10.1148/radiol.2019191022

Appendix E1

Materials and Methods

Cohorts

COPDGene is an observational, longitudinal study on the genetic epidemiology of COPD (ClinicalTrials.gov: NCT00608764) funded by the National Heart, Lung, and Blood Institute (5,13). CT scans were obtained without intravenous contrast while the participant was coached to full inspiration and relaxed expiration. CT protocols for image acquisition and reconstruction were standardized across the multiple COPDGene centers. Inspiratory scans were acquired with a fixed dose of 200 mAs, which is considered high dose by today's standards but was typical at the time the COPDGene study was initiated. CT series were calculated with two different reconstruction filters, one smoother (ie, low spatial frequency) and one edge enhancing (ie, high spatial frequency). Our study used low spatial frequency reconstructions. Prospective data were collected through the Longitudinal Follow-up (LFU) portion of COPDGene. Participants were contacted in person or via e-mail every six months following enrollment and asked to report on intercurrent hospitalizations and medical events. Deaths were reported to the central study from the clinical centers. Information from the Social Security Death Index (SSDI) and the COPDGene longitudinal follow-up program was used to determine a survival or censoring time for each subject, taking care to avoid ascertainment bias. Those participants in the LFU program received regularly occurring, biannual phone calls, and for this reason, for those still undergoing active follow up, vital status was back censored six months prior to dataset generation. Those whose follow up time terminated in death were included if their contact in the prior six months indicated that they were being actively followed at the time of death. For those participants with vital status ascertained using the SSDI, deaths and vital status were back censored three months to account for the expected lag time between a death and its appearance in the SSDI dataset (18).

The ECLIPSE study was a three-year multicenter, observational study of 2,164 participants with GOLD stage 2–4 COPD, and 582 nonsmoker and smoker control participants that was completed in 2011 (4,17). The trial was approved by the ethics and review boards at participating centers and written informed consent was obtained from all participants. Individuals with known respiratory diseases other than COPD or severe alpha-1 antitrypsin deficiency were excluded. Participants ages 40 to 75 years were recruited to the COPD study group if they had a smoking history of ≥ 10 pack-years, a postbronchodilator forced expiratory volume in 1 second (FEV1) less than 80% predicted, and a postbronchodilator ratio between FEV1 and forced vital capacity (FVC) ≤ 0.7. Smoking (≥ 10 pack-years) and nonsmoking (< 1 pack-year) control participants were enrolled if they were ages 40 to 75 years and had normal lung function (postbronchodilator FEV1 > 85% predicted and FEV1/FVC > 0.7). Participants with known respiratory diseases other than COPD or severe alpha-1 antitrypsin deficiency were excluded. CT scanning was performed without bronchodilatation within 1 day of lung function testing. Images were acquired at suspended full inspiration without administration of intravenous contrast. Exposure settings were 120 kVp and 40 mAs and images were reconstructed using thin (≤ 1.25 mm) contiguous slices and a low spatial frequency reconstruction algorithm (4,17). The present

study included 1,962 ECLIPSE participants with available baseline CT, spirometry, and mortality data. Participants with unreadable CT, identified by thoracic radiologists during the original study, were omitted. Table E1 compares COPDGene and ECLIPSE testing cohorts.

Statistical Analysis

The resampling-based test (20) for assessment of model calibration was carried out by simulating an emphysema level (integer between 0 and 5) for each participant using the probability profiles predicted by the algorithm. A χ^2 (χ^2) statistic was computed by comparing simulated and expected counts across participants. This was repeated 10,000 times to create a null distribution. The observed levels from the visual scoring were then compared with the null distribution using the same χ^2 statistic, replacing the simulated data with the actual visual scores, with small values indicating that visual and deep learning scores are consistent. In this test is the *P* value is the proportion of values in the null distribution with values at least as large as the observed one.

Deep Learning Algorithm

Convolutional neural networks (CNN) are a class of computational models in deep learning that are particularly well-suited to image analysis. Development of CNNs for processing volumetric CT at full resolution can challenge the capabilities of current computer hardware. Memory constraints of consumer grade graphics processing units (GPUs), which are relied upon to perform intensive computations in training CNNs, impose limits on the amount of image data that can be used per participant. To help address memory issues, our deep learning algorithm combines convolutional neural network (CNN) and long-short term memory (LSTM) architectures (33). An initial lung segmentation is used to determine three-dimensional bounding box around the lungs. The lung segmentation is generated by a deep learning algorithm consisting of a fully convolutional DenseNet architecture (34) that was trained on full resolution axial and coronal CT from normal and diseased lungs. A preprocessing step standardized CT pixel data to z-scores using mean and standard deviation calculated within the lung segmentation volumes. Twenty-five (25) full resolution axial images sampled over the standardized lung volume (omitting the superior and inferior 5 mm) are separately processed by the convolutional blocks of the network to extract features. Convolutional blocks consist of two-dimensional (2D) convolutions, rectified linear unit (RELU) activation and max pooling. The four 2D convolutional layers have 32 6 × 6, 96 3 × 3, 256 3 × 3 and 384 3 × 3 filters, respectively. The first two max pooling layers have stride 3 and the second two max pooling layers have stride 4. The convolutional blocks used shared weights so each of the 25 axials is processed in the same way. Output of convolutional blocks for the 25 images are flattened and concatenated into a sequence and passed to the LSTM layer, which learns representations of sequences that are useful for classification. This reduces the dimensionality of the whole CT features from 25 × 3456 to 1 × 1228. Following the LSTM layer are two linear dense layers (1024 and 6 nodes respectively) and negative log likelihood was used as the loss function. Dropout is used between the LSTM and Dense layers. The output layer produces the predicted probabilities ($0.0 \le p_i \le 1.0$) for classification in each category (c_i, with i = 0,1,2,3,4 or 5 representing the categories absent, trace, mild, moderate, confluent and advanced destructive, respectively). This is treated as a discrete probability distribution and the final prediction (c_{pred}) is the probability-weighted average (35) of the categories rounded to the nearest integer. The model was trained on CT on 2,507 COPDGene participants with three geometric augmentations (random variation in axial slice sampling, in-plane translation, and in-plane rotation) for an effective training set size of

7,521. Table E2 describes characteristics of participants used for training. Computer systems used for algorithm development, training and testing included Intel Core i9 7980XE CPU, Dual NVIDIA GeForce GTX 1080Ti and Titan XP GPUs, and 64GB RAM. Training time for the algorithm was approximately four days.

Results

In comparing visual and deep learning scores in the COPDGene test cohort, the greatest discordance was in individuals without visual evidence of emphysema that were classified by the deep learning algorithm as having trace emphysema (ie, the two leftmost cells along the first row of Table 1). Table E3 compares individuals with no visual evidence of emphysema that were classified by the deep learning algorithm as absent or trace emphysema. Compared with participants classified by both visual assessment and deep learning as having no emphysema (n = 637), participants classified as having trace emphysema by the deep learning algorithm (n = 1,495) had diminished FEV1% predicted (90.7 versus 93.7, $P < .001$) and FEV1/FVC (0.77 versus 0.79, $P < .001$). They also had greater LAA-950 (2.3 versus 2.0, $P = .013$) and greater smoking exposure (35.6 versus 32.1 pack-years, $P < .001$).

Figure E1 shows Kaplan-Meier plots of survival in the COPDGene cohort with axes limits matching Figure 3 in the main paper.

Table E5 shows results of multivariable Cox modeling using visual emphysema scores as predictors. Confirming results of a prior study, base models adjusted for race, sex, age, weight, height, smoking pack-years, current smoking status, and education level show lower survival associated with higher visual emphysema grade. Estimated hazard ratios (HR) were 1.2 (95% CI 1.0, 1.6), 1.4 (95% CI 1.2, 1.8), 2.2 (95% CI 1.7, 2.7), 4.6 (95% CI 3.7, 5.8), or 4.7 (95% CI 3.5, 6.2) for visual grades of trace, mild, moderate, confluent or advanced destructive emphysema, respectively.

Tables E6a and E6b show cause of death in the COPDGene test cohort (n = 740) by visual and deep learning emphysema scores, respectively. Cause of death is encoded using TORCH (Toward a Revolution in COPD Health [36]) Underlying Cause of Death (UCD) conventions. χ^2 tests of independence show an association between categories of emphysema severity and cause of death ($\chi^2 = 167.1$ $P < .001$ and $\chi^2 = 173.7$ $P < .001$ for visual and deep learning emphysema classifications, respectively).

COPDGene® Investigators – Core Units

Administrative Center: James D. Crapo, MD (PI); Edwin K. Silverman, MD, PhD (PI); Barry J. Make, MD; Elizabeth A. Regan, MD, PhD

Genetic Analysis Center: Terri Beaty, PhD; Ferdouse Begum, PhD; Peter J. Castaldi, MD, MSc; Michael Cho, MD; Dawn L. DeMeo, MD, MPH; Adel R. Boueiz, MD; Marilyn G. Foreman, MD, MS; Eitan Halper-Stromberg; Lystra P. Hayden, MD, MMSc; Craig P. Hersh, MD, MPH; Jacqueline Hetmanski, MS, MPH; Brian D. Hobbs, MD; John E. Hokanson, MPH, PhD; Nan Laird, PhD; Christoph Lange, PhD; Sharon M. Lutz, PhD; Merry-Lynn McDonald, PhD; Margaret M. Parker, PhD; Dmitry Prokopenko, Ph.D; Dandi Qiao, PhD; Elizabeth A. Regan, MD, PhD; Phuwanat Sakornsakolpat, MD; Edwin K. Silverman, MD, PhD; Emily S. Wan, MD; Sungho Won, PhD

Imaging Center: Juan Pablo Centeno; Jean-Paul Charbonnier, PhD; Harvey O. Coxson, PhD; Craig J. Galban, PhD; MeiLan K. Han, MD, MS; Eric A. Hoffman, Stephen Humphries, PhD; Francine L. Jacobson, MD, MPH; Philip F. Judy, PhD; Ella A. Kazerooni, MD; Alex Kluiber; David A. Lynch, MB; Pietro Nardelli, PhD; John D. Newell, Jr., MD; Aleena Notary; Andrea Oh, MD; Elizabeth A. Regan, MD, PhD; James C. Ross, PhD; Raul San Jose Estepar, PhD; Joyce Schroeder, MD; Jered Sieren; Berend C. Stoel, PhD; Juerg Tschirren, PhD; Edwin Van Beek, MD, PhD; Bram van Ginneken, PhD; Eva van Rikxoort, PhD; Gonzalo Vegas Sanchez-Ferrero, PhD; Lucas Veitel; George R. Washko, MD; Carla G. Wilson, MS;

PFT QA Center, Salt Lake City, UT: Robert Jensen, PhD

Data Coordinating Center and Biostatistics, *National Jewish Health, Denver, CO*: Douglas Everett, PhD; Jim Crooks, PhD; Katherine Pratte, PhD; Matt Strand, PhD; Carla G. Wilson, MS

Epidemiology Core, *University of Colorado Anschutz Medical Campus, Aurora, CO*: John E. Hokanson, MPH, PhD; Gregory Kinney, MPH, PhD; Sharon M. Lutz, PhD; Kendra A. Young, PhD

Mortality Adjudication Core: Surya P. Bhatt, MD; Jessica Bon, MD; Alejandro A. Diaz, MD, MPH; MeiLan K. Han, MD, MS; Barry Make, MD; Susan Murray, ScD; Elizabeth Regan, MD; Xavier Soler, MD; Carla G. Wilson, MS

Biomarker Core: Russell P. Bowler, MD, PhD; Katerina Kechris, PhD; Farnoush Banaei-Kashani, PhD

COPDGene® Investigators – Clinical Centers
Ann Arbor VA: Jeffrey L. Curtis, MD; Perry G. Pernicano, MD

Baylor College of Medicine, Houston, TX: Nicola Hanania, MD, MS; Mustafa Atik, MD; Aladin Boriek, PhD; Kalpatha Guntupalli, MD; Elizabeth Guy, MD; Amit Parulekar, MD;

Brigham and Women's Hospital, Boston, MA: Dawn L. DeMeo, MD, MPH; Alejandro A. Diaz, MD, MPH; Lystra P. Hayden, MD; Brian D. Hobbs, MD; Craig Hersh, MD, MPH; Francine L. Jacobson, MD, MPH; George Washko, MD

Columbia University, New York, NY: R. Graham Barr, MD, DrPH; John Austin, MD; Belinda D'Souza, MD; Byron Thomashow, MD

Duke University Medical Center, Durham, NC: Neil MacIntyre, Jr., MD; H. Page McAdams, MD; Lacey Washington, MD

HealthPartners Research Institute, Minneapolis, MN: Charlene McEvoy, MD, MPH; Joseph Tashjian, MD

Johns Hopkins University, Baltimore, MD: Robert Wise, MD; Robert Brown, MD; Nadia N. Hansel, MD, MPH; Karen Horton, MD; Allison Lambert, MD, MHS; Nirupama Putcha, MD, MHS

Lundquist Institute for Biomedical Innovationat Harbor UCLA Medical Center, Torrance, CA: Richard Casaburi, PhD, MD; Alessandra Adami, PhD; Matthew Budoff, MD; Hans Fischer, MD; Janos Porszasz, MD, PhD; Harry Rossiter, PhD; William Stringer, MD

Michael E. DeBakey VAMC, Houston, *TX*: Amir Sharafkhaneh, MD, PhD; Charlie Lan, DO

Minneapolis VA: Christine Wendt, MD; Brian Bell, MD; Ken M. Kunisaki, MD, MS

Morehouse School of Medicine, Atlanta, GA: Marilyn G. Foreman, MD, MS; Eugene Berkowitz, MD, PhD; Gloria Westney, MD, MS

National Jewish Health, Denver, CO: Russell Bowler, MD, PhD; David A. Lynch, MB

Reliant Medical Group, Worcester, MA: Richard Rosiello, MD; David Pace, MD

Temple University, Philadelphia, PA: Gerard Criner, MD; David Ciccolella, MD; Francis Cordova, MD; Chandra Dass, MD; Gilbert D'Alonzo, DO; Parag Desai, MD; Michael Jacobs, PharmD; Steven Kelsen, MD, PhD; Victor Kim, MD; A. James Mamary, MD; Nathaniel Marchetti, DO; Aditi Satti, MD; Kartik Shenoy, MD; Robert M. Steiner, MD; Alex Swift, MD; Irene Swift, MD; Maria Elena Vega-Sanchez, MD

University of Alabama, Birmingham, AL: Mark Dransfield, MD; William Bailey, MD; Surya P. Bhatt, MD; Anand Iyer, MD; Hrudaya Nath, MD; J. Michael Wells, MD

University of California, San Diego, CA: Douglas Conrad, MD; Xavier Soler, MD, PhD; Andrew Yen, MD

University of Iowa, Iowa City, IA: Alejandro P. Comellas, MD; Karin F. Hoth, PhD; John Newell, Jr., MD; Brad Thompson, MD

University of Michigan, Ann Arbor, MI: MeiLan K. Han, MD MS; Ella Kazerooni, MD MS; Wassim Labaki, MD MS; Craig Galban, PhD; Dharshan Vummidi, MD

University of Minnesota, Minneapolis, MN: Joanne Billings, MD; Abbie Begnaud, MD; Tadashi Allen, MD

University of Pittsburgh, Pittsburgh, PA: Frank Sciurba, MD; Jessica Bon, MD; Divay Chandra, MD, MSc; Carl Fuhrman, MD; Joel Weissfeld, MD, MPH

University of Texas Health, San Antonio, San Antonio, TX: Antonio Anzueto, MD; Sandra Adams, MD; Diego Maselli-Caceres, MD; Mario E. Ruiz, MD; Harjinder Singh

References

33. Humphries SM, Notary AM, Centeno JP, Lynch DA. Automatic Classification of Centrilobular Emphysema on CT Using Deep Learning: Comparison with Visual Scoring. In: Stoyanov D, Taylor Z, Kainz B, et al, eds. Image Analysis for Moving Organ, Breast, and Thoracic Images. RAMBO 2018, BIA 2018, TIA 2018. Lecture Notes in Computer Science, vol 11040. Cham, Switzerland: Springer, 2018; 319–325.

34. Jégou S, Drozdzal M, Vazquez D, Romero A, Bengio Y. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway, NJ: IEEE, 2017; 1175–1183.

35. Casella G, Berger RL. Statistical inference. Vol 2. Pacific Grove, Calif: Duxbury, 2002.

36. Vestbo J; TORCH Study Group. The TORCH (towards a revolution in COPD health) survival study protocol. Eur Respir J 2004;24(2):206–210.

Table E1

Characteristics of COPDGene participants (*n* = 2,507) used to train deep learning algorithm stratified by visual emphysema score

	Emphysema grade–visual scoring						
Parameter	Absent	Trace	Mild centrilobular	Moderate centrilobular	Confluent	Adv. destructive	*P* value*
No. participants† (*n* = 2507)	782 (31)	431 (17)	477 (19)	430 (17)	276 (11)	111 (4)	
No. deaths‡ (*n* = 431)	50 (6)	55 (13)	91 (19)	90 (21)	95 (34)	50 (45)	

Demographic data							
Age (years)	56.5 ± 8.6	56.5 ± 8.4	58.9 ± 8.9	63.2 ± 9.2	66.4 ± 7.5	64.6 ± 7.8	<0.001
BMI	30.6 ± 6.2	29.4 ± 6.3	28.0 ± 5.8	27.2 ± 5.4	26.7 ± 5.8	24.4 ± 5.3	<0.001
No. of men	267	173	190	171	120	55	0.01
No. of pack-years smoked	36.1 ± 20.0	42.6 ± 21.4	47.9 ± 23.7	55.7 ± 30.0	59.0 ± 30.2	52.0 ± 23.9	<0.001
Functional parameters							
FEV1% pred.	90.2 ± 18.2	82.2 ± 19.0	78.5 ± 21.2	61.1 ± 24.9	46.2 ± 23.0	31.9 ± 15.0	<0.001
FEV1/FVC	0.77 ± 0.08	0.73 ± 0.11	0.69 ± 0.12	0.55 ± 0.16	0.43 ± 0.14	0.35 ± 0.11	<0.001
LAA-950	1.6 ± 2.2	1.6 ± 2.7	2.5 ± 3.5	8.4 ± 8.0	20.4 ± 11.2	36.0 ± 12.1	<0.001

Table elements are number of participants with percentage in parentheses or mean ± SD, as indicated. Percentages were calculated as number of participants in a table cell divided by number of participants classified in that grade of emphysema.

* Welch two sample *t* test.

† Percentages calculated as number of participants with a given emphysema grade divided by number of participants included in the training set.

‡ Mortality data not available on 710 of the training participants.

Table E2

Comparison of COPDGene (*n* = 7,143) and ECLIPSE (*n* = 1,962) testing cohorts

	COPDGene	ECLIPSE	*P* value*
Age (years)	59.8 ± 8.9	61.5 ± 8.4	<0.001
Number of deaths	982 (14)	155 (8)	
FEV1% predicted	77.8 ± 24.9	58.3 ± 30.4	<0.001
FEV1/FVC	0.67 ± 0.16	0.51 ± 0.17	<0.001
MMRC	1.28 ± 1.42	1.33 ± 1.14	0.13
Smoking pack-years	43.7 ± 24.6	41.5 ± 28.8	0.002
GOLD			
Nonsmoker controls	37 (0)	183 (9)	
PRISM	832 (12)	0 (0)	
0	3159 (44)	250 (13)	
1	573 (8)	2 (0)	
2	1378 (19)	690 (35)	
3	780 (11)	624 (32)	
4	339 (5)	213 (11)	

Table elements are number of participants with percentages in parentheses or mean ± SD. Percentages were calculated as number of participants in a table cell divided by total number of participants in that cohort.

* Welch two sample *t* test.

FEV1% = forced expiratory volume in one second percent predicted for age and sex, FVC = forced vital capacity, MMRC = modified Medical Research Council, GOLD = Global Initiative for Obstructive Lung Disease, PRISm = Preserved ratio impaired spirometry.

Table E3

Demographic and clinical parameters for COPDGene testing participants without visual evidence of emphysema on CT (*n* = 2132) classified as having no or trace emphysema

	Deep learning classifications				
	Absent		Trace		*P* value*
	mean	95% CI	mean	95% CI	
Number of participants	*n* = 637		*n* = 1495		
Age (years)	56.9	56.3, 57.6	57.4	57.0, 57.9	0.21
FEV1% predicted	93.9	92.8, 94.9	90.7	89.9, 91.6	<0.001
FEV1/FVC	0.79	0.79, 0.80	0.77	0.76, 0.77	<0.001
%LAA-950	2.01	1.82, 2.20	2.31	2.17, 2.45	0.01
Smoking pack-years	32.1	30.8, 33.4	35.6	34.6, 36.6	<0.001
SGRQ	14.9	13.6, 16.3	18.0	17.0, 19.0	<0.001
MMRC	0.71	0.63, 0.80	0.85	0.79, 0.92	0.01
6MWD (m)	472.5	464.9, 480.1	456.3	450.5, 462.0	<0.001

Table elements are mean values and 95% confidence intervals, except first row.

* Welch two sample *t* test.

Table E4

Mortality, demographics, functional parameters and comorbidities in COPDGene testing cohort (*n* = 7,143) according to visual grade of emphysema

	Emphysema grade–visual scoring						
Parameter	Absent	Trace	Mild centrilobular	Moderate centrilobular	Confluent	Adv. destructive	*P* value*
No. participants†	2499 (35)	1322 (19)	1409 (20)	1049 (15)	656 (9)	208 (3)	
No. deaths (*n* = 982)	162 (6)	116 (9)	163 (12)	203 (19)	254 (39)	84 (40)	
Demographic data							
Age (years)‡	57.5 ± 8.6	58.1 ± 8.8	59.9 ± 8.6	63.0 ± 8.3	65.0 ± 7.7	65.0 ± 7.5	<0.001
BMI (kg/m^2)	30.7 ± 6.5	29.7 ± 6.2	27.9 ± 5.7	27.7 ± 5.5	26.1 ± 5.5	25.2 ± 4.6	<0.001
No. of men)	1217 (49)	741 (56)	779 (55)	558 (53)	319 (49)	120 (58)	<0.001
Race							
Non-Hispanic white	1707 (68)	834 (63)	924 (66)	742 (71)	530 (81)	174 (84)	<0.001
African American	792 (32)	488 (37)	485 (34)	307 (29)	126 (19)	34 (16)	
No. of pack-years smoked (*n* = 7104)	35.3 ± 19.9	41.2 ± 23.3	46.9 ± 23.7	52.5 ± 26.5	55.2 ± 27.6	58.0 ± 28.3	<0.001
Current smoker	1204 (48)	799 (60)	861 (61)	524 (50)	171 (26)	32 (15)	<0.001
Education high school or less	770 (31)	515 (39)	585 (42)	437 (42)	243 (37)	83 (40)	<0.001
Functional parameters							
GOLD stage (*n* = 7098)						χ^2 = 3375.7††	<0.001
Nonsmoker control	34 (1)	3 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
PRISM	390 (16)	220 (17)	153 (11)	59 (6)	10 (2)	0 (0)	
0	1665 (67)	698 (53)	579 (41)	194 (18)	22 (3)	1 (0)	
1	126 (5)	102 (8)	153 (11)	134 (13)	49 (7)	9 (4)	

2	222 (9)	201 (15)	348 (25)	369 (35)	189 (29)	49 (24)	
3	43 (2)	76 (6)	142 (10)	217 (21)	236 (36)	66 (32)	
4	6 (0)	7 (1)	24 (2)	75 (7)	144 (22)	83 (40)	
FEV1% pred. (*n* = 7098)	89.9 ± 17.8	83.7 ± 19.7	78.6 ± 21.9	66.1 ± 24.0	48.4 ± 22.5	39.9 ± 20.5	<0.001
FEV1/FVC (*n* = 7098)	0.77 ± 0.08	0.73 ± 0.10	0.67 ± 0.12	0.58 ± 0.14	0.45 ± 0.14	0.39 ± 0.12	<0.001
6-Minute walk distance (m) (*n* = 7070)	454.3 ± 111.9	429.4 ± 119.1	425.1 ± 111.8	391.1 ± 120.3	348.6 ± 121.2	327.8 ± 115.4	<0.001
SGRQ	18.3 ± 19.5	22.8 ± 21.1	25.2 ± 21.7	32.95 ± 22.8	42.15 ± 20.5	45.9 ± 18.9	<0.001
MMRC dyspnea score (*n* = 7131)	0.87 ± 1.26	1.07 ± 1.34	1.23 ± 1.37	1.65 ± 1.45	2.35 ± 1.33	2.67 ± 1.21	<0.001
LAA-950 (%)	2.38 ± 2.96	2.39 ± 3.42	3.68 ± 4.76	8.66 ± 7.76	21.83 ± 11.27	33.4 ± 11.44	<0.001
Comorbidities							
Chronic bronchitis	334 (13)	212 (16)	292 (21)	266 (25)	179 (27)	42 (20)	<0.001
Severe exacerbations last year	135 (5)	105 (8)	145 (10)	164 (16)	151 (23)	52 (25)	<0.001
Coronary artery disease	115 (5)	88 (7)	98 (7)	92 (9)	67 (10)	18 (9)	<0.001
Diabetes	361 (14)	207 (16)	150 (11)	105 (10)	60 (9)	11 (5)	<0.001
Congestive heart failure	51 (2)	35 (3)	50 (4)	37 (4)	26 (4)	4 (2)	0.016

Table elements are number of participants with percentage in parentheses or mean ± SD. Percentages were calculated as number of participants in table cell divided by number of participants classified in that grade of emphysema (ie, values in top row).

* *P* value for differences across emphysema grades, calculated with χ^2 test for categoric variables and with *F* test from analysis of variance for continuous variables.

† Percentages are according to total number of participants.

†† Chi-squared test statistic comparing emphysema grade and GOLD stage

Table E5

Cox multivariable models for predicting mortality in COPDGene test cohort (*n* = 7,143 participants)

		Model 1: Base model			Model 2: Base model + LAA-950		
Parameter	Referent group	Hazard ratio	95% CI	*P* value	Hazard ratio	95% CI	*P* value
Trace	Absent	1.23	0.97, 1.57	0.0930	1.23	0.96,1.56	0.10
Mild	Absent	1.44	1.15, 1.80	0.0016	1.37	1.09, 1.71	0.006
Moderate	Absent	2.15	1.72, 2.67	<0.0001	1.66	1.32, 2.08	<0.001
Confluent	Absent	4.63	3.72, 5.77	<0.0001	2.22	1.68, 2.91	<0.001
Adv. destructive	Absent	4.67	3.52, 6.21	<0.0001	1.45	0.99, 2.11	0.06
LAA-950					1.04	1.03, 1.05	<0.001

Models were fit using visual emphysema classification scores.

Note.—Models are adjusted for age, race, sex, weight, height, smoking pack-years, current smoking status at enrollment, and education level.

Table E6

Cause of death analysis in COPDGene participants with TORCH (Toward a Revolution in COPD Health (10)) Underlying Cause of Death (UCD), single cause data available (*n* = 740)

a) Visual scores (χ^2 = 167.1, *P* < .001)

	Visual emphysema score					
	0	1	2	3	4	5
No. of deaths	110	88	122	152	206	62
UCD						
Cancer	25 (23)	27 (31)	39 (32)	42 (28)	28 (14)	14 (23)
Cardiovascular	30 (27)	19 (22)	26 (21)	28 (18)	18 (9)	2 (3)
Respiratory	14 (13)	7 (8)	26 (21)	62 (41)	124 (60)	40 (64)
Other	32 (29)	30 (34)	25 (20)	16 (11)	27 (13)	6 (10)
Unknown	9 (8)	5 (6)	6 (5)	4 (3)	9 (4)	0 (0)

b) Deep learning emphysema scores (χ^2 = 173.7, *P* < .001)

	Deep learning emphysema scores					
	0	1	2	3	4	5
No. of deaths	56	80	131	193	208	72
UCD						
Cancer	14 (25)	24 (30)	48 (37)	45 (23)	35 (17)	9 (13)
Cardiovascular	14 (25)	15 (19)	29 (22)	42 (22)	18 (9)	5 (7)
Respiratory	3 (5)	7 (9)	20 (15)	71 (37)	122 (59)	50 (69)
Other	21 (38)	27 (34)	28 (21)	28 (15)	24 (12)	8 (11)
Unknown	4 (7)	7 (9)	6 (5)	7 (4)	9 (4)	0 (0)

Data are counts (% of total). χ^2 tests of independence were used to compare emphysema grade and UCD. Table elements are number of participants with percentage in parentheses, as indicated. Percentages were calculated as number of participants in table cell divided by number of participants classified in that grade of emphysema (ie, values in top row). Emphysema scores are 0 = absent, 1 = trace, 2 = mild, 3 = moderate, 4 = confluent, 5 = advanced destructive.