

# Quantifying the Extent of Emphysema:

## *Factors Associated with Radiologists' Estimations and Quantitative Indices of Emphysema Severity Using the ECLIPSE Cohort*

Hester A. Gietema, MD, PhD, Nestor L. Müller, MD, PhD, Paola V. Nasute Fauerbach, MD, Sanjay Sharma, MSc, Lisa D. Edwards, PhD, Pat G. Camp, PhD, Harvey O. Coxson, PhD, on behalf of the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) investigators

**Rationale and Objectives:** This study investigated what factors radiologists take into account when estimating emphysema severity and assessed quantitative computed tomography (CT) measurements of low attenuation areas.

**Materials and Methods:** CT scans and spirometry were obtained on 1519 chronic obstructive pulmonary disease (COPD) subjects, 269 smoker controls, and 184 nonsmoker controls from the Evaluation of COPD Longitudinally to Identify Surrogate Endpoints (ECLIPSE) study. CT scans were analyzed using the threshold technique ( $\% < -950\text{HU}$ ) and a low attenuation cluster analysis. Two radiologists scored emphysema severity (0 to 5 scale), described the predominant type and distribution of emphysema, and the presence of suspected small airways disease.

**Results:** The percent low attenuation area (%LAA) and visual scores of emphysema severity correlated well ( $r = 0.77, P < .001$ ). %LAA, low attenuation cluster analysis, and absence of radiologist described gas trapping, distribution, and predominant type of emphysema were predictors of visual scores of emphysema severity (all  $P < .001$ ). CT scans scored as showing regions of gas trapping had smaller lesions for a similar %LAA than those without ( $P < .001$ ).

**Conclusions:** Visual estimates of emphysema are not only determined by the extent of LAA, but also by lesion size, predominant type, and distribution of emphysema and presence/absence of areas of small airways disease. A computer analysis of low attenuation cluster size helps quantitative algorithms discriminate low attenuation areas from gas trapping, image noise, and emphysema.

**Key Words:** Emphysema; chronic obstructive pulmonary disease; computed tomography; quantitative CT; small airways disease.

©AUR, 2011

The introduction of computed tomography (CT) has changed the way that clinicians diagnose and quantify the extent of emphysema in living individuals. It was

recognized early on that the frequency distribution of x-ray attenuation values in a CT image of the lung (CT densitometry) could produce an estimate of the extent of emphysema (1–3). Even though there has been a great deal of attention given to densitometric assessment of emphysema (4–8), the daily clinical routine is still to visually grade disease extent and severity.

Both densitometry (1,3,4) and visual grading of emphysema extent (9–11) have been shown to correlate well with the extent of emphysema on histology specimens. However, the estimations produced by densitometry have been reported to be similar to the extent of emphysema in histology specimens, whereas visual estimations tend to overestimate the extent of emphysema (12,13). Visual scores, on the other hand, have been reported to show stronger associations with spirometry data (14), the core diagnostic test to detect and stage chronic obstructive pulmonary disease (COPD) (15).

Another advantage of visual scoring is the ability of radiologists to distinguish emphysema from other causes of lowered attenuation such as image noise and hyperinflated regions

**Acad Radiol 2011; 18:661–671**

From the Department of Radiology, University of British Columbia, 855 West 12th Ave., Room CP-G22A, Vancouver General Hospital, Vancouver, BC V5Z 1M9, Canada (H.A.G., N.L.M., P.V.N.F., H.O.C.); University of British Columbia James Hogg Research Centre, Providence Heart + Lung Institute, St. Paul's Hospital, Vancouver, BC, Canada (H.A.G., P.G.C., H.O.C.); Department of Radiology, University Medical Centre, Utrecht, The Netherlands (H.A.G.); GlaxoSmithKline R&D, Research Triangle Park, NC (S.S., L.D.E.). Received October 21, 2010; accepted January 17, 2011. The ECLIPSE study is entirely funded by GlaxoSmithKline. ECLIPSE is registered at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) no. NCT00292552; GSK Study Code SCO104960. H.A.G. was supported in part by a grant from the American Thoracic Society during the time of this study. H.O.C. is supported by a GSK Clinical Scientist Award and, in part, by the University of Pittsburgh COPD SCCOR NIH 1P50 HL084948 and R01 HL085096 from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, to the University of Pittsburgh. H.O.C. was a CIHR/BC Lung Association New Investigator during the time of this research. **Address correspondence to:** H.O.C. e-mail: [harvey.coxson@vch.ca](mailto:harvey.coxson@vch.ca)

©AUR, 2011

doi:10.1016/j.acra.2011.01.011

consistent with “gas-trapping” caused by small airways disease. Although densitometry can quantify the spatial distribution of low attenuation voxels, it cannot discriminate other features within the CT scan that may help decide the cause of the low attenuation areas. For these reasons, investigators have attempted to develop new techniques using multiple features within a CT image (16). The most simple of these approaches examines how the low attenuation voxels are clustered together (17,18). Low attenuation areas consistent with gas-trapping and image noise are typically characterized by small clusters of voxels with decreased attenuation, whereas emphysema often results in larger clusters, especially in the advanced stages. Radiologists tend to look for those larger areas of low attenuation areas and use the size of destroyed areas in their estimation of the extent of emphysema, but the use of cluster analysis in computer based quantification of emphysema is limited (17,19,20). Although low attenuation cluster (LAC) analysis alone has been shown not to be able to estimate extent of emphysema (19), we hypothesized that the LAC analysis may be able to help to understand the relationship between densitometry results and visual scores.

The purpose of the current analysis is to investigate the relationship between densitometry results and visual scores by assessing factors that determine estimations of emphysema extent as seen by radiologists. To accomplish this we used the baseline CT scans from the Evaluation Of COPD Longitudinally To Identify Predictive Surrogate Endpoints (ECLIPSE) (NCT00292552; GSK Study Code SCO104960) cohort, which is a large cohort of smoking and nonsmoking subjects with and without airflow limitation (21–23).

## MATERIAL AND METHODS

### Subjects

The ECLIPSE Study is a 3-year noninterventional longitudinal study designed to discover and validate novel and robust metrics of disease and disease progression (23). The trial was approved by the ethics and review boards at participating centers and written informed consent was obtained from all subjects. A list of all participating centers can be found in Appendix 1. Individuals ages 40 to 75 years were recruited to the COPD study group if they had a smoking history of  $\geq 10$  pack-years, a postbronchodilator forced expiratory volume in 1 second ( $FEV_1$ ) less than 80% predicted, and a postbronchodilator ratio between  $FEV_1$  and forced vital capacity (FVC)  $\leq 0.7$ . Smoking ( $\geq 10$  pack-years) and nonsmoking ( $< 1$  pack-year) control subjects were enrolled if they were ages 40 to 75 years and had normal lung function (postbronchodilator  $FEV_1 > 85\%$  predicted and  $FEV_1/FVC > 0.7$ ). Individuals recruited to the study were genotyped for  $\alpha_1$ -antitrypsin deficiency and six homogenous protease inhibitor variant Z (PiZZ) and 11 heterozygous variants SZ (PiSZ) individuals were identified and excluded from the analysis (24).

### Pulmonary Function Testing

Postbronchodilator spirometry (ie,  $FEV_1$  and FVC) was performed following the American Thoracic Society guidelines. Subjects were asked to withhold bronchodilator medication for at least 6 hours before testing. For descriptive purposes, COPD subjects were staged according to GOLD guidelines (15).

### CT and Quantitative Analysis (Densitometry)

CT scanning was performed without bronchodilatation within 1 day of lung function testing. All CT scans were acquired using multidetector-row CT scanners (GE Healthcare, Milwaukee, WI or Siemens Healthcare, Erlangen, Germany) with a minimum of four rows at suspended full inspiration without administration of intravenous contrast. Exposure settings were 120 kVp and 40 mA; images were reconstructed using 1.0 mm (Siemens) or 1.25 mm (GE) contiguous slices and a low spatial frequency reconstruction algorithm (GE: standard; Siemens: b35f) at a  $512 \times 512$  matrix. All CT scans were analyzed using Pulmonary Workstation 2.0 software (VIDA Diagnostics, Iowa City, IA). Lungs were segmented from the thoracic wall, the heart, and main pulmonary vessels, followed by segmentation of the individual lobes and the airways (25). The extent of emphysema was estimated using the threshold technique quantifying the percent of voxels with an apparent x-ray attenuation value below  $-950$  HU (%LAA) (5). The size of the emphysematous lesions was estimated using a LAC analysis (17,18). In short, the cumulative size of the lesion (the number of connected low attenuation voxels) is plotted against the cumulative number of lesions (clusters of a given size) on a log-log scale. The slope of this relationship, the power law exponent (D), is an indication of average lesion size: The steeper the slope, the smaller the lesions.

### Qualitative Analysis (Visual Scoring)

All images were displayed on a standalone medical imaging workstation (eFilm 3.0, Merge Healthcare, Milwaukee, WI), enabling the readers to view the images in different planes and using a standard window width of 1000 and a window level of  $-700$  HU. The qualitative analysis was performed on images obtained using high spatial frequency reconstruction algorithm (GE: bone; Siemens: b60f). Emphysema was visually scored by two experienced chest radiologists, who independently scored all CT scans as follows: 0 = no emphysema, 1 =  $< 5\%$  (trivial), 2 = 5–25% (mild), 3 = 26–50% (moderate), 4 = 51–75% (severe), and 5 =  $> 75\%$  involvement of both lungs (very severe). In case of disagreement, the mean score of the two readers was used as the final score (0.5, 1.5, 2.5, 3.5, or 4.5) unless the results differed more than one category. For those cases consensus was reached in a separate reading session and that score was used. The distributions of emphysema, if present, was described as upper lobe predominant, lower lobe predominant or diffuse emphysema; and the predominant type of

**TABLE 1. Demographic Data by Study Group (A) and GOLD Stage (B)**

Table 1A

	COPD Subjects	Smoker Controls	Non-smoker Controls	P Values			
				Overall	COPD vs Smoker Controls	COPD vs. Nonsmoker Controls	Smoker Controls vs. Nonsmoker Controls
<i>n</i>	1519	269	184				
Age (y)	63.2 (6.9)	54.8 (8.9)	54.1 (9.1)	<.001	<.001	<.001	.329
Sex (M/F)	960/559	151/118	65/119	<.001	.028	<.001	<.001
Height	169.5 (8.9)	172.1 (9.1)	167.9 (8.8)	<.001	<.001	.027	<.001
Weight	74.7 (17.3)	78.8 (14.9)	76.9 (16.1)	<.001	<.001	.093	.238
Body mass index	25.9 (5.2)	26.5 (4.2)	27.2 (4.9)	.001	.055	.001	.175
Pack-years	48.0 (26.3)	31.5 (22.6)	0.0 (0.1)	<.001	<.001	<.001	<.001
Current smoker (%)	38%	63%	0	<.001	<.001	<.001	<.001
Never smoked (%)	0	0	96%				
FEV1% predicted	48.8 (15.8)	109.0 (11.5)	115.6 (13.6)	<.001	<.001	<.001	<.001
FEV1/FVC	44.4 (11.5)	79.1 (5.1)	81.4 (5.2)	<.001	<.001	<.001	.022
% LAA	18.1 (12.5)	2.3 (3.0)	4.2 (4.2)	<.001	<.001	<.001	.076

Table 1B

	GOLD II	GOLD III	GOLD IV	P Values			
				Overall	GOLD II vs GOLD III	GOLD II vs GOLD IV	GOLD III vs GOLD IV
<i>n</i>	692	635	190				
Age (y)	63.2 (7.1)	63.5 (6.8)	61.6 (6.9)	.005	.486	.005	.001
Sex (M/F)	404/288	423/212	132/58	.001	.002	.006	.461
Height	169.0 (9.1)	169.7 (8.7)	170.2 (8.9)	.202	.184	.119	.509
Weight	76.7 (17.3)	73.7 (16.6)	70.9 (18.5)	<.001	.002	<.001	.044
Body mass index	26.7 (5.1)	25.5 (5.1)	24.3 (5.3)	<.001	<.001	<.001	.004
Pack-years	47.0 (28.1)	49.1 (24.3)	48.5 (25.9)	.335	.146	.467	.807
Current smoker (%)	39%	38%	29%	.047	.825	.016	.025
FEV1% predicted	63.2 (8.4)	40.2 (5.8)	24.5 (3.7)	<.001	<.001	<.001	<.001
FEV1/FVC	52.3 (8.8)	39.6 (8.5)	31.5 (7.5)	<.001	<.001	<.001	<.001
% LAA	12.3 (9.7)	21.0 (11.8)	29.5 (12.6)	<.001	<.001	<.001	<.001

Results are given as mean  $\pm$  standard deviation.

FEV<sub>1</sub>, forced expiratory volume in 1 second; FVC, forced vital capacity; %LAA, percent low attenuation area.

**TABLE 2. Radiologist Graded Emphysema Extent Scores by Study Group (A) and GOLD Stage (B)**

Table 2A

Emphysema Extent	Emphysema Grade	COPD Subjects <i>n</i> (%)	Smoker Controls <i>n</i> (%)	Nonsmoker Controls <i>n</i> (%)
No emphysema	0	72 (4.7)	19 (7.1)	68 (37.0)
	0.5	3 (0.2)	0	2 (1.1)
<5%, trivial	1	283 (18.6)	227 (84.4)	110 (59.8)
	1.5	105 (6.9)	9 (3.3)	3 (1.6)
5–25%, mild	2	216 (14.2)	9 (3.3)	0
	2.5	100 (6.6)	1 (0.4)	1 (0.5)
26–50%, moderate	3	195 (12.8)	4 (1.5)	0
	3.5	109 (7.2)	0	0
51–75%, severe	4	184 (12.1)	0	0
	4.5	105 (6.9)	0	0
>75%, very severe	5	147 (9.7)	0	0

Table 2B

Emphysema Extent	Emphysema Grade	GOLD II <i>n</i> (%)	GOLD III <i>n</i> (%)	GOLD IV <i>n</i> (%)
No emphysema	0	52 (7.5)	18 (2.8)	2 (1.1)
	0.5	2 (0.3)	1 (0.2)	0
<5%, trivial	1	197 (28.5)	76 (12.0)	10 (5.3)
	1.5	59 (8.5)	41 (6.5)	4 (2.1)
5–25%, mild	2	126 (18.2)	74 (11.7)	15 (7.9)
	2.5	45 (6.5)	42 (6.6)	13 (6.8)
26–50%, moderate	3	88 (12.7)	86 (13.5)	21 (11.1)
	3.5	37 (5.3)	58 (9.1)	14 (7.4)
51–75%, severe	4	43 (6.2)	108 (17.0)	33 (17.4)
	4.5	20 (2.9)	55 (8.7)	30 (15.8)
>75%, very severe	5	23 (3.3)	76 (12.0)	48 (25.3)

Computed tomography scans were graded using a 6-point grading scale: 0 = no emphysema, 1 = <5% (trivial), 2 = 5–25% (mild), 3 = 26–50% (moderate), 4 = 51–75% (severe), and 5 = >75% involvement of both lungs (very severe). In case of disagreement, the mean score of the two readers was used as the final score (0.5, 1.5, 2.5, 3.5, or 4.5). Values are shown as the number of subjects in each grade and as the percent of total number of subjects in the group.

COPD, chronic obstructive pulmonary disease.

emphysema was described as paraseptal, centrilobular or panacinar (26). Regions of the lung exhibiting decreased attenuation and vascularity that are suggestive of “small airways disease” or gas-trapping were also noted and are referred to as gas-trapping throughout this report. CT scans were visually checked for other causes of airflow limitation, such as interstitial lung disease or severe bronchiectasis and, if present, the subjects were excluded from further analysis.

### Statistics

Data evaluation were performed using the SAS statistical software package version 9.1 (Cary, NC). *P* values < .05 were considered significant. Pearson’s correlation coefficients and Fisher’s *z* transformation were calculated to investigate the correlation between lung function and CT parameters. Kappa statistics were used to assess inter-observer agreement (27). Analyses of variance were conducted to compare LAC values. For COPD subjects only, regression analysis was carried out to determine the predictors of visually scored extent of emphysema from among the following variables: %LAA, LAC, distribution of emphysema, type of emphysema, and presence

of gas-trapping. Because the visual score was a categorical variable and the proportional odds assumption did not hold for all covariates, we used a partial proportional odds model to assess predictors of visual scores (28). Odds ratios and their associated 95% confidence intervals were calculated for each predictor. For comparison purposes, we have included the proportional odds model estimates.

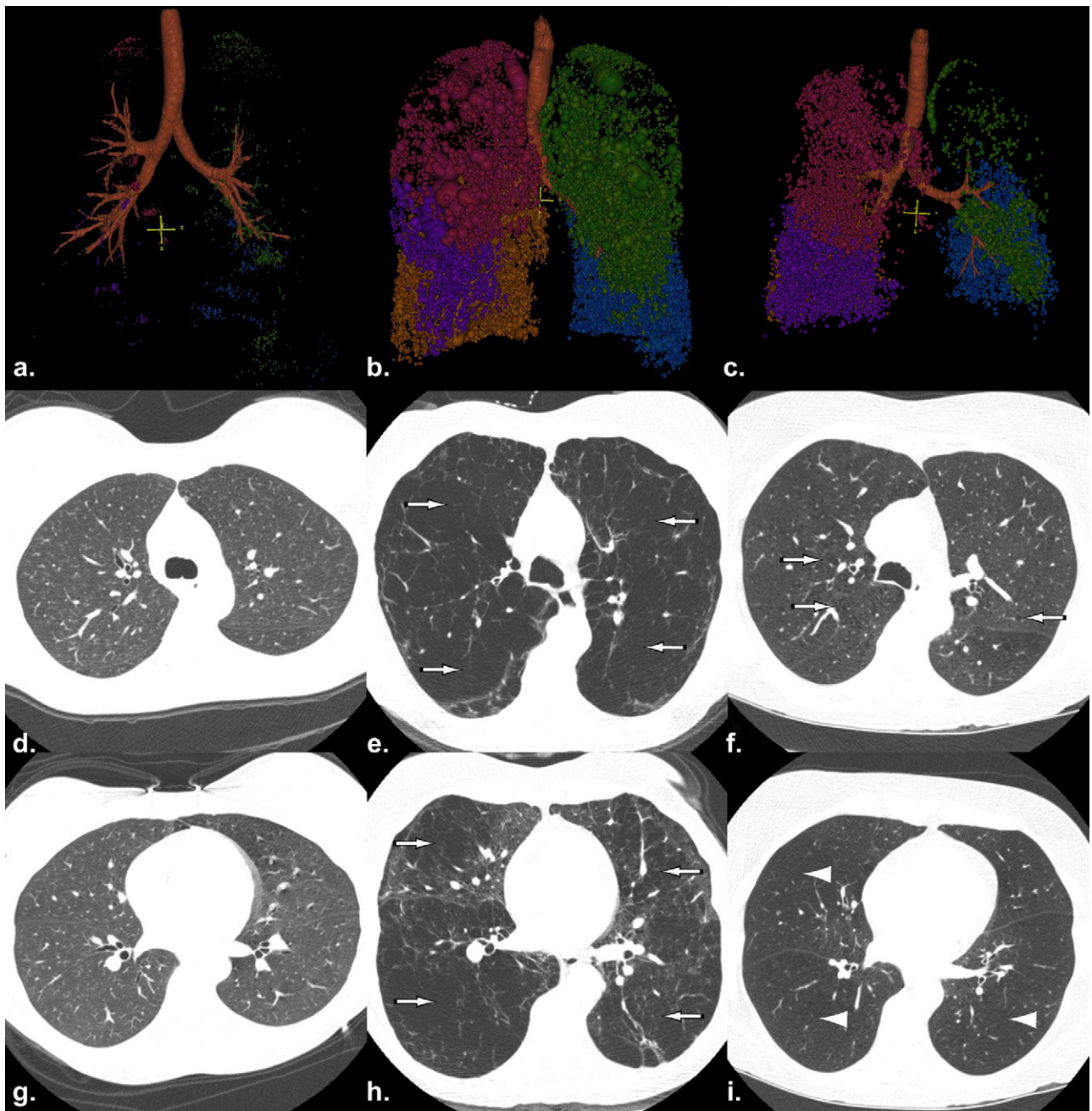
### Ethics Approval

The trial was approved by the ethics and review boards at participating centers and written informed consent was obtained from all subjects.

## RESULTS

### Demographics

The subject demographics are shown in Table 1. In total, 2290 subjects completed spirometry and had their CT scans analyzed by the software and read by both observers. A total of 233 CT scans were excluded from analysis because of the



**Figure 1.** This figure shows three individuals with the low attenuation cluster analysis of emphysema shown in (a,b,c). The colored regions represent connected regions of the lung voxels below  $-950$  HU. Each color represents a different lobe. Corresponding computed tomography (CT) images from the carina (d,e,f) and the inferior pulmonary vein (g, h, i) are also shown. Subject A is considered to have trivial emphysema by radiologists and %LAA assessment as indicated by the lack of colored regions in the three-dimensional reconstruction of the LAC analysis (a) and the lack of low attenuating regions in CT images. Subject B has severe emphysema indicated by the large colored regions (b) and the large low attenuating areas indicated by arrows ( $\leftarrow\rightleftarrows$ ) (e,h). Subject C has mild emphysema with small lesions indicated individual colored regions in the upper lobes (c) and marked by arrows ( $\leftarrow\rightleftarrows$ ) (f) and large diffuse areas of decreased attenuation suggesting “gas-trapping” from small airways disease (arrowhead  $\blacktriangleleft$ ) (i).

following reasons: respiratory motion artifacts ( $n = 189$ ) and parts of the lungs not included on the scan ( $n = 44$ ). An additional 62 were excluded because of severe bronchiectasis and 23 remaining subjects were excluded because of protocol violations, leaving 1972 CT scans available for analysis. The resulting study group consisted of 1519 COPD subjects,

269 smoker controls, and 184 nonsmoker controls. A total 439 scans (22%) had a visual score, which was one category different for the two observers. Twenty (1.0%) scans showed more than one category difference in the visual scores of the two observers. The readers showed excellent agreement (weighted  $\kappa = 0.83$  [95% confidence interval 0.81–0.84]).

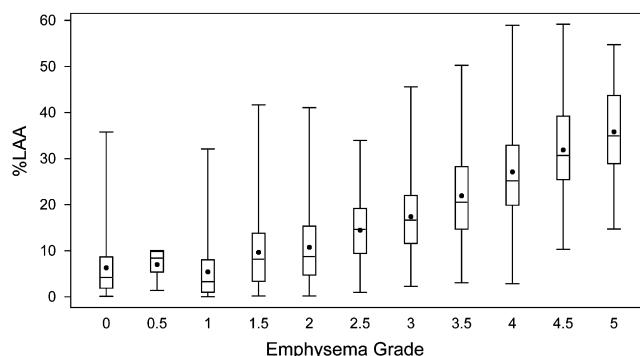
The radiologists scored 1813 (92%) CT scans as showing emphysema and 159 (8%) as showing no emphysema. The vast majority of smoker controls (91%) and nonsmoker controls (98%) were scored as having no or at most trivial emphysema. Table 2 shows the visual scores for each study group and GOLD stage.

### Quantitative versus Qualitative Analysis and Impact of Lesion Size

An example of the quantitative analysis and qualitative assessment of emphysema is shown in Figure 1. These data show that although there is an association between %LAA and visual scores ( $r = 0.77$ ,  $P < .001$ ; Fig 2) there is considerable variation in how the radiologists score emphysema compared with the quantitative approach. In the less severe categories (ie, less than a score of 2), radiologists tend to underestimate the extent of CT emphysema compared with the quantitative analysis, whereas in the more severe categories they tend to overestimate the extent of CT emphysema. Some hyperinflation was reported for 621 (31%) CT scans and of these, 327 (53%) CT scans were scored as showing no or trivial emphysema. Table 3 shows the frequency of presence of gas-trapping for each visual score and the mean %LAA for each study group. For CT scans without hyperinflation, the correlation coefficient for %LAA and visually estimated emphysema severity increased ( $P < .001$ ) to  $r = 0.81$  ( $P < .001$ ). For the total population, both %LAA and visual scores correlated well with FEV<sub>1</sub>%predicted ( $r = -0.63$  and  $r = -0.61$  respectively, both  $P < .001$ ). For COPD subjects alone, both associations were weaker ( $P < .001$ ): the associations between %LAA and FEV<sub>1</sub>%predicted ( $r = -0.52$ ,  $P < .001$ ), and between visual scores and FEV<sub>1</sub>% predicted ( $r = -0.44$ ,  $P < .001$ ) were moderate.

The LAC analysis was associated with %LAA ( $r = 0.73$ ,  $P < .001$ ), indicating increasing emphysematous lesion size with increasing extent of CT emphysema. The LAC was also associated with visual emphysema scores ( $r = 0.73$ ,  $P < .001$ ). The associations between LAC and emphysema were virtually the same when subjects were divided into those with and those without gas-trapping ( $r = 0.75$ ;  $P < .001$  and  $r = 0.76$ ;  $P < .001$ ). To determine if subjects with gas-trapping affect the correlation between the LAC and the qualitative and quantitative assessments of emphysema these subjects were removed and the analysis repeated. Significant increases in the correlations for visual emphysema scores (0.73–0.76;  $P = .022$ ), and for %LAA (0.73–0.75;  $P = .046$ ) were noted upon the removal of the subjects exhibiting signs of gas-trapping.

Table 4 demonstrates type of emphysema by study group and GOLD stage; Table 5 demonstrates distribution of emphysema. Subjects with panacinar emphysema had the largest lesions ( $D = -1.63 \pm 0.22$ ) compared with subjects with centrilobular emphysema ( $D = -1.99 \pm 0.47$ ;  $P < .001$ ) and paraseptal emphysema ( $D = -2.12 \pm 0.46$ ;  $P < .001$ ). Subjects with lower lobe predominance and diffuse emphysema had smaller lesions ( $D = -1.80 \pm 0.35$  and  $-1.82$



**Figure 2.** Box plot of radiologist-assessed emphysema grade vs. quantitative estimation (%LAA) of emphysema severity. Note that visual scores tend to overestimate extent of emphysema compared to %LAA. Computed tomography scans were graded using a six point grading scale: 0 = no emphysema, 1 = <5% (trivial), 2 = 5–25% (mild), 3 = 26–50% (moderate), 4 = 51–75% (severe), and 5 = >75% involvement of both lungs (very severe). In case of disagreement, the mean score of the two readers was used as the final score (0.5, 1.5, 2.5, 3.5, or 4.5).

$\pm 0.34$ ) compared with subjects with upper lobe predominance ( $D = -2.05 \pm 0.48$ ;  $P < .001$ ). After adjusting for visual scores, CT scans scored as showing regions suggesting gas-trapping had on average 3.6% ( $\pm 0.4\%$ ) greater %LAA than those without gas-trapping. CT scans scored as showing some gas-trapping showed a higher frequency of lower lobe predominant emphysema (39% vs 27%;  $P < .001$ ) but no difference in the average LAC (difference in slopes =  $-0.039 \pm 0.02$ ;  $P = .101$ ) was seen. There was also no difference in hole size distribution, upper lobe vs lower lobe, between those with and without gas-trapping (difference in slopes =  $0.035 \pm 0.02$ ;  $P = .090$ ).

### Factors Predicting Radiologists' Results

In the COPD subjects, regression analysis that included predictors %LAA, distribution of emphysema, presence/absence of gas-trapping, LAC (D) (surrogate for lesion size as seen by radiologists), and type of emphysema yielded a model with all parameters being predictors of visual scores (all  $P < .001$ ). The odds ratios for each predictor in a partial proportional odds model are shown in Table 6A, whereas the odds ratios for a proportional odds model are presented in Table 6B. For each percentage point increase in %LAA, the odds of a higher visual score versus a lower score are 1.11 (95% CI: 1.09, 1.12) times. Lack of evidence of gas-trapping was associated with increased odds of a higher visual score (OR 5.02; 95% CI: 3.95, 6.38), whereas noncentrilobular emphysema was associated with decreased odds of a higher visual score (OR 0.24; 95% CI: 0.17, 0.34). The odds of a visual score 5 are 1.81 (95% CI: 1.24, 2.64) greater for subjects with diffuse rather than upper lobe emphysema, but the odds of a score of 4 or greater are 2.26 (95% CI: 1.59, 3.21) for subjects with diffuse rather than upper lobe emphysema.

**TABLE 3. Extent of Low-attenuation Areas for Each Emphysema Score**

Emphysema Extent	Emphysema Grade	%LAA (–950 HU)	5–95% Confidence Intervals	Decreased Attenuation <i>n</i> (%)
No emphysema	0	6.3 (6.5)	0.4–20.9	92 (57.9)
	0.5	7.0 (3.7)	1.4–10.1	4 (80.0)
<5%, trivial	1	5.4 (5.9)	0.2–18.2	231 (37.3)
	1.5	9.7 (8.1)	0.5–25.8	61 (52.1)
5–25%, mild	2	10.8 (7.6)	1.5–25.8	99 (44.0)
	2.5	14.5 (7.4)	2.5–27.3	38 (37.3)
26–50%, moderate	3	17.4 (7.7)	5.0–30.6	47 (23.6)
	3.5	21.9 (9.7)	7.9–39.3	22 (20.2)
51–75%, severe	4	27.1 (10.4)	12.6–45.9	15 (8.2)
	4.5	31.9 (9.4)	15.3–46.8	11 (10.5)
>75%, very severe	5	35.8 (9.4)	20.3–50.7	1 (0.7)

The last column represents the frequency of computed tomography scans scored as showing areas of decreased attenuation consistent with “gas-trapping” from small airways disease for each emphysema score. Results are shown as means and standard deviation (SD) except the decreased attenuation column which is the number of subjects and the percent of the group. The same grading system is used as described in Table 2.

HU, Hounsfield units; %LAA, percent low attenuation area.

**TABLE 4. Frequencies of Types of Emphysema by Study Group (A) and GOLD Stage (B)**

Table 4A

Emphysema Type	COPD Subjects <i>n</i> (%)	Smoker Controls <i>n</i> (%)	Nonsmoker Controls <i>n</i> (%)
Centrilobular	1252 (82.4)	177 (65.8)	100 (54.3)
Paraseptal	151 (9.9)	73 (27.1)	16 (8.7)
Panacinar	44 (2.9)	0	0
Not applicable	72 (4.7)	19 (7.1)	68 (37.0)

Table 4B

Emphysema Type	GOLD II <i>n</i> (%)	GOLD III <i>n</i> (%)	GOLD IV <i>n</i> (%)
Centrilobular	530 (76.6)	549 (86.5)	171 (90.0)
Paraseptal	103 (14.9)	43 (6.8)	5 (2.6)
Panacinar	7 (1.0)	25 (3.9)	12 (6.3)
Not applicable	52 (7.5)	18 (2.8)	2 (1.1)

Results are shown as the number of subjects and the percent (%) of the total number of subjects in each group. For those who did not show emphysema, type is “not applicable.” If emphysema was present, but the radiologists were not sure about the type of emphysema, type was scored as “unknown.”

## DISCUSSION

In this radiologic study of a large cohort of subjects we demonstrate there is a good association between quantitative and qualitative assessments of emphysema. We also report that visual scores are not only determined by the extent of low attenuation voxels, used in this analysis as surrogate for the extent of “dark areas” seen by radiologists, but also by how these low attenuation areas are spatially arranged in the lung. The data in this study show that a quantitative analysis of how these LACs have the highest proportional odds ratio to predict an increase in the visual score assigned by the radiologist (Table 6A, B). Furthermore, the emphysema type (centrilobular or noncentrilobular), distribution (upper, lower, or diffuse) and the radiologist’s judgment as to whether the decreased attenuation is caused by “emphysematous destruction” or what is considered to be “small airways

disease” (Table 6, Fig 1) also predicts the severity score assigned by the radiologist. These results indicate that estimating the extent of emphysema is a far more complicated task than just detecting low attenuation areas.

These data also show that for moderate to severe emphysema visual estimations tend to overestimate the extent of emphysema compared with a quantitative assessment of low attenuation areas, whereas for smaller amounts of CT emphysema (ie, a score <2) radiologists tend to underestimate the extent of CT emphysema. A possible explanation for these results could be that small isolated low attenuating voxels are easily missed by a reader, or attributed to some other cause such as image noise. Furthermore, it is likely that radiologists interpret certain low attenuation areas as from gas trapping or small airway disease, whereas the computer can only detect that the voxel has a low attenuation value. Although our data are not conclusive because the radiologists did not try

**TABLE 5. Distribution of Emphysema as Scored by the Radiologists by Study Group (A) and GOLD Stage (B)**

Table 5A

Emphysema Distribution	COPD Subjects <i>n</i> (%)	Smoker Controls <i>n</i> (%)	Nonsmoker Controls <i>n</i> (%)
Upper lobe	1083 (71.3)	232 (86.2)	110 (59.8)
Lower lobe	83 (5.5)	5 (1.9)	4 (2.2)
Diffuse	281 (18.5)	13 (4.8)	2 (1.1)
Not applicable	72 (4.7)	19 (7.1)	68 (37.0)

Table 5B

Emphysema Distribution	GOLD II <i>n</i> (%)	GOLD III <i>n</i> (%)	GOLD IV <i>n</i> (%)
Upper lobe	521 (75.3)	432 (68.0)	128 (67.4)
Lower lobe	25 (3.6)	46 (7.2)	12 (6.3)
Diffuse	94 (13.6)	139 (21.9)	48 (25.3)
Not applicable	52 (7.5)	18 (2.8)	2 (1.1)

Results are shown as the number of subjects and the percent of the total number of subjects in each group. For those who did not show emphysema, distribution is "not applicable." If emphysema was present, but the radiologists were not sure about the distribution of emphysema, distribution was scored as "unknown."

**TABLE 6. Odds Ratios (95% CI) for Each Cumulative Logit in Partial Proportional Odds Model (A) and Odds Ratios (95% CI) for Proportional Odds Model (B)**

Table 6A

	Score = 5	Score $\geq 4$	Score $\geq 3$	Score $\geq 2$
%LAA	1.11 (1.09–1.12)	1.11 (1.09–1.12)	1.11 (1.09–1.12)	1.11 (1.09–1.12)
Trapping: no vs. yes	5.02 (3.95–6.38)	5.02 (3.95–6.38)	5.02 (3.95–6.38)	5.02 (3.95–6.38)
Diffuse vs. upper	1.81 (1.24–2.64)	2.26 (1.59–3.21)	3.58 (2.37–5.42)	4.38 (2.42–7.96)
Lower vs. upper	2.35 (1.01–5.51)	6.25 (3.22–12.15)	4.45 (2.32–8.56)	3.46 (1.33–9.00)
Low attenuation cluster analysis	4.38 (1.86–10.33)	14.96 (7.54–29.69)	17.43 (9.95–30.52)	24.88 (13.55–45.67)
Centrilobular: no vs. yes	0.24 (0.16–0.34)	0.24 (0.16–0.34)	0.24 (0.16–0.34)	0.24 (0.16–0.34)

Table 6B

	Proportional Odds Model
%LAA	1.10 (1.08–1.11)
Trapping: no vs. yes	4.73 (3.74–5.99)
Diffuse vs. upper	2.68 (2.06–3.48)
Lower vs. upper	4.15 (2.58–6.67)
Low attenuation cluster analysis	18.91 (12.13–29.48)
Centrilobular: No vs. yes	0.24 (0.17–0.34)

%LAA, percent low attenuation area.

to quantify the extent of gas trapping, this suggestion is supported by the observation that half of the CT scans scored as showing some gas trapping were also scored as showing no or trivial emphysema.

It should also be noted that many of the "nonsmoking control" subjects showed radiologist scored emphysema. These findings may be at odds with what a reader would expect to see in nonsmoking subjects but it should be noted that there could be many reasons for these findings. First, the CT scans used in this analysis were "low-dose" CT scans and because the radiologist knew that this was an emphysema study may have over called the presence of emphysema if any low attenuating voxels were seen. Second, these subjects were in their sixth decade (and older) and it should be expected that

some age related changes would be observed. Third, intraparenchymal "cysts" have been seen in normal nonsmoking older subjects (29) and may have been interpreted as emphysema. Furthermore, as discussed above, focal areas of gas-trapping may have influenced the readers' interpretation of the low attenuation region.

Therefore, we think that a LAC has the potential to discriminate emphysema from decreased attenuation due to small airways disease (gas-trapping) and image noise. It is assumed that if a subject has small airways disease these regions will show gas-trapping on expiration and many of these regions will have voxels with low attenuation values that are not ordered in a specific pattern in inspiration. The data from this study show that the average quantitative lesion size

(D) is smaller in those subjects with regions that are judged by the radiologist to exhibit decreased attenuation and vascularity without obvious tissue destruction (26) (ie, no visible “hole,” suggestive of small airways disease). We report that CT scans with areas of gas-trapping have on average 3.6% (of total lung volume) more low attenuation areas than those without, for a similar extent of emphysema as estimated by radiologists. These results indicate that the computer ‘sees’ both entities, but cannot separate them. It also suggests that the decreased attenuation areas detected in this study were typically small. Because emphysema and areas of gas-trapping often occur together, analysis of total lungs using this technique is a poor discriminator between these entities.

Studies investigating the relation between subjective and objective (ie, software based) estimations of emphysema severity are limited. Bankier et al showed moderate to good associations ( $r = 0.54$  to  $0.72$ ) of visual grading with densitometric assessments but also showed that human observers tend to overestimate the extent of emphysema compared with histology specimens, whereas %LAA demonstrated no significant differences compared with histology (12). We find a better association ( $r = 0.77$ ) between qualitative and quantitative assessments of CT emphysema than Bankier et al reported, but the reported association is slightly weaker than those reported by Gurney et al and Park et al (both  $r = 0.82$ ) (13,14). All previous studies reported that visual scores tend to overestimate emphysema severity compared with %LAA, in accordance with the current report. However, in the current analysis this overestimation was only true for moderate to severe emphysema, whereas in subjects with trivial to mild emphysema an opposite trend was observed. To the best of our knowledge, no previous studies have investigated the factors influencing subjective scores other than the extent of low attenuation areas.

This study has some limitations. First, as in most CT studies of emphysema, there is no gold standard to compare the emphysema metrics against. However, there is a significant body of literature that supports the use of quantitative CT scanning to measure emphysema (3–5,30) and although these data are not irrefutable; it is highly supportive of the use of threshold analysis to assess the extent of emphysema. Second, we used  $-950$  HU as threshold to quantify the extent of emphysema. Although Madani et al reported that  $-960$  HU may be a better cutoff value for CT scans performed on a multidetector-row CT scanner (7), the threshold of  $-950$  HU is still commonly used (20,24,31,32) and any cutoff is arbitrary.

Third, gas-trapping was not strictly assessed in this study, but included as another feature of the CT scan. By study design, radiologists assessed the extent and type and distribution of emphysema as primary metrics but also included some notes as to the presence or absence of other abnormalities such as gas-trapping, which was scored as a simple “yes” or “no” statement, and no attempt was undertaken to assess the extent of these low attenuation areas or to make comments on the regional distribution of these areas. Therefore, subjects anno-

tated as having gas-trapping may in fact have been areas with early emphysema since emphysematous regions often show air trapping. Furthermore, expiratory CT scans were not performed within ECLIPSE, so the assumption that these low attenuation areas actually do correspond to regions of gas-trapping could not be confirmed. Therefore, the analysis of LAC analysis for discriminating emphysema from small airways disease is limited but suggestive and requires further analysis using proper CT protocols.

In conclusion, several factors influence severity of emphysema as estimated by radiologists other than low attenuation areas, including lesion size, type and distribution of emphysema. Most of these parameters can be provided by computer software, and therefore more sophisticated models are required to estimate emphysema severity than %LAA alone. Such models may be able to replace or augment radiologists in estimating emphysema severity. Our data suggest that lesion size assessed using the LAC analysis is one of these computed parameters and may be helpful in distinguishing emphysema from regions of low attenuation and image noise.

## ACKNOWLEDGMENTS

The authors thank all the participants for their willingness to contribute to advance medical science in the field of COPD. The authors would like to thank Tara Candido, Sebastian Cogswell, Heather Davis, Nima Farzaneh, Lukas Holy, Natasha Krowchuk, Helena Lee, Evan Phillips, Nerissa Tai, Anh-Toan Tran, Nghia Tran, Eugene Wang, and Tomonori Yokogawa for technical assistance with the CT analysis and data management as well as Juerg Tschirren and Greg Gallardo at VIDA Diagnostics for assistance with the Pulmonary Workstation software.

The authors would also like to acknowledge the ECLIPSE steering and scientific committees. The ECLIPSE study is funded by GlaxoSmithKline (ClinicalTrials.gov identifier: NCT00292552; GlaxoSmithKline study code SCO104960).

ECLIPSE Steering Committee: Harvey Coxson (Canada), Lisa Edwards (GlaxoSmithKline, USA), Katharine Knobil (cochair, GlaxoSmithKline, UK), David Lomas (UK), William MacNee (UK), Edwin Silverman (USA), Ruth Tal-Singer (GlaxoSmithKline, USA), Jørgen Vestbo (Co-chair, Denmark), Julie Yates (GlaxoSmithKline, USA).

ECLIPSE Scientific Committee: Alvar Agusti (Spain), Peter Calverley (UK), Bartolome Celli (USA), Courtney Crim (GlaxoSmithKline, USA), Bruce Miller (GlaxoSmithKline, UK), William MacNee (Chair, UK), Stephen Rennard (USA), Ruth Tal-Singer (GlaxoSmithKline, USA), Emiel Wouters (The Netherlands), Julie Yates (GlaxoSmithKline, USA).

## REFERENCES

- Gould GA, MacNee W, McLean A, et al. CT measurements of lung density in life can quantitate distal airspace enlargement—an essential defining feature of human emphysema. *Am Rev Respir Dis* 1988; 137:380–392.
- Hayhurst MD, Flenley DC, McLean A, et al. Diagnosis of pulmonary emphysema by computerized tomography. *Lancet* 1984; 2:320–322.

3. Müller NL, Staples CA, Miller RR, et al. "Density mask". An objective method to quantitate emphysema using computed tomography. *Chest* 1988; 94:782-787.
4. Coxson HO, Rogers RM, Whittall KP, et al. A quantification of the lung surface area in emphysema using computed tomography. *Am J Respir Crit Care Med* 1999; 159:851-856.
5. Gevenois PA, de Maertelaer V, De Vuyst P, et al. Comparison of computed density and macroscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med* 1995; 152:653-657.
6. Gevenois PA, De Vuyst P, de Maertelaer V, et al. Comparison of computed density and microscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med* 1996; 154:187-192.
7. Madani A, De Maertelaer V, Zanen J, et al. Pulmonary emphysema: radiation dose and section thickness at multidetector CT quantification—comparison with macroscopic and microscopic morphometry. *Radiology* 2007; 243:250-257.
8. Shaker SB, Dirksen A, Laursen LC, et al. Short-term reproducibility of computed tomography-based lung density measurements in alpha-1 antitrypsin deficiency and smokers with emphysema. *Acta Radiol* 2004; 45:424-430.
9. Bergin C, Muller N, Nichols DM, et al. The diagnosis of emphysema. A computed tomographic-pathologic correlation. *Am Rev Respir Dis* 1986; 133:541-546.
10. Hruban RH, Meziane MA, Zerhouni EA, et al. High resolution computed tomography of inflation-fixed lungs: pathologic-radiologic correlation of centrilobular emphysema. *Am Rev Respir Dis* 1987; 136:935-940.
11. Kuwano K, Matsuba K, Ikeda T, et al. The diagnosis of mild emphysema. Correlation of computed tomography and pathology scores. *Am Rev Respir Dis* 1997; 141:169-178.
12. Bankier AA, De Maertelaer V, Keyzer C, et al. Pulmonary emphysema: subjective visual grading versus objective quantification with macroscopic morphometry and thin-section CT densitometry. *Radiology* 1999; 211:851-858.
13. Gurney JW, Jones KK, Robbins RA, et al. Regional distribution of emphysema: correlation of high-resolution CT with pulmonary function tests in unselected smokers. *Radiology* 1992; 183:457-463.
14. Park KJ, Bergin CJ, Clausen JL. Quantitation of emphysema with three-dimensional CT densitometry: comparison with two-dimensional analysis, visual emphysema scores, and pulmonary function test results. *Radiology* 1999; 211:541-547.
15. Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007; 176:532-555.
16. Uppaluri R, Mitsa T, Sonka M, et al. Quantification of pulmonary emphysema from lung computed tomography images. *Am J Respir Crit Care Med* 1997; 156:248-254.
17. Coxson HO, Whittall KP, Nakano Y, et al. Selection of patients for lung volume reduction surgery using a power law analysis of the computed tomographic scan. *Thorax* 2003; 58:510-514.
18. Mishima M, Hirai T, Itoh H, et al. Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease. *Proc Natl Acad Science U S A* 1999; 96:8829-8834.
19. Madani A, Van Muylem A, de Maertelaer V, et al. Pulmonary emphysema: size distribution of emphysematous spaces on multidetector CT images—comparison with macroscopic and microscopic morphometry. *Radiology* 2008; 248:1036-1041.
20. Martinez FJ, Foster G, Curtis JL, et al. Predictors of mortality in patients with emphysema and severe airflow obstruction. *Am J Respir Crit Care Med* 2006; 173:1326-1334.
21. Agusti A, Calverley PM, Celli B, et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir Res* 2010; 11:122.
22. Lomas DA, Silverman EK, Edwards LD, et al. Serum surfactant protein D is steroid sensitive and associated with exacerbations of COPD. *Eur Respir J* 2009; 24:95-102.
23. Vestbo J, Anderson W, Coxson HO, et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE). *Eur Respir J* 2008; 31:869-873.
24. Lomas DA, Silverman EK, Edwards LD, et al. Evaluation of serum CC-16 as a biomarker for COPD in the ECLIPSE cohort. *Thorax* 2008; 63:1058-1063.
25. Zhang L, Hoffman EA, Reinhardt JM. Atlas-driven lung lobe segmentation in volumetric X-ray CT images. *IEEE Trans Med Imaging* 2006; 25:1-16.
26. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008; 246:697-722.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-174.
28. Stols ME, Davis CS, Koch GG. Categorical data analysis using the SAS system. Cary, NC: SAS Institute Inc, 2000.
29. Copley SJ, Wells AU, Hawtin KE, et al. Lung morphology in the elderly: comparative CT study of subjects over 75 years old versus those under 55 years old. *Radiology* 2009; 251:566-573.
30. Madani A, Zanen J, de Maertelaer V, et al. Pulmonary emphysema: objective quantification at multi-detector row CT - comparison with macroscopic and microscopic morphometry. *Radiology* 2006; 238:1036-1043.
31. Dransfield MT, Washko GR, Foreman MG, et al. Gender differences in the severity of CT emphysema in COPD. *Chest* 2007; 132:464-470.
32. Patel BD, Coxson HO, Pillai SG, et al. Airway wall thickening and emphysema show independent familial aggregation in COPD. *Am J Respir Crit Care Med* 2008; 178:500-505.

**APPENDIX 1. PRINCIPAL INVESTIGATORS AND CENTERS PARTICIPATING IN ECLIPSE (NCT00292552, SCO104960)**

Bulgaria: Yavor Ivanov, Pleven; Kosta Kostov, Sofia.

Canada: Jean Bourbeau, Montreal, Que Mark Fitzgerald, Vancouver, BC; Paul Hernandez, Halifax, NS; Kieran Killian, Hamilton, Ontario; Robert Levy, Vancouver, BC; Francois Maltais, Quebec City, Que; Denis O'Donnell, Kingston, Ontario.

Czech Republic: Jan Krepelka, Praha.

Denmark: JØrgen Vestbo, Hvidovre.

Netherlands: Emiel Wouters, Maastricht.

New Zealand: Dean Quinn, Wellington.

Norway: Per Bakke, Bergen.

Slovenia: Mitja Kosnik, Golnik.

Spain: Alvar Agusti, Palma de Mallorca.

Ukraine: Yuri Feschenko, Kiev; Vladimir Gavrisyuk, Kiev; Lyudmila Yashina, Kiev; Nadezhda Monogarova, Donetsk.

United Kingdom: Peter Calverley, Liverpool; David Lomas, Cambridge; William MacNee, Edinburgh; David Singh, Manchester; Jadwiga Wedzicha, London.

United States of America: Antonio Anzueto, San Antonio, TX; Sidney Braman, Providence, RI; Richard Casaburi, Torrance CA; Bart Celli, Boston, MA; Glenn Giessel, Richmond, VA; Mark Gotfried, Phoenix, AZ; Gary Greenwald, Rancho Mirage, CA; Nicola Hanaia, Houston, TX; Don Mahler, Lebanon, NH; Barry Make, Denver, CO; Stephen Rennard, Omaha, NE; Carolyn Rochester, New Haven, CT; Paul Scanlon, Rochester, MN; Dan Schuller, Omaha, NE; Frank Sciruba, Pittsburgh, PA; Amir Sharafkhaneh, Houston, TX; Thomas Siler, St. Charles, MO, Edwin Silverman, Boston, MA; Adam Wanner, Miami, FL; Robert Wise, Baltimore, MD; Richard ZuWallack, Hartford, CT.