyses, this reduction in viable tissue reduces the model's precision for predicting local emphysema progression, with the mean absolute error (MAE) increasing from 1.6% in Prism participants (mean emphysema = 1.51%) to 10.7% in GOLD 4 participants (mean emphysema = 23%) (section “Local Prognostic Model Performance and Reproducibility”). Despite this decline in predictive precision, the LEP score remains associated with mortality in individuals with greater emphysema burden, because the relative contribution of emphysema to mortality increases as its extent grows. While spirometric measures (such as FEV₁) generally have a stronger influence on mortality than emphysema alone, in subjects with more extensive emphysema—such

as those with %LAA-950 greater than 10% (mean emphysema values for GOLD 3 and GOLD 4 in COPDGene are 13% and 23%)—emphysema becomes a more prominent driver of mortality risk. In this setting, the LEP score continues to reflect clinically meaningful risk, even if its ability to predict future tissue destruction is reduced. To further support this point, Fig. 6 presents the association between the LEP score and mortality specifically in GOLD 4 participants. To ensure sufficient statistical power, we defined the at-risk population as those with LEP scores above the mean, identifying 42 high-risk individuals out of 84 GOLD 4 participants. Using a higher percentile threshold (e.g., the 70th percentile) would have reduced the number of events and limited the power of the Cox proportional hazards model. As expected, the results show that the LEP score retains its prognostic value, while FEV₁ loses its statistical significance—likely due to reduced power—despite maintaining a high hazard ratio. These findings affirm the prognostic significance of the LEP score in identifying patients at a higher risk of adverse outcomes, thus offering a potent tool for personalized patient management and intervention strategies.

Discussion

In this study, we introduce a novel Localized Emphysema Progression (LEP) model, utilizing a temporal autoencoder with an attention mechanism to define and evaluate image embeddings, or “deep features”, as potential tools for use in more advanced prognostic models to predict COPD outcomes, such as changes in %LAA-950. Our model uses baseline CT images to predict future progression, and it highlights the potential of utilizing baseline

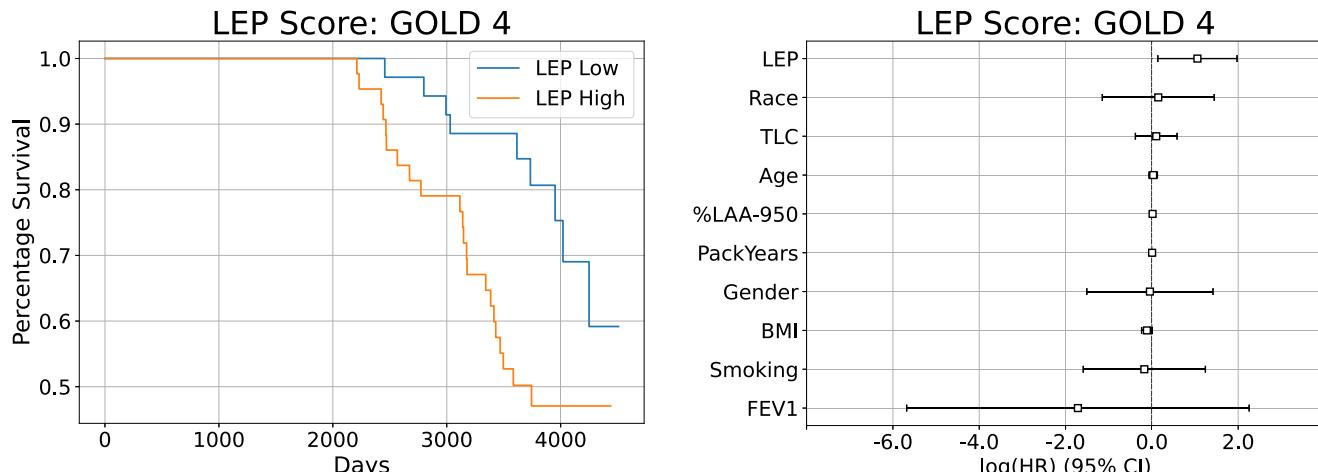


Fig. 6 | Kaplan–Meier and Cox Proportional-Hazards analysis for high/low LEP risk in COPDGene participants with GOLD stage 4 ($n = 84$).

imaging data to anticipate disease trajectory. The LEP score may prove to be an invaluable tool for monitoring disease progression and informing treatment strategies for patients with COPD. The study's findings also open avenues for future research. For example, this approach can be integrated with follow-up exams in future models to enhance monitoring capabilities and treatment planning, representing a significant step towards more sophisticated and practical prognostic tools in clinical settings.

Our intent is not to provide the most finely tuned prognostic model or an activity map of lung destruction. The necessity of our model being a true auto-encoder is crucial; it must have a bottleneck to ensure that the encoded information of the evolution of emphysema progression (EP) is entirely encapsulated in the latent space (image embeddings) rather than being distributed across multiple layers. The U-Net architecture does not adhere to this requirement, as information flows through the skip connections at various levels, which could dilute the purpose of encoding progression into a compact latent representation.

The model's ability to accurately predict 5-year local emphysema progression from baseline CT scans underscores its potential in early identification of patients at elevated risk. Its capacity to localize the evolution of emphysema and detect subjects on a rapid progression trajectory is a distinctive feature, offering invaluable insight not easily attained by other methodologies³⁰. However, to the best of our knowledge, there are currently no existing models specifically designed to predict regional or lobar emphysema progression without follow-up imaging, and therefore no established baseline for direct comparison. We acknowledge this as a limitation. Nonetheless, the proposed LEP model may serve as an initial reference framework for future research aiming to develop and validate standardized benchmarks. Establishing such benchmarks will be essential for enabling comprehensive evaluation and comparison of emerging methodologies in this field.

We propose an emphysema risk score—referred to as the LEP score—by aggregating localized progression data generated by our model to provide a subject-level risk score. A key finding of our study is that high LEP scores are associated with both increases and decreases in %LAA-950 over five years, reflecting the underlying heterogeneity of COPD phenotypes. Specifically, a high LEP score accompanied by an increase in %LAA-950 is indicative of true emphysema progression, likely driven by localized parenchymal destruction and consistent with biological mechanisms such as protease-antiprotease imbalance, oxidative stress, and chronic inflammation. In contrast, a high LEP score accompanied by a decrease in %LAA-950 may reflect airway-predominant phenotypes or PRISM, non-parenchymal remodeling, underlying inflammation, or technical factors such as variability in lung inflation during image acquisition. This seemingly counter-intuitive finding is consistent with emerging theories that emphasize disease activity and unregulated inflammatory responses as central to emphysema

progression^{24,31}. In such cases, increased inflammation may transiently elevate lung density, thereby reducing %LAA-950 and potentially masking true structural decline on longitudinal imaging^{32,33}. Notably, the LEP score appears capable of identifying these individuals despite misleading imaging trends. Integrating additional clinical and imaging biomarkers—such as gas trapping, airway wall thickness, and spirometry—may further improve the interpretability of LEP-based predictions and enhance our ability to distinguish between emphysema- and airway-predominant phenotypes.

Our findings also suggest that in high-risk individuals, a lung function decline which is more pronounced in high-risk groups identified by the LEP score compared to slow and fast progressors identified by $\Delta\%$ LAA-950 (-46.1 vs. -38.2 and -38.5). While some changes may not indicate clinical improvement per se, they can reflect an underlying trend toward worsening condition—such as dyspnea—particularly when evaluated across groups stratified by risk. For example, in Table 2, the mMRC change from Q1 to Q4 of the LEP score (from -0.1 to 0.2) suggests a progressive increase in perceived breathlessness. This supports the notion that higher LEP scores are associated with a trajectory of worsening respiratory symptoms, even if the magnitude of change falls below the MCID threshold. These findings underscore the nuanced capability of our model in differentiating progression trajectories in COPD patients and highlight the potential of the LEP model in recognizing patients who might benefit from early interventions.

Model performance declined with increasing COPD severity as defined by GOLD stages. Although this may appear counterintuitive, advanced COPD—particularly at GOLD 4—is characterized by extensive parenchymal destruction and markedly reduced lung tissue density. This loss of density limits the information available for prediction, making it more difficult for the model to correlate LEP scores with emphysema progression. The non-linear trend in correlations across disease stages reflects the progressive degradation of lung structure and the resulting loss of distinct density patterns essential for accurate predictions. Despite the reduced predictive precision at later stages, the LEP score remains associated with mortality in individuals with greater emphysema burden, as the impact of emphysema on mortality increases with disease extent. To enhance model performance in advanced disease, we are investigating an extension of our local foundational model to a fully 3D representation of the lung, which may better capture the spatial complexity of severe emphysema. However, such a transition would require a larger sample size to avoid overfitting and ensure generalizability.

Fast-slow and high-low risk categories for emphysema progression were initially defined at the 70th percentile of participants in the “LEP score” test, though alternative thresholds are possible. We decided to use this percentile rather than the median value (50%) for two main reasons: (1) to define a population at higher risk than just the mean value and (2) to define a

population with a reasonable prevalence of 30%. However, the 75th percentile with a lower prevalence (25%) is also a valid approach. Using the 75th percentile we observed a decline in the model's performance by approximately 3% due to the reduced prevalence of subjects in the fast-low and high-low risk categories for emphysema progression.

Effects of variability, such as CT scanner type, acquisition and reconstruction, will influence our method to predict emphysema progression. Our findings indicate that reduced dose impacts the accuracy of the method by approximately 10%. To address the variability introduced by different CT scanner types, we trained our model using a multicenter cohort. This approach helps to mitigate this issue by encompassing a diverse range of scanner types and settings, enhancing the generalizability and robustness of our model across different clinical environments. The method used for co-registration was specifically designed to be robust against large deformations, with an error below 1.3 mm on the DIR-Lab dataset (expiration to inspiration)³⁴. Additionally, our registration is inherently less challenging since we are aligning longitudinal inspiration images, where the degree of lung volume change is less significant than in expiration to inspiration. Evaluating all possible confounders is challenging, as it would require scanning several patients with varying COPD statuses on different scanners and using different kernels. To overcome this limitation, we studied the effect of high and low doses to provide an idea of the impact of these confounders on our method. The data presented on reproducibility under significant changes in protocol show that the model is robust to changes in noise. To further investigate the impact of variability due to CT scanner type, we evaluated the model's performance across different scanner brands within the COPDGene cohort. The correlation between the LEP score and $\Delta\%LAA-950$ varied by approximately 12% relative to the correlation computed across all COPDGene participants. However, because the same subjects were not scanned across different scanner manufacturers, definitive conclusions regarding the effect of scanner type on variability cannot be drawn. Nonetheless, this comparison provides useful insights into the potential influence of scanner-related variability on model performance.

Clinical implications and integration into practice

Risk Stratification and Prognosis: The LEP score offers an objective, imaging-based biomarker for identifying individuals at elevated risk of emphysema progression without the need for longitudinal studies. This is particularly important in early stages of COPD (GOLD 1–2) and Pre-COPD (excluding PRISM), where traditional spirometry may not fully capture the risk of structural lung damage. By quantifying the likelihood of progression before significant lung function decline occurs, the LEP score enables early identification of high-risk individuals who may benefit from closer monitoring and earlier intervention.

Personalized Treatment Planning: The LEP activity maps generated by our model offer valuable insight into the spatial distribution of emphysema progression, thus offering a more comprehensive understanding of the disease. These maps could help clinicians understand lobe-specific patterns of disease progression, potentially enabling the development of personalized treatment strategies before irreversible damage occurs—for example, lobe-targeted interventions such as lung volume reduction. Additionally, our model opens new avenues for the mechanistic understanding of emphysema progression through imputed phenotypes. While new molecular platforms, such as spatial transcriptomics, are offering unique insights into the mechanisms underlying various diseases³⁵ destructive nature of longitudinal sampling. However, our model circumvents this limitation by providing an alternative, non-invasive means of assessing emphysema progression, leveraging spatial data without the need for repeated lung biopsies.

Clinical Trials and Drug Development: In research and drug development contexts, the LEP score could serve as an enrichment tool for clinical trials targeting emphysema-modifying therapies. By selecting participants with higher predicted progression rates, trials can improve statistical power and evaluate therapeutic effects more efficiently. Additionally,

the score could potentially serve as a surrogate endpoint to assess treatment response in early-phase studies.

While our study benefits from a large sample size that includes a substantial proportion of African American participants (30%), other racial and ethnic groups remain underrepresented. Moreover, our replication cohort consists predominantly of Non-Hispanic White individuals. These factors may limit the generalizability of our findings, particularly to populations not adequately represented in our training and validation cohorts. Future studies should aim to include more diverse populations to ensure equitable applicability of predictive models across all patient groups.

In addition, while our findings support the clinical potential of the LEP score, it is important to note that the model was trained and evaluated primarily on cohorts of smokers. As such, its applicability to non-smoker phenotypes—such as those observed in individuals with Alpha-1 antitrypsin deficiency or biomass exposure—remains to be validated. Ensuring relevance across the full spectrum of COPD phenotypes will require extending model development to include a broader range of exposure profiles and clinical characteristics.

Variability introduced by technical factors—such as differences in CT scanner type, acquisition protocols, and reconstruction algorithms—may also influence our model's ability to predict emphysema progression. Addressing these sources of heterogeneity will be critical for translating our findings into clinical practice—for example, through image harmonization techniques.

The heterogeneity and often non-linear nature of COPD progression underscore the importance of incorporating both imaging and functional markers when interpreting model predictions. Future work will aim to further investigate the biological mechanisms underlying our findings by integrating additional imaging features—such as gas trapping and airway wall metrics—alongside molecular and clinical data, to more precisely characterize the pathophysiological basis of the model's predictions across diverse COPD phenotypes. Expanding the size and diversity of the study population will also create opportunities to extend our local foundational model to a fully 3D representation of the lung. This advancement could enhance predictive accuracy, particularly in individuals with advanced disease (e.g., GOLD 4), by capturing the complex spatial patterns of parenchymal destruction that are often missed in lower-dimensional models.

Methods

Local emphysema progression (LEP) definition and ground truth generation

The method proposed was designed to identify local deep phenotypes by learning the local emphysema evolution according to the percentage of low-attenuation area (%LAA-950). In this way, we defined the Local Emphysema Progression (LEP) as those pixels that change from lung tissue at baseline CT scan to air at a 5-year follow-up co-registered CT scan, i.e. the pixel density decays from a value of > -950 HU at baseline to a value of ≤ -950 HU at 5-year follow-up co-registered CT scan. Supplementary Fig. 5 shows an example of the LEP definition where the pixels that change from lung tissue at baseline (a) to air (b) are plotted in green over the follow-up co-registered CT scan (c).

Deep local emphysema progression model

The method proposed for predicting the local evolution of the emphysema at 5-years combines a local foundational model of density evolution with a multilayer perceptron (Fig. 1)³⁶. The local FM is fitted to encode useful information about the evolution of the local density patterns at a 5-year follow-up co-registered CT scan. In particular, the local FM is trained in a neighborhood of 32 by 32 pixels (0.64 mm²) to reconstruct the same neighborhood in a 5-year follow-up co-registered CT scan. Using 2D patches instead of 3D patches allowed us to leverage a larger amount of training data, which is crucial for building a robust model. The trade-off between the complexity of predicting local 3D volumes and the training data requirements was carefully considered. Given the current computational constraints and the complexity of 3D reconstruction over a five-year period, we determined that a 2D approach would be more effective and stable for our

purposes. It is worth noting that our results based on a 2D local autoencoder are quite promising. Once the foundational model was successfully trained (examples of the reconstruction can be seen in Supplementary Fig. 2), a multilayer perceptron (MLP) was trained to regress the local emphysema progression or LEP, according to the %LAA-950. This process is carried out using a conditional strategy in three stages, i.e., for each batch in a particular epoch, the whole model is fitted in three sequential steps as follows:

- First, the MLP is fitted with the embeddings from the local FM, i.e., the output of the encoder (see Fig. 1).
- Second, the encoder part of the local FM is fine tuned to improve the embeddings representation of emphysema progression. In this step, the decoder part of the FM and the MLP model are not trained; they remain fixed.
- Third, the local FM is re-trained to maintain an acceptable local reconstruction. In this step the MLP model remains fixed.

The conditional training strategy allows the LEP model to keep a balance between encoding useful information of the emphysema progression into the embeddings, z , the local density reconstruction in the follow-up CT scan, and the predicted LEP. The reconstruction loss function, i.e., the local FM loss, is defined to incorporate information about the emphysema measured at the 5-year follow-up, the emphysema progression and the emphysema subtypes measured at baseline. The rationale for introducing emphysema subtyping is to acknowledge that different subtypes exhibit distinct patterns of progression based on their severity. The subtype masks are pre-computed at the pixel level using a local histogram-based emphysema (LHE) method described in refs. 37,38. These masks are only used for computing the loss function and are not involved in the inference process.

Model Architecture

The MLP is defined as a sequence of four dense blocks with 300 neurons and one with 200 neurons. Each block consists of a dense layer, a batch normalization and a dropout strategy with a ratio of 40%. According to the results described in ref. 39 we decided to use the swish function as activation. The output of the MLP corresponds to a sigmoid function, and the mean absolute error was proposed as a loss function to regress the LEP. In contrast to the simplicity of the MLP model, the autoencoder proposed makes use of two Kernel Sharing Atrous Convolution (KSAC)²⁶ with rates of 2 and 4, and five Convolutional Block Attention Modules (CBAM)²⁵ with different strides. These attention blocks are critical for capturing detailed information about emphysema progression and lung density decline. This mechanism is not required in the decoder path because our primary objective is not image reconstruction or segmentation but rather encoding useful information about density evolution and emphysema progression. This design choice allows us to focus on capturing detailed information about the formation process of parenchymal destruction without the influence of global features that could bias the results and reduce interpretability. A detailed description of this architecture can be found in the Supplementary Table 1. Like the MLP, the local FM uses the swish function as activation for all the layers except for the embeddings, z , and the output, where a linear activation is used. After each convolution, including those in KSAC and CBAM, a batch normalization strategy is applied.

Emphysema subtype specific loss function

The reconstruction loss function, i.e. the FM loss, is defined to incorporate information about the local emphysema measured at the 5-years follow-up, its subtypes measured at baseline and its progression. Let $\mathbf{y}^{(i)}$ a patch of 32 by 32 pixels in a 5-years co-register follow-up scan, and $\hat{\mathbf{y}}^{(i)}$ its predicted reconstruction. Then, the FM loss is defined as follows:

$$\begin{aligned} L_{\text{FM}} = & \frac{1}{N} \sum_j \lambda_0 L_{r||\text{lep}}(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}) + \lambda_1 L_{r||e}(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}) + \lambda_2 L_r(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}) \\ & + \lambda_3 L_{r||\text{lung}}(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}) \\ & + \lambda_4 L_{r||\text{epr}}(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}) + \sum_s \lambda_5 S_w(s) L_{r||s}(\mathbf{y}^{(i)}, \hat{\mathbf{y}}^{(i)}) \end{aligned} \quad (1)$$

where L_r corresponds to the mean absolute error

$$L_r(\mathbf{y}, \hat{\mathbf{y}}) = \sum_j |\mathbf{y}_j - \hat{\mathbf{y}}_j| \quad (2)$$

And $L_{r||(\cdot)}$ corresponds to the mean absolute error of the pixels with emphysema progression ($L_{r||\text{lep}}$), emphysema at 5-year ($L_{r||e}$), lung tissue ($L_{r||e}$), i.e., $> -950\text{HU}$, pixels predicted as emphysema ($L_{r||s}$), i.e., $\leq -950\text{ HU}$, or pixels with a particular emphysema subtype ($L_{r||s}$). We propose to use the emphysema subtype^{37,38} to weight each patch, $\mathbf{y}^{(i)}$, according to the weights, S_w , as follows: 0.4, 0.8, 1, 0.3, 0.2 and 0.2 for paraseptal emphysema, centrilobular emphysema (mild, moderate, and severe), panlobular emphysema and normal parenchyma respectively. The notion for introducing emphysema subtype is to acknowledge that different subtypes have different patterns of progression based on their severity. The weights were set a priori to reflect that severity, so different local patches belonging to different emphysema subtypes would be a different evolution.

Training

Our model was trained using 6.888 million of lung patches of 32 by 32 pixels (0.64 m^2) from 984 COPDGene participants randomly selected from our cohort (only 7000 patches by participant) keeping a balance between neighborhoods with a high/low number of pixels with emphysema progression (high $> 25\%$). A z-score normalization was applied to the lung regions according to the mean and std values of the training dataset. The optimization was performed using a stochastic gradient descent (Adaptive Moment Estimation) with a learning rate of 5×10^{-5} for pre-training training the AE, and the learning rates used in the conditional training strategy were set to 1×10^{-4} , 1×10^{-5} , and 1×10^{-5} for the MLP (stage 1), the encoder (stage 2) and the AE (stage 3) respectively.

Data augmentation was used to teach the model the desired invariance and robustness properties. Heterogeneity in the neighborhoods is needed for the network to learn flips and rotation. With this intention, the input of the network was randomly deformed by means of spatial flips (both axes) and a rotation of $\pm 90^\circ$. One or more deformation can be sequentially combined during the training and no interpolation is needed.

Local emphysema progression activity map

The inference of the LEP activity map is done by sliding the encoder plus the MLP across the whole lung with a stride of 4 pixels and aggregating the predicted probabilities. Finally, the local emphysema activity map is normalized according to the overlapping due to the sliding process.

Local emphysema progression (LEP) risk score

The LEP risk score is defined according to the following procedure: (1) the local emphysema progression at 5-year is estimated across the entire lung parenchyma with a stride of 4 pixels using the temporal encoder; (2) each region is defined as stable or progressing according to the threshold that maximizes the Youden's index on the training population (probability $\text{th} = 0.158$); (3) The LEP score is computed as the number of regions within the lung parenchyma predicted to progress at 5 years, as defined in (2), normalized by the total number of regions in the lung. This value ranges from 0, indicating that no regions are predicted to progress, to 1, indicating that all regions are predicted to progress. It represents the proportion of lung tissue at risk of future emphysema progression.

Data and statistical analyses

Multivariate linear models were used to explore the association of the LEP risk score and the change of the %LAA-950 ($\Delta\%LAA-950$) in relation to various COPD outcomes. The linear models were adjusted using baseline covariate variables commonly applied in studies of emphysema progression, such as the one described in¹⁰: age, gender, race, emphysema, BMI, TLC, smoking status, and packs per year. Results were reported as the adjusted covariance means and the ordinary least square

difference (Diff.) for each group (low/high risk of LEP, and slow/fast progressors according to $\Delta\%LAA-950$). Statistically significant results are considered with $p\text{-value} \leq 0.05$ on a Student's t-test. All data are presented as the mean (SD).

Data availability

The data that support the findings of this study are not openly available because they are the property of the COPDGene study. However, researchers can request access to the data upon request.

Code availability

The source code for deep-phenotypes detection and validation will be incorporated into the Chest Imaging Platform (www.chestimagingplatform.org) –the open-source software library that we have developed in our group over the last ten years and continue to support.

Received: 22 May 2025; Accepted: 31 July 2025;

Published online: 28 August 2025

References

1. Marin, L., Colombo, P., Bebawy, M., Young, P. M. & Traini, D. Chronic obstructive pulmonary disease: patho-physiology, current methods of treatment and the potential for simvastatin in disease management. *Expert Opin. Drug Deliv.* **8**, 1205–1220 (2011).
2. Valipour, A. et al. Patterns of emphysema heterogeneity. *Respiration* **90**, 402–411 (2015).
3. Castaldi, P. J. et al. Heterogeneity and progression of chronic obstructive pulmonary disease: emphysema-predominant and non-emphysema-predominant disease. *Am. J. Epidemiol.* **192**, 1647–1658 (2023).
4. World Health Organization. Chronic obstructive pulmonary disease (COPD). [https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-\(copd\)](https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)). 12, 2025 (2024).
5. Vos, T. et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1204–1222 (2020).
6. Martinez, F. J. et al. Predictors of mortality in patients with emphysema and severe airflow obstruction. *Am. J. Respir. Crit. Care Med* **173**, 1326–1334 (2006).
7. Gietema, H. A. et al. Distribution of emphysema in heavy smokers: Impact on pulmonary function. *Respir. Med* **104**, 76–82 (2010).
8. Lynch, D. A. et al. CT-based visual classification of emphysema: Association with mortality in the COPDGene study. *Radiology* **288**, 859–866 (2018).
9. Tang, X. et al. Emphysema is an independent risk factor for 5-year mortality in patients with bronchiectasis. *Clin. Respiratory J.* **11**, 887–894 (2017).
10. Ash, S. Y. et al. Relationship between emphysema progression at CT and mortality in ever-smokers: results from the COPDGene and ECLIPSE Cohorts. *Radiology* **299**, 222–231 (2021).
11. Newell, J. D., Hogg, J. C. & Snider, G. L. Report of a workshop: quantitative computed tomography scanning in longitudinal studies of emphysema. *Eur. Respiratory J.* **23**, 769–775 (2004).
12. Gietema, H. A. et al. Monitoring of smoking-induced emphysema with CT in a lung cancer screening setting: detection of real increase in extent of emphysema. *Radiology* **244**, 890–897 (2007).
13. Milne, S. & King, G. G. Advanced imaging in COPD: Insights into pulmonary pathophysiology. *J. Thorac. Dis.* <https://doi.org/10.3978/j.issn.2072-1439.2014.11.30> (2014).
14. Parr, D. G., Sevenoaks, M., Deng, C., Stoel, B. C. & Stockley, R. A. Detection of emphysema progression in alpha 1-antitrypsin deficiency using CT densitometry; Methodological advances. *Respir. Res.* <https://doi.org/10.1186/1465-9921-9-21> (2008).
15. Dirksen, A. et al. Exploring the role of CT densitometry: a randomised study of augmentation therapy in alpha 1-antitrypsin deficiency. *Eur. Respiratory J.* **33**, 1345–1353 (2009).
16. van Ginneken, B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radio. Phys. Technol.* **10**, 23–32 (2017).
17. Uppaluri, R., Mitsa, T., Sonka, M., Hoffman, E. A. & McLennan, G. Quantification of pulmonary emphysema from lung computed tomography images. *Am. J. Respir. Crit. Care Med.* **156**, 248–254 (1997).
18. Ginsburg, S. B., Lynch, D. A., Bowler, R. P. & Schroeder, J. D. Automated texture-based quantification of centrilobular nodularity and centrilobular emphysema in chest CT images. *Acad. Radio.* **19**, 1241–1251 (2012).
19. Serrano, G. G., Washko, G. R. & Estépar, R. S. J. Deep learning for biomarker regression: application to osteoporosis and emphysema on chest CT scans, in *Medical Imaging 2018: Image Processing*, E. D. Angelini and B. A. Landman, Eds., SPIE, 2018. <https://doi.org/10.1117/12.2293455>.
20. Singla, S. et al. Subject2Vec: generative-discriminative approach from a set of image patches to a vector. In *Medical Image Computing and Computer Assisted Intervention MICCAI 2018*, Springer International Publishing, 2018, 502–510.
21. Gorbunova, V. et al. Early detection of emphysema progression. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2010*, Springer Berlin Heidelberg, 2010, 193–200.
22. Staring, M. et al. Towards local estimation of emphysema progression using image registration, in *SPIE Proceedings*, J. P. W. Pluim and B. M. Dawant, Eds., SPIE, 2009. <https://doi.org/10.1117/12.811576>.
23. Coxson, H. O. et al. The presence and progression of emphysema in COPD as determined by CT scanning and biomarker expression: a prospective analysis from the ECLIPSE study. *Lancet Respir. Med.* **1**, 129–136 (2013).
24. Tanabe, N. et al. Longitudinal study of spatially heterogeneous emphysema progression in current smokers with chronic obstructive pulmonary disease, *PLoS One* <https://doi.org/10.1371/journal.pone.0044993> (2012).
25. Woo, S., Park, J., Lee, J. Y. & Kweon, I. S. CBAM: convolutional block attention module. *CoRR* <http://arxiv.org/abs/1807.06521> (2018)
26. Huang, Y. et al. See more than once: Kernel-sharing atrous convolution for semantic segmentation. *Neurocomputing* **443**, 26–34 (2021).
27. Regan, E. A. et al. Genetic epidemiology of COPD (COPDGene) study design. *COPD: J. Chronic Obstr. Pulm. Dis.* **7**, 32–43 (2010).
28. Global Initiative for Chronic Obstructive Lung Disease, Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease, 2023 report. Available at: https://goldcopd.org/wp-content/uploads/2022/12/GOLD-2023-ver-1.1-2Dec2022_WMV.pdf (Accessed: 12 August 2025).
29. JVestbo J. et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *Eur. Respir. J.* <https://doi.org/10.1183/09031936.00111707> (2008).
30. Tsutsumi, A. et al. Characteristics of chronic obstructive pulmonary disease patients with robust progression of emphysematous change. *Sci. Rep.* <https://doi.org/10.1038/s41598-021-87724-8> (2021).
31. Shaker, S. B. et al. Factors influencing the decline in lung density in a Danish lung cancer screening cohort. *Eur. Respir. J.* <https://doi.org/10.1183/09031936.00207911> (2012).
32. Jobst, B. J. et al. Effect of smoking cessation on quantitative computed tomography in smokers at risk in a lung cancer screening population. *Eur. Radiol.* <https://doi.org/10.1007/s00330-017-5030-6> (2018).
33. Shaker, S. B., Stavngaard, T., Laursen, L. C., Stoel, B. C. & Dirksen, A. Rapid fall in lung density following smoking cessation in COPD. *COPD: J. Chronic Obstruct. Pul. Dis* <https://doi.org/10.3109/15412555.2010.541306> (2011).

34. Tian, L. et al. GradICON: approximate diffeomorphisms via gradient inverse consistency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 18084–18094 (2023).
35. Zhou, R., Yang, G., Zhang, Y. & Wang, Y. Spatial transcriptomics in development and disease. *Mol. biomedicine* **4**, 32 (2023).
36. Curiale, A. H. & San Jose Estepar, R. Detection of local emphysema progression using conditional CNN, in *SPIE Medical Imaging* (San Diego, 2024).
37. Castaldi, P. J. et al. Distinct quantitative computed tomography emphysema patterns are associated with physiology and function in smokers. *Am. J. Respir. Crit. Care Med* **188**, 1083–1090 (2013).
38. Mendoza, C. S. et al. Emphysema quantification in a multi-scanner HRCT cohort using local intensity distributions. *9th IEEE Int. Symp. Biomed. Imaging (ISBI)* **2012**, 474–477 (2012).
39. Ramachandran, P., Zoph, B. & Le, Q. V. Searching for activation functions, *arXiv preprint arXiv:1710.05941* (2017).

Acknowledgements

This work was supported by U.S. National Institutes of Health (NIH) grant 1R01HL149877 and Alpha-1 Foundation grant 1037165.

Author contributions

Study concept and design: A.H.C. and R.S.J. E.A. cquisition, analysis, or interpretation: A.H.C., C.P., R.S.J.E., A.D., G.W. and R.S.J.E. Drafting of the manuscript: A.H.C. and R.S.J.E. Critical revision of the manuscript for important intellectual content: A.H.C., C.P., R.S.J.E., A.D., G.W. and R.S.J.E. Statistical analysis: A.H.C. and R.S.J.E. Read and approved the manuscript: A.H.C., C.P., R.S.J.E., A.D., G.W. and R.S.J.E. A.H.C. and R.S.J.E. had full access to data included in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-01917-3>.

Correspondence and requests for materials should be addressed to Ariel H. Curiale or Raúl. San José Estépar.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025