LUNG FUNCTION AND EMPHYSEMA IN A LARGE LUNG CANCER CASE SERIES

by

Maria Cecilia Crisanti

Medical Doctor, University of Buenos Aires, Argentina, 2002

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Maria Cecilia Crisanti

It was defended on

March 31st, 2011

and approved by

Thesis Advisor: Joel Weissfeld M.D. M.P.H., Associate Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

John Wilson PhD, Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Brenda Diergaarde PhD, Assistant Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

LUNG FUNCTION AND EMPHYSEMA IN A LARGE LUNG CANCER CASE SERIES

Maria Cecilia Crisanti, M.S.

University of Pittsburgh, 2011

Chronic obstructive pulmonary disease (COPD) is the most frequent chronic disease in developed countries and is predicted to be the third cause of death in 2020. Lung cancer is the leading cause of cancer death both in men and women. A vast majority of patients diagnosed with lung cancer have COPD, a history of tobacco use, or both. Shared inflammatory pathways may govern the pathogenesis of COPD and lung cancer. Several studies imply a relationship between COPD and lung cancer, but there is very limited information in the literature about emphysema and lung cancer risk. The Carinal Registry is a prospectively collected case series of patients diagnosed with lung cancer at the University of Pittsburgh. Among other data, it contains information on COPD that was acquired from the medical record (yes/no entry). Our goal was to evaluate the data quality and decide if the COPD variable could be used to score for emphysema. For this purpose, we adapted a subjective, semi-quantitative, visual emphysema scoring method (VESM) to score emphysema severity on C T scans and compared this emphysema severity score with the presence or absence of emphysema as defined in the Carinal Registry. We defined the best CT to score emphysema to be obtained preferably within one year preoperatively, with lung edge-enhancing reconstruction algorithm. Training in VESM showed high reproducibility scores and high sensitivity of the trainee to detect emphysema when compared to the standard expert score. Our results showed that there was poor correlation between the COPD status as recorded from the COPD variable from medical records and the

VESM. The VESM was a more accurate measure of COPD status among lung cancer patients enrolled in the Carinal Registry. Moreover, we compared the distribution of COPD among lung cancer patients and the community adapted from Wilson et al. manuscript who carried out a community based screening study for lung cancer among smokers. Our results showed that the distribution of COPD was similar among both populations, suggesting the possibility of an underlying common pathway of lung cancer and emphysema.

Public Health Significance:

The public health significance of this study is clearly explained by the high frequency of both emphysema and lung cancer and the dismal prognosis of lung cancer. We have studied a sample of a large case series of lung cancer patients and scored their emphysema severity with a semi-quantitative method based on CT scan reading. We have also compared this method with simply retrieving emphysema data from the medical record and assessed the validity of these methods. All the above mentioned are very important reasons that can affect the public health as well as research purposes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Acknowledgements:

Special thanks to Pam Sufka for providing research support, Dr. Steven Fisher and Dr. Carl Fuhrman for participating in radiology training and as expert reader of CT scans, Stentor help desk for invaluable assistance.

References are listed as by the American Journal of Epidemiology guidelines for authors.

1.0 RATIONALE AND BACKGROUND

Chronic obstructive pulmonary disease (COPD) is the most frequent chronic disease in developed countries and is predicted to be the third cause of death in 2020 (1). According to the Third National Health and Nutrition Examination Survey (NHANES III), COPD affects approximately 10% of the United States adult population. COPD is a multicomponent disease that can be evaluated and characterized by spirometry to assess functional capacity and imaging to assess distribution and extent of the disease (2).

Lung cancer has a worldwide incidence of 1.2 million cases. In the United States the estimated annual death rate due to this disease for 2010 (157,300) approximates its annual incidence rate (222,520), making it the leading cause of cancer deaths in both men and women in the United States (3).

A vast majority of patients diagnosed with lung cancer have COPD, a history of tobacco use, or both (40-70%). However, only a minority of long term smokers (10-15%) develops COPD (4). The pathophysiology of COPD and emphysema is characterized by luminal airway narrowing and destruction of lung parenchyma driven by inflammation. Thus, shared inflammatory pathways may govern the pathogenesis of COPD and lung cancer (5,6). The role of chronic airway inflammation induced by cigarette smoke is an active area of research. It remains unclear whether COPD is in the causal pathway of lung cancer or whether both COPD and lung cancer are related to an underlying exposure, or a combination of both. However, it has

also been shown that even in non-smokers, the presence of COPD alone increases the risk of developing lung cancer (7,8).

Several studies imply a relationship between chronic obstructive pulmonary disease and lung cancer, but there is very limited information in the literature about specifically emphysema and lung cancer risk. A nested case control study by Maldonado et al. (9) from the Mayo Clinic, in which patients were pulled from a screening study and scored for emphysema with software in an objective manner, revealed no association between radiographic emphysema and the risk of lung cancer. However, Wilson et al (10) studied subjects in the frame of the Pittsburgh Lung Screening Study (PLuSS), which is a subproject of the University of Pittsburgh Lung Cancer Specialized Program of Research Excellence (SPORE). The aforementioned is a community-based study of lung cancer screening that used low dose multi-detector helical CT (LDCT) and pulmonary function tests (PFT) as part of its assessment of emphysema. The authors reported an increased frequency of lung cancer in subjects with emphysema, with the highest frequency observed in subjects with both emphysema and moderate–severe airflow obstruction. Their main finding was that for any level of tobacco exposure, patients with chronic airflow obstruction were at greater risk for lung cancer than smokers without airflow obstruction. This relationship proved to be severity dependent; where individuals with the worst lung function showed the highest risk. De Torres et al. studied a prospective cohort of individuals enrolled in a lung cancer screening study that also used LDCT and found that the incidence density of lung cancer among individuals with emphysema on LDCT was 25.0 per 1,000 person-years and only 7.5 per 1,000 person-years among individuals without emphysema (11). Littman et al. as well reported a positive association between radiographic emphysema and the risk of lung cancer (12).

The Carinal Registry is a prospectively collected case series of patients diagnosed with lung cancer that started in 1990 at the University of Pittsburgh. These patients underwent surgery and/or had blood/tissue collected. The Carinal Registry possesses deficient emphysema information. The current problems that researchers face are based upon one fact: the data were not collected for research purposes, but in the clinical environment. These data were not collected using a standard protocol. For example, COPD data in the Carinal Registry comes from medical records and not from direct CT scan study reading. In addition, it is a retrospective analysis. Researchers also face data quality problems that can be partially due to variability in equipment, such as advances in electronic medical records, CT image acquisition and CT image reconstruction protocol changes throughout the years, CT technology improvement between 2002 and 2009, technical problems such as motion artifact, and other unexpected issues such as interim health problems obscuring the image (pleural effusion, pneumonia, post-operative changes).

The objectives of this study are three fold as described by the Specific Aims.

1.1 SPECIFIC AIM 1

Data analysis: To perform a systematic evaluation of research procedures designed to retrieve structural emphysema data for participants in a lung cancer database.

1.1.1 Systematic Database Evaluation

Using 2002-2009 Carinal Registry patients as our study base, we described the proportion of different gender, ethnic groups, age group, and histological types of lung cancer. The completeness of data entered from the medical history as emphysema/COPD, asthma, bronchitis and PFT results was evaluated. The frequency of smokers, both current and former and the distribution of pack years among them were analyzed. We looked at the family history and specifically described how many patients have a positive family history of lung cancer, other tobacco related cancers, or non-tobacco related cancer. All this information was evaluated in the Pilot Study (random sample from Carinal Registry, resource driven) as well.

1.1.2 Systematic Analysis of CT information

We analyzed the availability of CT studies. CT scans were divided in pre-diagnosis, post-diagnosis, pre-operative and post-operative. Specific reconstruction parameters (edge-enhancing lung algorithm) mA dose and slice thickness used in these studies were also evaluated.

1.2 SPECIFIC AIM 2

Emphysema and Lung Cancer: To describe the distribution of emphysema among lung cancer patients.

We described the distribution and frequency of COPD and emphysema among lung cancer patients. All patients enrolled in the Carinal Registry have lung cancer and most

underwent CT scan study. CT scans were retrieved and scored for emphysema based on the National emphysema Treatment Trial (NETT) protocol (see Methods) (13). These results were compared to the emphysema scores reported by Wilson et al. (10).

1.3 SPECIFIC AIM 3

Validation of Data in a Lung Cancer Database

1.3.1 To validate data in the Carinal Registry database

Data validation was carried out by comparison of the frequencies and distribution of the COPD variable from the Pilot Study (information obtained from the medical records) with the Visual Emphysema Score Method (VESM) after reading the CT scans for emphysema according to the NETT standards.

1.3.2 To assess the reproducibility of the visual emphysema scoring method

Intra-rater variability of VESM was assessed as a measure of reliability. CT scans were independently read and results were compared using a weighted kappa statistic. Inter-rater variability of VESM was assessed as well as a measure of reliability, by comparing CT scan scoring from the trainee vs. the expert scoring of the three training modules and the Pilot Study CT scans.

2.0 METHODS

2.1 DATABASE AND PATIENT POPULATION

The database used for this study is the Carinal Registry version 12, downloaded July 2nd 2010. It is a prospectively collected database of lung cancer patients who had tissue and/or blood sample collected. It contains demographic information, specimen information (sample received date, sample type, pathology information, lung cancer staging), social history, family history, medical history, and Pulmonary Function Test (PFT) results. The sub-sample used in this study is a stratified random sample of 64 patients from an original sample of 548 patients diagnosed with lung cancer that had blood and tissue sample collected between 2002 and 2009. S pecific inclusion criteria as shown in Figure 1 are: (1) unique SSN Medical Record Number; (2) any lung cancer; (3) first lung cancer sample received between years 2002-2009; (4) lung cancer histology non-missing and non-carcinoid; (5) sample received within 12 months of diagnosis; (6) patient not enrolled in PLuSS; (7) patient had blood collected; (6) age at diagnosis $\geq$50; (7) positive smoking history; (8) 10+ pack-years.

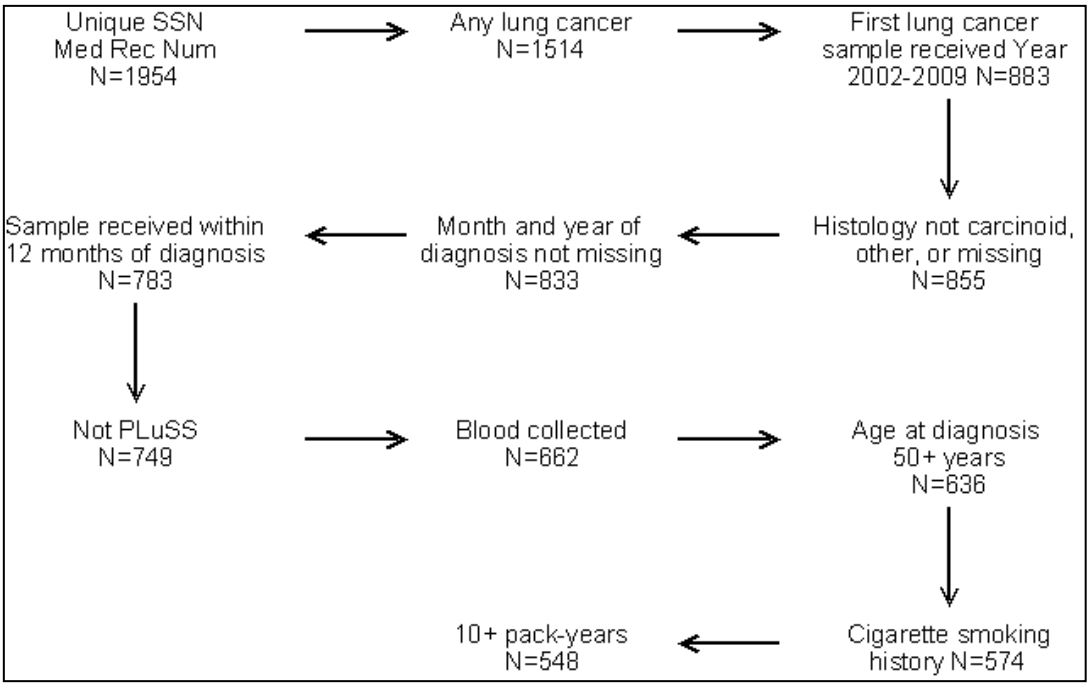


Figure 1. Sample Selection Strategy.

Algorithm followed to select Carinal Registry Sample Set of 548 patients.

2.2 SAMPLE SELECTION STRATEGY

A Pilot Sample was selected according to two stratifying variables. These were year of enrollment in database (8 years, from 2002 to 2009) and whether or not the patient had a Pulmonary Function Test (PFT) performed with available result present in the database; giving a total of 16 strata. This pilot sample was randomly selected using an equal allocation technique with a selection of a fixed number of sampling units (n=4) from each of the stratum (14) (Figure 2). Our goal was to obtain a balanced sample of the entire database since we were interested in having good representation from all patients. The selection of a smaller Pilot Sample was resource driven.

a. Carinal Registry Sample

Year	Yes	No	All
2002	33	41	47
2003	47	29	76
2004	52	47	99
2005	29	51	80
2006	43	42	85
2007	19	26	45
2008	37	20	57
2009	36	23	59
All	296	252	548

b. Hypothetical Sampling Fractions Under Proportionate Sampling Fraction of 11.6%

Year	Yes	No	All
2002	4	2	6
2003	5	3	8
2004	6	5	11
2005	3	6	9
2006	5	5	10
2007	2	3	5
2008	4	2	6
2009	4	3	7
All	33	29	64

Year	Yes	No	All
2002	0.060	0.026	0.086
2003	0.086	0.053	0.139
2004	0.095	0.086	0.181
2005	0.053	0.093	0.146
2006	0.078	0.077	0.155
2007	0.035	0.047	0.082
2008	0.068	0.036	0.104
2009	0.066	0.042	0.108
All	0.540	0.460	1.000

d. Pilot Sample Selection (equal allocation)

Year	Yes	No	All
2002	4	4	8
2003	4	4	8
2004	4	4	8
2005	4	4	8
2006	4	4	8
2007	4	4	8
2008	4	4	8
2009	4	4	8
All	32	32	64

Figure 2. Pilot Sample Selection Technique.

Disproportionate Sampling Fraction with equal allocation (n=4).

2.3 CT IMAGING INTERPRETATION PROTOCOL AND TRAINING IN THE VISUAL EMPHYSEMA SCORING METHOD

All chest CT scans were obtained following standard protocols at the University of Pittsburgh Medical Center (UPMC). However, the CT scans used in this study were not obtained under one single CT image acquisition protocol.

CT images were viewed on a Picture Archiving and Communication System (PACS) monitor display system (iSite Radiology Launcher, Stentor iSite Enterprise and Single View, Philips, 15) using standard and lung window settings.

Scoring procedures followed a semi-quantitative five level scale based on National Emphysema Treatment Trial (NETT) standards, representing no, trace, mild, moderate and severe emphysema (levels 0-4). CT scans with emphysema corresponding to the NETT standards 1.5-3.0 and >3.0 were placed into ***moderate*** and ***severe*** categories, respectively. ***Mild*** emphysema was defined by the clear presence of emphysema (usually seen only in upper lobes), which was less severe than the NETT 1.5 standard. ***Trace*** emphysema was defined by a definite finding of minimal emphysema, typically characterized by scattered centrilobular lucencies in the upper lobes (13,16).

A training module of 96 CT scans was developed by an expert thoracic radiologist for the visual emphysema scoring method. This training module was divided into 5 work-lists with CT scans pertaining at each of the five above mentioned emphysema severity categories. After reviewing the cases, these were joined in one single list in alphabetical order where there was no information available on the expert's score. The CT scans were scored twice, in two independent occasions and compared to the expert score. A third and final training session was created by randomly pulling 24 CT scans from PLuSS. These CT scans were new to the reader; the reader

had never seen neither the set nor the expert score before. Agreement was calculated with weighted kappa.

2.4 RETRIEVING CT SCAN DATA FOR THE PILOT STUDY

An automated search strategy using UPMC MARS Electronic Search Procedure (UPMC MARS ESP) was developed to find all CT scan studies of the 548 patients enrolled in the Carinal Registry. The search strategy was validated manually and it showed to be accurate. A missing CT field was defined as no CT report in MARS and no CT study images in the entire PACS system. A separate spreadsheet was built filtered to any thoracic CT with or without contrast (exam type including angiography, abdominal, total body CT, PET scan), sequence (before and after date of lung cancer diagnosis), patient identifiers, exam date, and hospital provider data for the 64 patients in the Pilot Study. ISite Radiology Launcher was used to build work lists. This application is hospital-specific requiring eight different work-lists to be built (one work list for each of 8 UPMC facilities). The ESP search procedure identified one or more thoracic CT scans for 60 patients in the Pilot Study- 4 patients did not have a CT scan in MARS, see below).

2.5 RESEARCH DESIGN

2.5.1 Visual Emphysema Scoring Method of CT studies (VESM)

The training module on emphysema scoring was reviewed and scored for emphysema in two independent occasions by the trainee. In addition, a final set of 24 independent and randomly selected CT scans was scored. Agreement was calculated by weighted kappa statistic. After completion of the training, the Pilot Study CT scans were scored for emphysema and compared to the expert reading as well. The expert score was used in further analysis due to poor agreement between the expert and the trainee scores.

3.5.1. Reliability

Intra-observer agreement was calculated between the first and second reading of the same training module by the same trainee. Inter-observer agreement was calculated between the expert's and trainee's readings of the three training sessions and the Pilot Study CT scan scoring (Appendix B).

2.5.2 Validity

The severity of emphysema as scored using the VESM both by the expert panel and trainee were compared. Sensitivity of the trainee to detect any emphysema was calculated for the three training sessions. The VESM score variable was condensed from 5 categories to only two, to detect either any emphysema or more than trace.

2.6 DATA ANALYSIS

2.6.1 Systematic Analysis of Database

To better characterize the study subjects, the demographics (age, race, and gender) among the individuals enrolled in the Carinal Registry and Pilot Study were analyzed. Additional variables in Carinal Registry and Pilot Study that we analyzed include: histological diagnosis, family history of cancer, relatives with lung cancer, relatives with other smoking-related cancer, relatives with non-tobacco-related cancer, smoking status, pack years smoked, medical history data on pulmonary disease (COPD, bronchitis, asthma), and availability of PFT data.

2.6.2 Software

Microsoft Access was used to develop the data collection forms and for storage and management of the data. Microsoft Visual Basic was used to number the samples. Stentor iSite Enterprise, iSite Radiology Launcher and Single View were used to view and score CT scans for emphysema (15). The electronic medical records resource used was MARS (Founded by John Vries, M.D. in 1989 at the University of Pittsburgh). All statistical analysis was performed using SAS software (SAS version 9.2; SAS Institute; Cary, NC).

2.6.3 Statistical Analysis

Comparisons between readers were made using weighted kappa score since observations were in the ordinal scale. Agreement was scored and interpreted as excellent if kappa was between 0.75 and 1.0, fair to good if kappa was between 0.4 and 0.75 and poor agreement if kappa was < 0.4 (17). Statistical tests (chi-square, Wilcoxon sum rank test) used a two-sided significance level of $p<0.05$. Separate analysis modeled the VESM variable as a 5-category variable (0: none, 1: trace, 2: mild, 3: moderate, 4: severe) to a 2 category variable (no/any emphysema) in order to compare with COPD variable (yes/no).

3.0 RESULTS

3.1 SPECIFIC AIM 1

3.1.1 Systematic Database Evaluation

Carinal Registry: Subject Characteristics

In the Carinal Registry, 50% of the participants are of female gender with a mean age of 68 (±8.25) (17%, 37.8%, 36.3%, and 8.9% of people aged 50-59, 60-69, 70-79, ≥80 respectively) and white race predominance (83% white, 7.5% black, 0.7% hispanic, 8.8% other). The different histological types of the tumors included Adenocarcinoma 42.7%, Squamous Cell Carcinoma 34.5%, Non-Small Cell Lung Carcinoma (NSCLC) 9.9%, Adenosquamous carcinoma 4.6%, Small Cell Lung Carcinoma (SCLC) 2.9%, Large Cell Carcinoma 2.6%, Neuroendocrine 1.6%, and Bronchiolo-Alveolar Carcinoma (BAlCa) 1.3%. All participants were ever smokers, 46.2% were current smokers and 51.8% former. Half of the study population smoked for 30 to 60 years (median 40, interquartile range of 20). The distribution according cigarette dose exposure (pack-years) was 16.8% (<30), 24.5% (30-44), 20.3% (45-59), 16.2% (60-74), 22.3% (≥75) with a mean of 53.9 (±27.83) and a median of 50 pack-years (interquartile range 38).

Almost 50% of the patients had a positive family history of cancer. Among the participants, 13.7% had a first degree relative with a lung tumor and 2.7% had a second degree

relative with a lung tumor. Ten percent had a history of one family member with a tobacco related cancer; while for non-tobacco related cancer 23.54% had one family member and 8.6% had two.

The great majority of the patients reported a history of COPD/emphysema (66.9%, with 41.6% missing data for this variable), bronchitis (13.99%, 74% missing data) or asthma (16.5%, 70% missing data). The information in the COPD variable, as for most of the variables in this database is obtained from the medical history. It is a variable coded as yes/no or missing.

Only 54% of patients had PFT information available in the Carinal Registry. Table 1 shows the demographics and smoking distribution according to the COPD status, excluding subjects with missing COPD data.

Table 1. Demographics and Smoking in the Carinal Registry population and Pilot Study sub-group by COPD status**.**

	Carinal Registry			**Pilot Sample**		
	COPD Status			**COPD Status**		
Variables	**No**	**Yes**	**Missing**	**No**	**Yes**	**Missing**
	N (%)	**N (%)**	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Sex						
Female	**55** (20.1)	**103** (37.6)	**116** (42.34)	**10** (31.3)	**10** (31.3)	**12** (37.5)
Male	**51** (18.6)	**111** (40.5)	**112** (40.9)	**5** (15.6)	**10** (31.3)	**17** (53.13)
Age categories						
50-59 years	**24** (25.8)	**25** (26.9)	**44** (47.3)	**4** (25)	**2** (12.5)	**10** (62.5)
60-69 years	**36** (17.4)	**75** (36.2)	**96** (46.4)	**5** (27.8)	**8** (44.4)	**5** (27.8)
70-79 years	**33** (16.6)	**96** (48.2)	**70** (35.2)	**4** (16.7)	**7** (29.2)	**13** (54.2)
>=80 years	**13** (26.5)	**18** (36.7)	**18** (36.7)	**2** (33.3)	**3** (50)	**1** (16.7)
Race						
White	**80** (17.6)	**189** (41.5)	**186** (40.9)	**12** (23.5)	**16** (31.4)	**23** (45.1)
Black	**11** (26.8)	**12** (29.3)	**18** (43.9)	**1** (14.3)	**2** (28.6)	**4** (57.1)
Other	**15** (29.2)	**13** (27.1)	**24** (43.6)	**2** (33.3)	**2** (33.3)	**2** (33.3)

Table 1 continued

Variables	Carinal Registry			Pilot Sample		
	COPD Status			COPD Status		
	No	Yes	Missing	No	Yes	Missing
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Smoker Status						
Current Smoker	**43** (17)	**100** (39.5)	**110** (43.5)	**4** (14.3)	**10** (35.7)	**14** (50)
Ex-smoker	**62** (21.8)	**112** (39.4)	**110** (38.7)	**10** (29.4)	**10** (29.4)	**14** (41.2)
Pack Years categories						
<30 Pack-Years	**24** (26.1)	**27** (29.4)	**41** (44.6)	**3** (30)	**2** (20)	**5** (50)
30-44 Pack-Years	**31** (23.1)	**39** (29.1)	**64** (47.8)	**2** (12.5)	**6** (37.5)	**8** (50)
45-59 Pack-Years	**17** (15.3)	**46** (41.4)	**48** (43.2)	**2** (15.4)	**5** (38.5)	**6** (46.2)
60-74 Pack-Years	**14** (15.7)	**41** (46.1)	**34** (38.2)	**2** (16.7)	**4** (33.3)	**6** (50)
>75 Pack-Years	**20** (16.4)	**61** (50)	**41** (33.6)	**6** (46.2)	**3** (23.1)	**4** (30.8)

Pilot Study: Subject Characteristics

The Pilot Study population was composed of 50% females, with a mean age of 68 ($\pm$8.85) and white race predominance (79.7% white, 10.9% black, 0% hispanic, 9.4% other). The different histological types of tumors included Adenocarcinoma 42.2%, Squamous Cell Carcinoma 32.8%, NSCLC 9.4%, Adenosquamous carcinoma 9.4%, SCLC 3.13%, Large Cell Carcinoma 0%, Neuroendocrine 1.6%, and BAlCa 1.6%. All participants were ever smokers, 43.8% were current and 53.1% former. Sixty percent of the study population smoked for 30 to 60 years (median 40, interquartile range 17). The distribution according to cigarette dose exposure (pack-years) was 15.6% (<30), 25% (30-44), 20.3% (45-59), 18.8% (60-74), 20.3% ($\geq$75) with a mean of 52 ($\pm$23.56) and a median of 50 pack-years (interquartile range 37).

Slightly more than half of the patients (51.6%) had a positive family history of cancer. Among the participants, 15.6% had a first degree relative with a lung tumor and 4.7 % had a second degree relative with a lung tumor. Ten percent had a history of one family member with a

tobacco related cancer; while for non-tobacco related cancer 21.9% had one family member and 10.9% had two.

A great majority of the patients reported a history of COPD/emphysema (57.1% with 45.3% of missing data), or asthma (26.1%, 64.1% missing data) (bronchitis 0%).

Since availability of PFT information in the Carinal Registry was our stratifying variable, 50% of the patients in the Pilot Study had PFT information available by default.

We can conclude that the Pilot Study is a good representation of its base sample, the Carinal Registry, given that the distribution of these variables does not significantly differ.

3.1.2 Systematic Analysis of CT information. Characterize CT information

The positive predictivity of the MARS ESP was 93.8%, given that in 60 of 64 cases we were able to find a CT study, leaving only 6.2% missing CT studies (4/64). Missing CT studies were seen among patients enrolled in earlier years. Between 2002 and 2005 there was 12% of missing CT studies, whereas the missing rate was null in the second half of the enrollment period (2006-2009). The relationship between year of enrollment (2002-2005 vs. 2006-2009) and availability of CT was not statistically significant ($p>0.05$). Of the 60 CT studies found, 37 patients (62%) had a CT study before (and after) diagnosis, 19 patients (32%) had a CT only after diagnosis. On the 60 patients that CT studies were found, we were able to find a total of 381 thoracic CT studies (6.35 mean number of thoracic CT studies per subject, range 1-19 per subject).

In some cases, the CT studies were suboptimal for emphysema scoring due to technical difficulties such as motion artifact, or disease-related problems such as pneumonia, pulmonary fibrosis, pleural effusion, and post-operative changes. However, only 2 patients had a single CT study available which had to be used for emphysema scoring regardless of image quality.

Forty nine patients (81.7%) had a pre-operative CT scan. These were defined by looking at the diagnosis and operative date, both available in Carinal Registry. A total of 120 pre-operative CT scans were found for these 49 patients, since in some instances there was more than one pre-operative CT scan. A total of 325 CT scans (85%) pertaining to 57 patients (95% of patients) were reconstructed using an edge-enhancing algorithm. Thirty-nine (65%) patients had both a pre-operative CT study with images reconstructed using the edge-enhancing algorithm (86 CT scans) (Table 2).

Table 2: Combination of CT scan acquisition sequence (before or after surgery) and edge-enhancing reconstruction protocol.

Before + Lung	Before + Other	After + Lung	After + Other	Frequency	Cumulative Frequency
0	**0**	**1**	**0**	7	7
0	**0**	**1**	**1**	4	11
0	**1**	**0**	**0**	1	12
0	**1**	**0**	**1**	2	14
0	**1**	**1**	**0**	6	20
0	**1**	**1**	**1**	1	21
1	**0**	**0**	**0**	8	29
1	**0**	**1**	**0**	15	44
1	**0**	**1**	**1**	2	46
1	**1**	**1**	**0**	10	56
1	**1**	**1**	**1**	4	60

Patients include the 60 patients from the Pilot Study with CT scans found. Legend: Before: CT scan obtained before surgery; After: CT scan obtained after surgery; Lung: edge-enhancing reconstruction available; Other: other reconstruction protocol available. Cumulative frequency adds up to 60, the number of patients for whom CT scans were found.

The slice thickness varied from 0.63mm to 8mm, but 5mm was the predominant CT scan slice thickness used among the Pilot Study series (79.4%). Eight CT studies (belonging to 7 different patients) had more than one slice thickness. Of these 8 CT scans, 7 had two different slice thicknesses and one had three different slice thicknesses.

The dose of the CT scan as measured in mA varied from 60mA (LDCT) to 751mA. Thirty-three percent (126/381) of the CT scans were done using CT image dose modulation, in which the dose is modulated throughout the thorax to obtain the best image quality at the lowest possible mA dose that varies depending on tissue density (18). These 126 CT scans pertained to 42 patients (70%).

We report a subjective finding from scoring patients with multiple CT studies, no information bias was encountered among those cases. The patients were scored equally effectively. The 381 cases scored for emphysema were not seen as independent studies but as series of studies related to one patient, since the CT studies could not be de-identified for this purpose.

3.2 SPECIFIC AIM 2

Specific Aim 2: To describe the distribution of emphysema scoring among lung cancer patients

3.2.1 Visual CT Scan Emphysema Score Training Modules and Agreement with Experts

The first training session showed excellent agreement with weighted kappa statistic k=0.82 (95% CI: 0.74-0.90). With the expert reading serving as gold standard, the trainee's reading achieved a

sensitivity of 98% to detect any emphysema and 78% to detect more than trace. The second training session resulted in excellent agreement as well with a k=0.83 (95% CI: 0.73-0.91), and with a sensitivity of 96% to detect any emphysema and 74% to detect more than trace emphysema. There was only one instance in which there was disagreement by more than one category. The third and final training session showed fair to good agreement with a k=0.59 (95% CI 0.39-0.79). The sensitivity to detect any emphysema was 89% (95% CI: 67-99) and a sensitivity to detect more than trace emphysema of 94% (95% CI: 70-99). There was no disagreement by more than two categories. At this point the training was considered completed and the Pilot Study CT scans were read and scored for emphysema (Appendix B1, sections 1-3).

3.2.2 Emphysema Scores Assessed by Visual Emphysema Scoring Method (VESM)

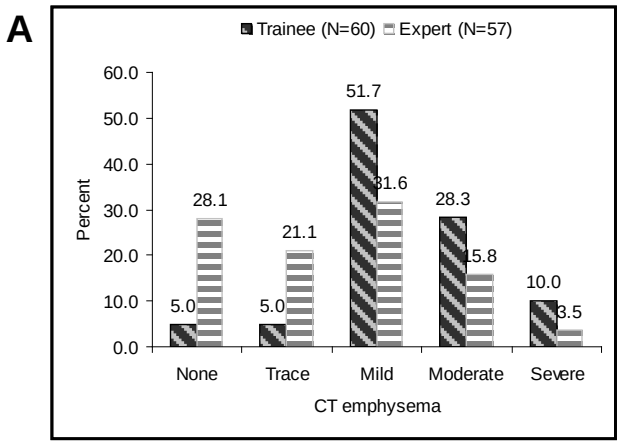
Patients in the Pilot Study were scored for their emphysema following the NETT criteria. Their scores were compared to the expert's score and weighted kappa showed poor agreement (k=0.27, 95% CI: 0.14-0.40) and asymmetry (p=0.012) (Appendix B1, section 4). Due to these results, the expert score was used in further analysis. The distribution and severity of emphysema for both raters are shown in Figure 3A. Twenty-eight percent of the study population had no emphysema, 21.1% trace, 31% mild, 15.8% moderate and 3.5% had severe emphysema. The frequency of missing data was only 10.9% (7/64). Demographics and smoking history organized by emphysema severity score are shown in Table 3.

Table 3: Pilot Study: Demographics and smoking by differential emphysema score (with percentages)**.**

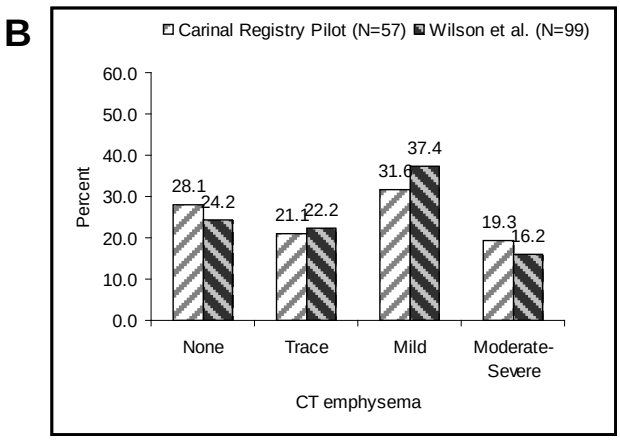
	Emphysema Score				
	None N (%)	**Trace** N (%)	**Mild** N (%)	**Moderate** N (%)	**Severe** N (%)
Sex					
Female	**9** (15.8)	**5** (8.8)	**8** (14.4)	**7** (12.3)	**1** (1.8)
Male	**7** (12.3)	**7** (12.3)	**10** (17.5)	**2** (3.5)	**1** (1.8)
Age categories					
50-59 years	**4** (7)	**4** (7)	**3** (5.3)	**2** (3.5)	**1** (1.8)
60-69 years	**4** (7)	**4** (7)	**4** (7)	**3** (5.3)	**1** (1.8)
70-79 years	**7** (12.3)	**3** (5.3)	**9** (15.8)	**3** (5.3)	**0**
>=80 years	**1** (1.8)	**1** (1.8)	**2** (3.5)	**1** (1.8)	**0**
Race					
White	**13** (22.8)	**11** (19.3)	**14** (24.6)	**7** (12.3)	**2** (3.5)
Black	**2** (3.5)	**0**	**2** (3.5)	**1** (1.8)	**0**
Other	**1** (1.8)	**1** (1.8)	**2** (3.5)	**1** (1.8)	**0**
Smoker Status					
Current Smoker	**7** (12.3)	**6** (10.5)	**7** (12.3)	**3** (5.3)	**1** (1.8)
Ex-smoker	**9** (15.8)	**6** (10.5)	**9** (15.8)	**6** (10.5)	**1** (1.8)
Pack Years categories					
<30 Pack-Years	**4** (7)	**1** (1.8)	**4** (7)	**0**	**0**
30-44 Pack-Years	**5** (8.8)	**2** (3.5)	**4** (7)	**0**	**1** (1.8)
45-59 Pack-Years	**2** (3.5)	**3** (5.3)	**4** (7)	**2** (3.5)	**0**
60-74 Pack-Years	**2** (3.5)	**2** (3.5)	**4** (7)	**4** (7)	**0**
>75 Pack-Years	**3** (5.3)	**4** (7)	**2** (3.5)	**3** (5.3)	**1** (1.8)

Wilson et al. studied the distribution of emphysema among patients at high risk for lung cancer in the frame of the PLuSS study (10). They screened 3,642 community volunteers using low dose CT (LDCT) scans with a visual method for emphysema scoring similar to the one described in this manuscript. Their emphysema findings were 75.8% for any emphysema while ours were 71.9%. We found no difference in the emphysema distribution between the two patient

populations (p>0.05) (Table 4). Figure 3B shows the distribution of the different emphysema severity scores among both patient populations from the Pilot Study and Wilson et al. study (10) to have no statistically significant reason to state that the medians differ (p>0.05).



Frequency of Radiographic Emphysema Score		
	Expert	Trainee
None	16 (28%)	3 (5%)
Trace	12 (21%)	3 (5%)
Mild	18 (31%)	31 (52%)
Moderate	9 (16%)	17 (28%)
Severe	2 (4%)	6 (10%)



Frequency of Radiographic Emphysema Score		
	Carinal	Wilson et al.
None	16 (28%)	24 (24.2%)
Trace	12 (21%)	22 (22.2%)
Mild	18 (31%)	37 (37.4%)
Moderate-Severe	11 (20%)	16 (16.2%)
Total	57	99

Figure 3. Distribution of Emphysema Scores in the Pilot Study Population.

CT scans were scored for emphysema based on the NETT criteria by the expert and trainee. **A**. Trainee vs. Expert CT scan Score Reading, Poor agreement as shown by low kappa score (see text); **B**. Carinal Registry Pilot Study emphysema score vs. Wilson et al. cases. Wilcoxon signed rank test: p=0.8574. Distributions do not differ significantly.

Table 4: Comparison of COPD variable and Visual Emphysema Score

COPD variable from the medical record, Visual Emphysema Score from visually scoring CT scans for emphysema

				Missing		Not missing		Emphysema	
	Study Group	N	Emphysema Measure	N	%	N	%	N	%
1	Carinal cases	548	Medical Records [1]	228	41.6	320	58.4	214	66.9
2	Carinal cases not in Random Subset	484	Medical Records [1]	199	41.1	285	58.9	194	68.1 *
3	Carinal Random Subset	64	Medical Records [1]	29	45.3	35	54.7	20	57.1 *
4	Carinal Random Subset	64	Visual CT Score [2]	7	10.9	57	89.1	41	71.9**
5	Carinal Random Subset	64	Visual CT Score >Trace [2]	7	10.9	57	89.1	29	50.9
6	Wilson cases	99	Visual CT Score [3]	0	0	99	100	75	75.8**

* p=0.1949 (chi-square) % of emphysema from medical record in Carinal Registry cases in random subset (Row 3) vs. Carinal Registry cases not in random subset (Row 2)
** p=0.5980 (chi-square) %emphysema by CT in Wilson et al. cases (Row 6) vs. Carinal Registry sub-sample cases (Row 4)
[1] Abstracted from medical records by laboratory investigators and recorded in Carinal Registry database
[2] Expert reader
[3] PLuSS procedure (8)

3.3 SPECIFIC AIM 3

3.3.1 Specific Aim 3a: To validate data in the Carinal Registry Database

Validation of data in Carinal Registry & Pilot Study was carried out by analysis of the COPD variable and the VESM variable (Table 4).

COPD was seen in a 66.9% of patients in the Carinal Registry, 68.1% of patients in the Carinal Registry (excluding the Pilot Study subset) and 57.1% of patients in the Pilot Study, with no significant difference in the distribution of disease ($p>0.05$), concluding that the Pilot Sample

is a good representation of the Carinal Registry. However, COPD data from the Carinal Registry and the Pilot Study yielded a frequency of missing data of 41.6% and 45.3% respectively. On the other hand, the VESM among patients in the Pilot Study showed a frequency of any emphysema of 71.9% (or 50.9% for more than trace) with only 10.9% of missing data.

Finally we also found that there was a 60% agreement between emphysema from the medical records (COPD/emphysema variable) and emphysema from CT scan reading (as per VESM variable) among the 64 patients in the Pilot Study subset (although the frequency of missing data from the COPD/emphysema variable was 39 = 60.1% of the data was missing).

3.3.2 Specific Aim 3b: To assess the reproducibility of VESM. Intra-rater and Inter-rater variability

Intra-rater variability of emphysema scores were assessed as a measure of reliability. As mentioned above, CT scans from the training module were independently read twice by the same reader. Results were compared yielding a weighted kappa statistic of 0.83 (95% CI: 0.75-0.91). The test of symmetry was not statistically significant (p = 0.991). There was one instance where the emphysema score differed by more than one category (see also section 4.2.1).

Inter-rater variability of emphysema scores was assessed as well. CT scans from the training module and from the Pilot Study were independently read twice and once respectively, by both the expert and trainee reader. The kappa scores of the training modules were presented above (sections 4.2.1 & 4.2.2, Figure 3A, and in more detail in Appendix B).

4.0 DISCUSSION

Emphysema is a subtype of COPD and it has been shown to increase the risk of developing lung cancer independently from smoking (10-12). Many pathways are involved but a seemingly convincing theory lays in the inflammatory process shared both by smoking and emphysema due to the repeated injury and repair mechanisms with high cell turnover and subsequent increased possibility of genetic errors leading to the development of a neoplasia (7,8).

We studied a large prospectively collected lung cancer case series and evaluated the data quality for the purposes of conducting research studies. A randomly selected Pilot Sample showed good correlation with its parental database, the Carinal Registry. The correlation of emphysema frequency (recorded from the medical history under the COPD/emphysema variable) among both the Carinal Registry and the Pilot Study patients was good showing no significant difference in COPD frequency distribution among both patients groups (Figure 4). These results strongly supported the design of the Pilot Sample as a sub-set of the Carinal Registry.

We studied the distribution of emphysema among lung cancer patients in the Pilot Sample case series as evaluated by the Visual CT scan score method, and observed no differences when compared to the community based study by Wilson et al. (8) (71.9% vs. 75.8% respectively, $p>0.05$). This finding suggests that there might be a shared pathophysiolological pathway between emphysema and lung cancer independently of smoking. The frequency of emphysema by analysis of COPD/emphysema (from medical records) was lower than that

obtained with the Visual CT emphysema score method (VESM), we speculated one reason could be the impact of the high frequency of missing COPD data (41.6 and 45.3%) (Table 4). Agreement between the frequency of COPD variable (yes/no) and VESM grouped (any/no) showed a high frequency of missing data (54.7%). However, we found a percent positive agreement of 60% and a percent agreement of 51.7% between these two measurements of COPD/emphysema in the same set of patients (Appendix C).

There are a few disadvantages in the semi-quantitative VESM method. One is that it requires practice and training. The training required in this method can be regarded as being time consuming and, in this present study, final scoring still required the reading of an expert radiologist. As seen in Figure 3, the expert scores differ from the trainee scores (poor agreement as seen by a kappa score of 0.27) despite the excellent level of agreement reached in the training sessions (kappa=0.82, 0.83 and 0.59 for the first, second and final sessions respectively). A possible explanation of the poor agreement is information bias since the reader was not blinded to the facts that the patients had a diagnosis of lung cancer and that they were smokers. In addition, the use of an edge enhancing reconstruction protocol to read most of the CT studies may have lowered the threshold to detect emphysema due to its better appreciation of lung parenchyma (20,21). Reproducibility of a method of research is critical in interpretation of results. This method showed high reproducibility and sensitivity to detect any emphysema (98, 96, 89% for the 1^{st}, 2^{nd} and final training sessions respectively) but ultimately, the inter-rater agreement was poor (k=0.27). On the other hand, the frequency of missing data in the VESM method is lower and the accuracy of emphysema diagnosis is higher, so these findings support the extra effort in order to obtain high quality data for research purposes.

The strategy of using an automated medical records resource such as MARS EPS maximized the CT scans found, which is essential for the quality of the database and validity of analysis with low missing data, as shown in Table 4 where the missing results from CT and the missing COPD/emphysema data are shown. In order to expand this Visual CT Scan Scoring method (VESM) to the remaining 484 patients in the Carinal Registry, one single best CT scan from each patient would have to be selected for emphysema scoring. The MARS EPS would play an important role in generating this CT scan list.

The best study would be defined as pre-operatively, preferably within one year of surgery/chemotherapy. The need for an edge-enhancing lung reconstruction remains controversial since in some studies it has shown to lower the threshold to detect mild emphysema (19,20). However, Vikgren et al. (21) found that edge-enhancing reconstruction is better than standard reconstruction to detect emphysema. As for the radiation dose, measured in mA, LDCT, in addition to lung cancer screening, has proven a good resource to evaluate the presence and severity of emphysema (22). Another strategy that has been successfully used to reduce the radiation dose to which the patients are exposed to is the dose modulation technique seen applied in several of the CT scans of this series (18). Slice thickness is another important factor in CT scanning. Cederlund et al. looked at CT scans during pre-operative evaluation for lung volume reduction surgery using an objective computer software method followed by subjective evaluation by 4 radiologists. They compared high resolution CT (HRCT, 2mm) with spiral CT (conventional 8-10mm) for classification of emphysema and found no difference between HRCT and spiral (60 & 62% agreement) (23). The study by Reske et al. agreed with their findings (20). On the other hand, another study by Cederlund et al. reported a slight benefit in using conventional spiral (47% vs. 40%) ($p<0.05$) (24). Both the slice thickness and the edge-

enhancing reconstruction as parameters to select the best CT scan to diagnose and rate severity of emphysema remain inconclusive.

Many factors may have contributed to limitations in this study. First, the CT training module is sub-optimally masked. It is blinded from the expert score, but it has personal identifiers (name, medical record number). The ideal training method would have been a de-identified and randomly ordered list of the 96 CT scans. However, PACS is a clinical entity/program that was not created for research purposes making it not a feasible option for the purposes of this manuscript. Furthermore, the fact that the PACS iSite system is site specific required the preparation of separate worklists; this fact not only makes training and reading more tedious but also can add lead population bias to the Visual CT Scan Scoring method, since the rater is not masked in terms of hospital site where the CT scan was acquired. In addition, the 381 CT studies were not viewed as independent studies but as sub-series of studies associated to one particular patient and this could have biased the scoring related to preconcepts related to different hospitals. Second, the emphysema scoring method is semi-quantitative and subjective, since it does not use software designed to score emphysema based upon shades of black and white. Also, the CT studies reviewed were not obtained under a single protocol, many indeed differ on slice thickness and image resolution, and these factors may have contributed as well. As mentioned by Friedman and Reske et al. (19,20) the edge-enhancing lung reconstruction protocols may lower the threshold for emphysema detection and favor higher detection and severity scores, as found by the trainee reading of the CT scans in the Pilot Study. Third, the fact that it is known to the investigators that all the patients have lung cancer either from inclusion criteria or from simply seeing the visual manifestations of lung cancer on CT scan, may have biased the frequency and

the level of severity of emphysema scoring towards a greater severity of emphysema. All the factors mentioned above may contribute to information bias.

5.0 CONCLUSION

This study shows that the Visual Emphysema Scoring Method is a very accurate method to score emphysema severity and carries a low frequency of missing data. When compared to the database information obtained from the medical record (COPD variable), the correlation between the two variables was poor. The VESM was a more accurate measure of COPD status among lung cancer patients enrolled in the Carinal Registry.

Training in VESM showed high reproducibility scores and high sensitivity of the trainee to detect emphysema when compared to the standard expert score.

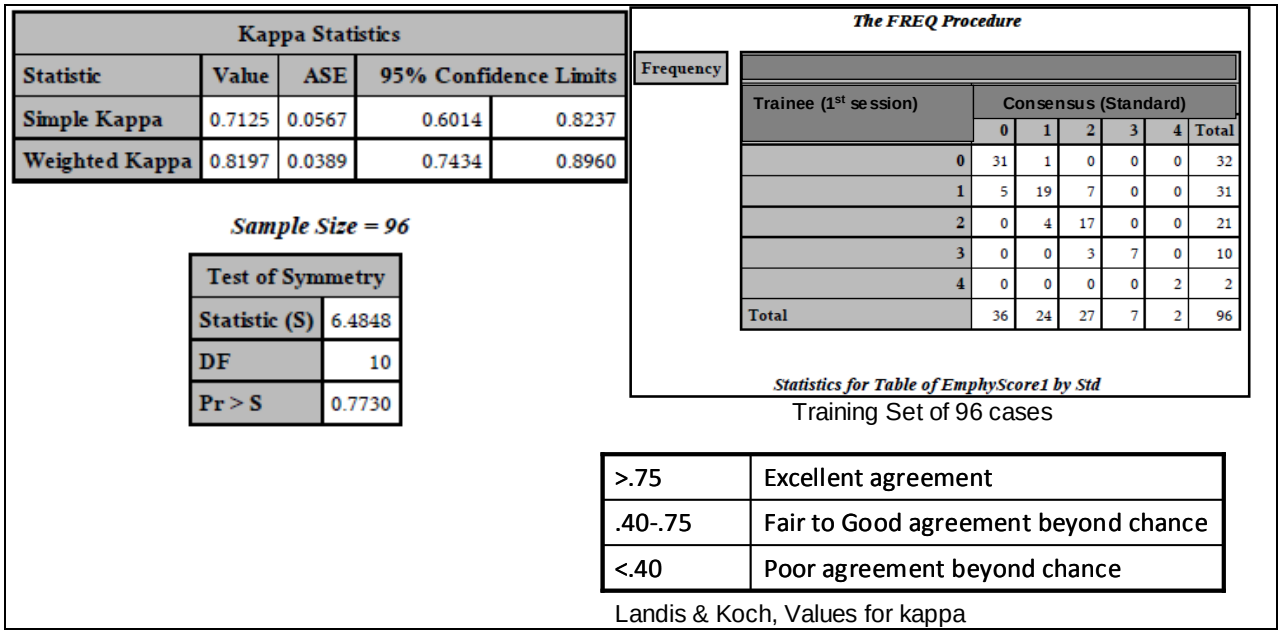
In other words, the COPD variable is not a reliable indicator of emphysema among the Carinal Registry patients. In order to better assess their emphysema severity score, the VEMS would have to be used. After performing VESM among the 64 Pilot Study patients, there are 484 remaining patients in the Carinal Registry. For practical reasons, we suggest that a single best CT scan has to be selected and we defined it as a preoperative study, within one year of surgery or diagnosis (if no surgery performed), preferably with edge-enhancing reconstruction and thin slices (less than 5mm).

APPENDIX A

ASSESSMENT OF AGREEMENT

We evaluated inter-observer agreement of the Visual emphysema Scoring Method (VESM) between the trainee and the expert. Kappa agreement scores and specific tables are presented below. There were three training sessions, the first two involved reading and scoring 96 C T scans that were used for training purposes as well. The third and final training session involved reading and scoring 24 cases never seen by the trainee before.

The first figure depicts the results of the trainee first time reading the training session of 96 cases vs. the consensus expert reading and its kappa score and confidence interval showing excellent agreement.



Kappa Statistics				
Statistic	Value	ASE	95% Confidence Limits	
Simple Kappa	0.7125	0.0567	0.6014	0.8237
Weighted Kappa	0.8197	0.0389	0.7434	0.8960

Sample Size = 96

Test of Symmetry	
Statistic (S)	6.4848
DF	10
Pr > S	0.7730

The FREQ Procedure

Frequency

Trainee (1st session)	Consensus (Standard)					
	0	1	2	3	4	Total
0	31	1	0	0	0	32
1	5	19	7	0	0	31
2	0	4	17	0	0	21
3	0	0	3	7	0	10
4	0	0	0	0	2	2
Total	36	24	27	7	2	96

Statistics for Table of EmphyScore1 by Std

Training Set of 96 cases

>.75	Excellent agreement
.40-.75	Fair to Good agreement beyond chance
<.40	Poor agreement beyond chance

Landis & Koch, Values for kappa

Figure 4: Inter-reader agreement.

Results of the trainee's first time reading of the training session were compared to the standard expert reading. Weighted kappa agreement was calculated and reference values used are shown in the right lower corner.

The second figure shows the trainee second time reading the training session of 96 cases compared to the consensus expert reading and the agreement reached by the kappa score and confidence interval, evidence of excellent agreement.

.

Kappa Statistics				
Statistic	Value	ASE	95% Confidence Limits	
Simple Kappa	0.7385	0.0553	0.6302	0.8469
Weighted Kappa	0.8266	0.0410	0.7462	0.9069

Sample Size = 96

Test of Symmetry	
Statistic (S)	4.3333
DF	10
Pr > S	0.9311

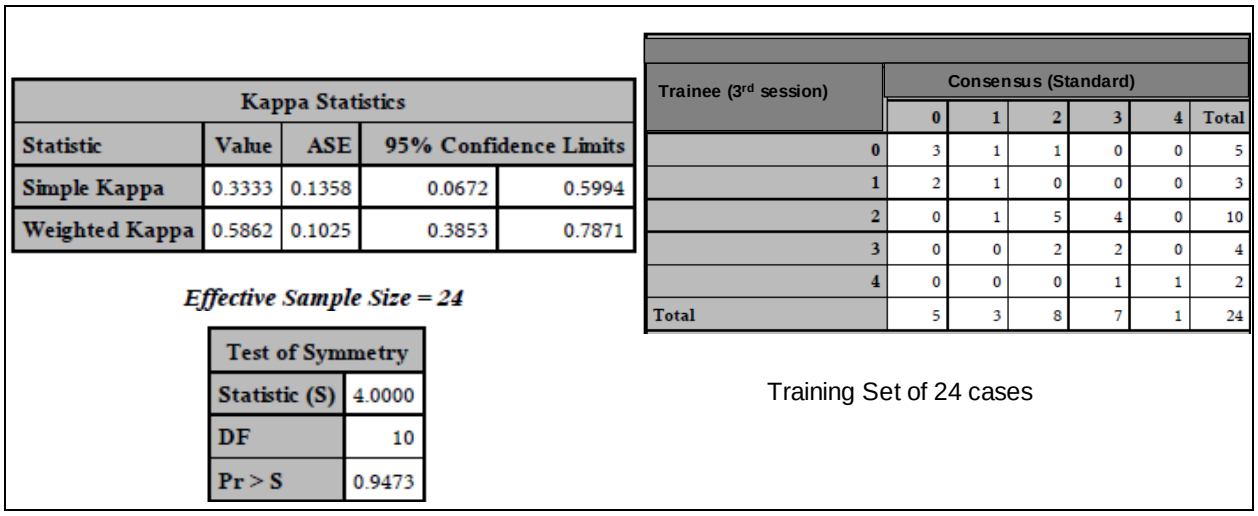
Trainee (2nd session)	Consensus (Standard)					
	0	1	2	3	4	Total
0	33	1	1	0	0	35
1	3	19	8	0	0	30
2	0	4	17	0	0	21
3	0	0	1	7	0	8
4	0	0	0	0	2	2
Total	36	24	27	7	2	96

Training Set of 96 cases

Figure 5: Inter-reader agreement.

Results of the trainee's second reading and comparison with the standard consensus expert score are shown. Agreement was calculated as weighted kappa score.

The third and final training session is shown below. Although the kappa score showed poor agreement, the sample size was significantly smaller (24 instead of 96) partially explaining the lower agreement score.



Kappa Statistics				
Statistic	Value	ASE	95% Confidence Limits	
Simple Kappa	0.3333	0.1358	0.0672	0.5994
Weighted Kappa	0.5862	0.1025	0.3853	0.7871

Effective Sample Size = 24

Test of Symmetry	
Statistic (S)	4.0000
DF	10
Pr > S	0.9473

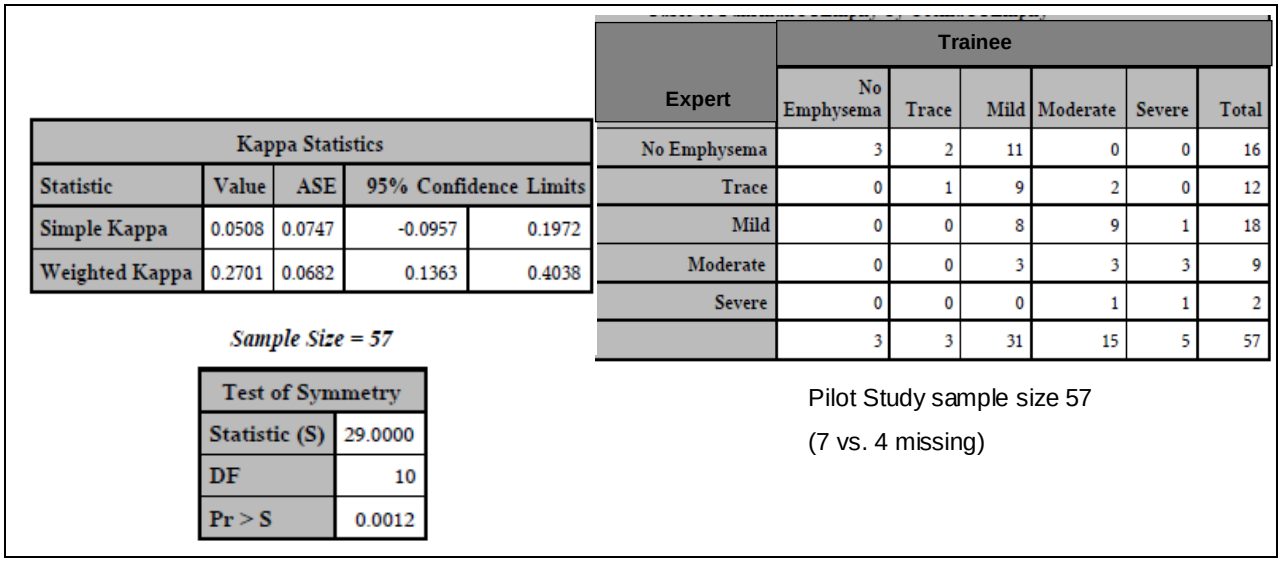
Trainee (3rd session)	Consensus (Standard)					
	0	1	2	3	4	Total
0	3	1	1	0	0	5
1	2	1	0	0	0	3
2	0	1	5	4	0	10
3	0	0	2	2	0	4
4	0	0	0	1	1	2
Total	5	3	8	7	1	24

Training Set of 24 cases

Figure 6: Inter-reader agreement.

Agreement between the trainee and consensus standard panel in the last training session is shown here as weighted kappa score.

This session shows the results of the Carinal Registry Pilot Sample scoring of the 64 cases using the VESM both by the trainee and expert. The agreement reached was poor, as shown by a low kappa score. Several factors may have influenced this result. .....



Kappa Statistics				
Statistic	Value	ASE	95% Confidence Limits	
Simple Kappa	0.0508	0.0747	-0.0957	0.1972
Weighted Kappa	0.2701	0.0682	0.1363	0.4038

Sample Size = 57

Test of Symmetry	
Statistic (S)	29.0000
DF	10
Pr > S	0.0012

Expert	Trainee					
	No Emphysema	Trace	Mild	Moderate	Severe	Total
No Emphysema	3	2	11	0	0	16
Trace	0	1	9	2	0	12
Mild	0	0	8	9	1	18
Moderate	0	0	3	3	3	9
Severe	0	0	0	1	1	2
	3	3	31	15	5	57

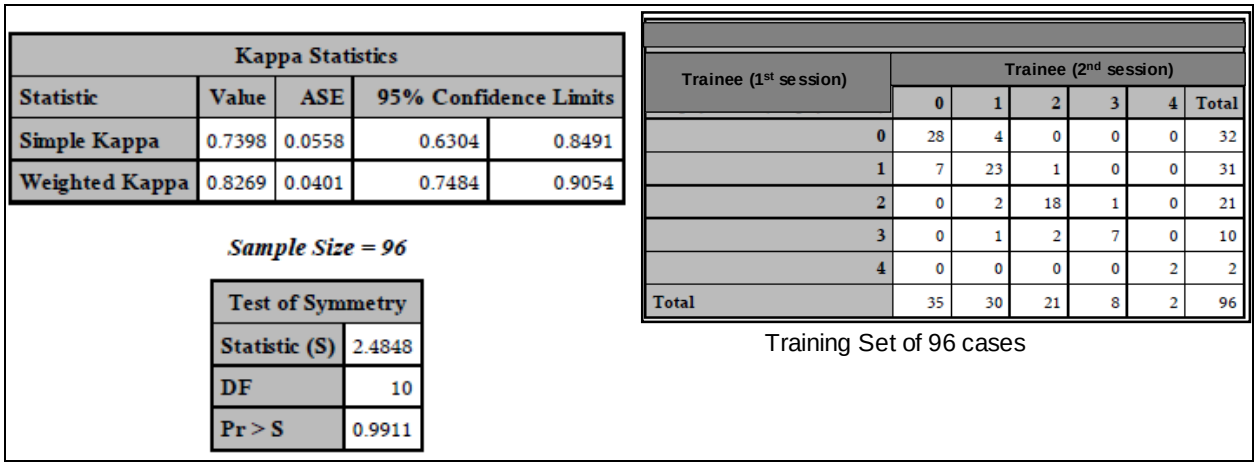
Pilot Study sample size 57

(7 vs. 4 missing)

Figure 7: Results of the Pilot Sample CT scan reading.

Agreement between the expert and trainee are shown as weighted kappa score.

We also evaluated intra-reader variability of the VESM by comparing the first and second trainee's reading of the training session of the set of 96 series. The reproducibility was significantly high achieving a kappa score of 0.83.



Kappa Statistics				
Statistic	Value	ASE	95% Confidence Limits	
Simple Kappa	0.7398	0.0558	0.6304	0.8491
Weighted Kappa	0.8269	0.0401	0.7484	0.9054

Sample Size = 96

Test of Symmetry	
Statistic (S)	2.4848
DF	10
Pr > S	0.9911

Trainee (1st session)	Trainee (2nd session)					
	0	1	2	3	4	Total
0	28	4	0	0	0	32
1	7	23	1	0	0	31
2	0	2	18	1	0	21
3	0	1	2	7	0	10
4	0	0	0	0	2	2
Total	35	30	21	8	2	96

Training Set of 96 cases

Figure 8: Intra-reader agreement.

Agreement between the first and second session was calculated as weighted kappa.

APPENDIX B

CORRELATION BETWEEN COPD VARIABLES

We evaluated the correlation between the two variables indicating COPD in the Carinal Registry Pilot Study by cross-tabulating the COPD/emphysema variable (obtained from medical records) vs. VESM variable (visual emphysema score method, obtained by reading CT scans and scoring for emphysema). Visual emphysema severity scores other than none were grouped under a unique "Any" category. Data shown was obtained from Carinal Registry Pilot Sample, including both the trainee and the expert readings. Agreement was calculated by percent positive agreement and percent agreement.

B.1 EXPERT READING

Table 5: Correlation between COPD and VESM variables (expert score).Table 5: Correlation between COPD and VESM variables (ex

COPD/Emphysema Medical Record	Any Emphysema by CT		
Frequency Percent	Yes %	No %	Total %
Yes %	12 41.38	4 13.79	16 55.17
No %	10 34.48	3 10.34	13 44.83
Total	22 75.86	7 24.14	29 100.00
Frequency Missing = 35			

Percent Positive Agreement = 60%; Percent Agreement = 51.7%

B.2 TRAINEE READING

Table 6: Correlation between COPD and VESM variables (trainee score).

COPD/Emphysema Medical Record	Trainee Any Emphysema by CT		
Frequency Percent	Yes %	No %	Total
Yes %	18 58.06	0 0.00	18 58.06
No %	13 41.94	0 0.00	13 41.94
Total	31 100.00	0 0.00	31 100.00
Frequency Missing = 33			

Percent Positive Agreement = 73.5%

BIBLIOGRAPHY

1. Hester A. Gietema, Pieter Zanen, Arnold Schilham, Bram van Ginneken, Rob J. van Klaveren, Mathias Prokop, Jan Willem J. Lammers. Distribution of emphysema in heavy smokers: Impact on pulmonary function. *Respiratory Medicine* 2010; 104, 76e82
2. Alfred P. Fishman, Jack A. Elias, Jay A. Fishman, Michael A. Grippi, Robert M. Senior, Allan I. Pack. *Fishman's Pulmonary Disease and Disorders*, 2007; Volume 1, Chapters 41 & 42, 4th edition, McGraw Hill.
3. American Cancer Society Statistics, online reference, http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-026238.pdf
4. R.P. Young, R.J. Hopkins, T. Christmas, P.N. Black, P. Metcalf and G.D. Gamble. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. *Eur Respir J* 2009; 34: 380–386
5. Koshiol J, Rotunno M, Consonni D, Pesatori AC, De Matteis S, et al. Chronic Obstructive Pulmonary Disease and Altered Risk of Lung Cancer in a Population-Based Case-Control Study. *PLoS ONE* 2009; 4(10): e7380. doi:10.1371/journal.pone.0007380
6. Tonya Walser, Xiaoyan Cui, Jane Yanagawa, Jay M. Lee, Eileen Heinrich, Gina Lee, Sherven Sharma and Steven M. Dubinett. Smoking and Lung Cancer: The Role of Inflammation. *Proc Am Thorac Soc* 2008; Vol 5. pp 811–815, DOI: 10.1513/pats.200809-100TH
7. Michelle C. Turner, Yue Chen, Daniel Krewski, Eugenia E. Calle, and Michael J. Thun. Chronic Obstructive Pulmonary Disease Is Associated with Lung Cancer Mortality in a Prospective Study of Never Smokers. *Am J Respir Crit Care Med* 2007; Vol 176. pp 285–290,
8. Gina Lee, Tonya C. Walser and Steven M. Dubinett. Chronic inflammation, chronic obstructive pulmonary disease, and lung cancer. *Current Opinion in Pulmonary Medicine* 2009, 15:303–307
9. Fabien Maldonado, Brian J. Bartholmai, Stephen J. Swensen, David E. Midthun, Paul A. Decker and James R. Jett. Are Airflow Obstruction and Radiographic Emphysema Risk Factors for Lung Cancer? A Nested Case-control Study Using Quantitative Emphysema Analysis. *Chest*. 2010;138(6):1295-302
10. Wilson DO, Weissfeld JL, Balkan A, Schragin JG, Fuhrman CR, Fisher SN, Wilson J, Leader JK, Siegfried JM, Shapiro SD, Sciurba FC. Association of radiographic emphysema and airflow obstruction with lung cancer. *Am J Respir Crit Care Med*. 2008;178(7):738-44. Epub 2008 Jun 19.
11. Juan P. de Torres, Gorka Bastarrika, Juan P. Wisnivesky, Ana B. Alcaide, Arantza Campo, Luis M. Seijo, Jesus C. Pueyo, Alberto Villanueva, Marıa D. Lozano, Usua Montes, Luis

Montuenga, and Javier J. Zulueta. Assessing the Relationship Between Lung Cancer Risk and Emphysema Detected on Low-Dose CT of the Chest. *Chest*. 2007;132(6):1932-8
12. Alyson J. Littman, Mark D. Thornquist, Emily White, Lisa A. Jackson, Gary E. Goodman, Thomas L. Vaughan. Prior lung disease and risk of lung cancer in a large prospective study. *Cancer Causes and Control* 2004; 15: 819–827.
13. National Emphysema Treatment Trial; internet source NIH http://www.nhlbi.nih.gov/health/prof/lung/nett/lvrsweb.htm visited on 12/1/10
14. Leslie Kish, Survey Sampling, 1965, John Wiley and Son, Inc.
15. Philips PACS System: http://www.healthcare.philips.com/us_en/products/healthcare_informatics/products/enterprise_imaging_informatics/index.wpd
16. Imaging of Diseases of the Chest - Hardcover (2005) by David M. Hansell, Peter Armstrong, David A. Lynch, H. Page McAdams; p752
17. Landis and Koch, Biometrics, 33, 159-174, 1977. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd edition. New York, John Wiley and Son, Inc., 1981, p 218
18. Lee CH, Goo JM, Ye HJ, Ye SJ, Park CM, Chun EJ, Im JG. Radiation dose modulation techniques in the multidetector CT era: from basics to practice. *Radiographics*. 2008;28(5):1451-9.
19. Friedman PJ. Imaging studies in emphysema. *Proc Am Thorac Soc*. 2008;5(4):494-500. Review.
20. Reske AW, Busse H, Amato MB, Jaekel M, Kahn T, Schwarzkopf P, Schreiter D, Gottschaldt U, Seiwerts M. Image reconstruction affects computer tomographic assessment of lung hyperinflation. *Intensive Care Med*. 2008;34(11):2044-53
21. Vikgren J, Friman O, Borga M, Boijsen M, Gustavsson S, Ekberg-Jansson A, Bake B, Tylén U. Detection of mild emphysema by computed tomography density measurements. *Acta Radiol*. 2005;46(3):237-45.
22. Bastarrika G, Wisnivesky JP, Pueyo JC, Díaz L, Arraiza M, Villanueva A, Alcaide AB, Campo A, Seijo L, de Torres JP, Zulueta JJ. Low-dose volumetric computed tomography for quantification of emphysema in asymptomatic smokers participating in an early lung cancer detection trial. *J Thorac Imaging*. 2009;24(3):206-11.
23. Cederlund K, Bergstrand L, Högberg S, Rasmussen E, Svane B, Aspelin P. Visual grading of emphysema severity in candidates for lung volume reduction surgery. Comparison between HRCT, spiral CT and "density-masked" images. Acta Radiol. 2002;43(1):48-53.
24. Cederlund K, Bergstrand L, Högberg S, Rasmussen E, Svane B, Tylén U, Aspelin P. Visual classification of emphysema heterogeneity compared with objective measurements: HRCT vs. spiral CT in candidates for lung volume reduction surgery. *Eur Radiol*. 2002;12(5):1045-51.