Development of CT Texture Analysis in COPD and Association with Visual Scoring and DL_{CO}

V. Lui,[1] W.C. Tan MD,[2] J.C. Hogg MD, PhD,[2] H.O. Coxson PhD[2] and M. Kirby PhD[3] | 1. Department of Computer Science, Ryerson University; 2. Centre for Heart Lung Innovation, University of British Columbia; 3. Department of Physics, Ryerson University.

Rationale

- In chronic obstructive pulmonary disease (COPD) patients, emphysema severity can be assessed using computed tomography (CT) qualitatively by Radiologist scoring and quantitatively using histogram-based density thresholding[1-2,5]
- However, studies have shown that qualitative CT emphysema scores predict outcomes such as lung cancer risk while quantitative CT emphysema does not [2]
- Qualitative CT emphysema scoring may therefore capture other information that is related to emphysema severity besides amount of emphysema, such as lesion size or clustering of emphysema, spatial distribution and heterogeneity[4-5,11]
- Novel textural constructs, such as gray level co-occurrence matrix (GLCM) and fractal dimensions (FD) have been used in various applications in medical imaging and might provide further information in the characterization of emphysema[3,6-9]

Objective & Hypothesis

Objectives

- To determine if CT texture features, such as GLCM and FD, can differentiate patients with COPD from healthy volunteers, and are related to lung function
- To determine if CT texture features are association with qualitative visual scoring
- To determine if CT texture features are significantly associated with COPD outcomes, independent of qualitative scoring and standard quantitative CT emphysema measurements

Hypothesis

- CT texture features can be developed to objectively aid in quantifying the severity of emphysema, and may provide information complementary to qualitative visual assessment

Methods

Study Population

- Clinical data and CT imaging from the visit 1 Canadian Cohort Obstructive Lung Disease (CanCOLD) study were utilized[10]
- A total of 1187 subjects were used based on the completeness of their clinical and imaging datasets
- COPD severity was defined using the Global initiative for Chronic Obstructive Lung Disease (GOLD) system. Spirometry measurements included: forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC). The diffusion capacity of the lung for carbon monoxide (DL_{CO}) was also measured
- CT images were visually scored by a trained radiologist on a scale of 0-4, emphysema prevalence is a binary variable of 0/1 indicating none/visible

Table 1. Subject Demographics

Parameter		Never-smoker (N) (n = 258)	At-Risk (R) (n = 255)	GOLD I (G1) (n = 411)	GOLD II+ (G2) (n = 263)
Sex (%)	Male	57%	56%	62%	54%
	Female	43%	44%	38%	46%
Age (Years)		66.90	66.57	67.06	66.14
BMI (kg/m^2)		27.43	28.07 ^{G1}	26.90 ^{G2}	28.32
Pack Year (Year)		<0.001 ^{R,G1,G2}	21.87 ^{G1,G2}	17.56 ^{G2}	27.13
DL_{CO} (ml_{CO}/min/mmHg)		22.22 ^{G2}	21.66 ^{G2}	21.97 ^{G2}	19.46

^{N} Significantly different from Never-smoker, ^{R} Significantly different from At-Risk, ^{G1} Significantly different from GOLD I, ^{G2} Significantly different from GOLD II

Imaging Processing and Feature Selection

Standard Quantitative CT (QCT) Measurements:

- The low attenuation areas of the lung below -950HU (LAA_{950}) and the and the low attenuation cluster (LAC) were generated

Texture Measurements:

- GLCM[12,13] consisting of Hounsfield unit from -1 to -1000 HU were created from the CT lung image, for which 23 texture features were extracted using MATLAB r2019a. Those includes the original features described by Haralick et al.[14], with later inclusion of Soh's & Tsatsoulis' [15], and Clausi's [16]. 10 FD first order statistics were derived through the image processing of the CT lung images via the *boxcount* [17] method in examining 16x16x16 window increments with *Blockproc3D* [18].

Statistical Analysis:

- Feature selection was performed using the generalized orthogonal matching pursuit (gOMP) algorithm via the MXM package using R [19]
- Statistical significance was determined via multivariate regressions models with selected features as the predictors and DL_{CO} as response

Results

Table 2. Features examined from CT lung images

First Order Statistics (of HU) and other extracted measures	Mean, standard deviation, skewness, kurtosis, 15th percentile (HU), percentage relative area less than (-950 HU, -910 HU, -856 HU), percentage above 0 HU, volume (total lung, tissue, air), LAC hole size (percentile, unique size, total count)
GLCM (0 HU to -1000 HU and -856 HU to -1000 HU)	Contrast (contr), Dissimilarity (dissi), Homogeneity (homom, homop), Inverse difference normalized (indnc), Inverse difference moment nomralized (idmnc), Energy (energy), Entropy (entro), Maximum probability (maxpr), Difference entropy (denth), Sum entropy (senth), Correlation (corm, corrp), Difference variance (dvarh), Sum of Squares: Variance (sosvh), Sum average (savgh), Sum variance (svarh), Information measure of correlation1 (inf1h), Information measure of correlation (inf2h), Cluster Prominence (cprom), Cluster Shade (cshad)
Fractal Dimensions (from LAA_{856}, LAA_{950}, LAC_{856}, LAC_{950})	Mean, Standard deviation, skewness, kurtosis, min, max, mode, percentile

Imaging Biomarker Development: Are CT texture measurements associated with COPD severity?

Table 3 Pearson's correlation plot of CT texture features with lung function and ANOVA

Feature	Pearson's r DL_{CO}	Pearson's r Pred DL_{CO}	Pearson's r FEV_1	ANOVA p-Value	Normal vs At-Risk	Normal vs GOLD I	Normal vs GOLD II+	At-Risk vs GOLD I	At-Risk vs GOLD II+	GOLD I vs GOLD II+
75th percentile Fractal Dimension - Lower Right Lobe - LAA856	-0.02	0.07	-0.11	***<0.001***	.986	***<0.001***	***.008***	***<0.001***	***.002***	***.002***
75th percentile Fractal Dimension - Middle Right Lobe - LAA856	-0.14	0.05	-0.27	***<0.001***	.979	***<0.001***	.252	***<0.001***	.113	***<0.001***
15th percentile Fractal Dimension - Right Lower Lobe - LAA950	0.12	0.17	0.10	***<0.001***	.982	.261	***<0.001***	.499	***<0.001***	***.029***
Mode Fractal Dimension - Upper Left Lobe - LAA950	0.00	0.08	-0.01	***<0.001***	.93	***.002***	***.026***	***.02***	.126	.964
Standard Deviation Fractal Dimension -Right Upper Lobe - LAC856	0.03	-0.16	0.17	***<0.001***	.138	***<0.001***	***.003***	***<0.001***	***<0.001***	.367
Standard Deviation Fractual Dimension -Right Middle Lobe - LAC950	-0.10	-0.03	-0.10	***<0.001***	.224	***<0.001***	***.016***	***<0.001***	.732	***.008***
Information measure of correlation 2 - Left Lower Thirds - GLCM (100X100)	-0.29	-0.25	-0.15	.20	.73	.996	.697	.796	.146	.475
Cluster Shade - Left Upper Thirds - GLCM(100X100)	-0.04	0.06	-0.11	.08	.994	.121	.416	.221	.58	.955
Max Probability - Left Upper Thirds - GLCM(100X100)	-0.18	-0.17	0.00	***<0.001***	.513	.143	.628	***.001***	.998	***.003***
Contrast - Left Middle Thirds - GLCM(-850HUto-1000HU)	0.10	0.08	-0.03	***.006***	.999	***.035***	.2	***.022***	.147	.95
Cluster Prominence - Right Lower Thirds - GLCM(-850HUto-1000HU)	-0.04	-0.16	-0.08	***<0.001***	.998	.984	***<0.001***	.943	***<0.001***	***<0.001***
Cluster Prominence - Right Upper Thirds - GLCM(-850HUto-1000HU)	-0.13	-0.23	-0.16	***<0.001***	.946	***<0.001***	***<0.001***	***.003***	***<0.001***	***<0.001***
Tissue Volume - Left Lower Thirds	0.45	0.07	0.51	***<0.001***	.762	***<0.001***	.251	***<0.001***	.826	***.015***
Tissue Volume - Left Middle Thirds	0.55	0.17	0.57	***.001***	.752	***.012***	.979	.205	.503	***.002***
1st Quartile Hole Size - Right Upper Lobe - LAC856	0.14	0.04	0.12	.78	.986	.951	.737	.999	.906	.934
Unique Hole Size - Right Upper Lobe - LAC856	0.01	0.08	-0.09	***<0.001***	.981	***<0.001***	***.032***	***<0.001***	.09	.067
3rd Quartile Hole Size - Right Middle Lobe - LAC950	0.16	-0.12	0.24	***<0.001***	.986	***<0.001***	***<0.001***	***<0.001***	***<0.001***	1.00

The Pearson correlation (r) shows the strength and direction of a linear relationship between two continuous variables. +/-1 shows a perfectly positive/negative relationship while 0 signifies no relationship. Its strength can be assessed as: .1 < |r| < .3 (weak correlation), .3 < |r| < .5 (moderate correlation), .5 < |r| (strong correlation).

-1 Negatively Correlated — 0 — +1 Positively Correlated

Are CT texture measurements associated with radiologist's visual score of emphysema prevalence?

Table 4.1. shows that in the covariates + QCT + CT texture model, QCT provided no additional information for predicting visual emphysema. Models with CT texture were shown to have greater AUC and lower AIC. Additionally, **Table 4.2** shows increase in the cluster prominence of the right upper thirds and being a current smoker showed the greatest increases in the odds ratio (OR) of emphysema prevalence.

Table 4.1. Logistic multivariable regression model for QCT and CT texture predictors with emphysema prevalence as response:

Model: Emphysema Prevalence Score	ROC AUC	AIC
Covariates: Age, sex, BMI, pack-years, smoking status	0.78	1195.4
+ QCT	0.79	1167.5
+ CT texture measurements	0.83	1052.7
+ QCT + CT texture measurements	0.83	1055.5

The AUC (Area under the Curve) is a measurement derived from the receiver operating characteristic (ROC) curve which serves as a metric for classification model performance. The Akaike's Information Criterion (AIC) measures the quality of the model for which the model is simulated on a different dataset, the most accurate model has the smallest AIC.

Table 4.2. Binary logistic regression model of emphysema prevalence, with covariates, QCT, and CT textures as predictors

Features	Estimate	Std. Error	z value	p-Value	OR	2.50%	97.50%
(Intercept)	-1.42	0.19	-7.54	***<0.001***	0.24	0.17	0.35
Female, Sex	-0.13	0.22	-0.57	0.57	0.88	0.57	1.36
Ex-Smoker, Tobacco Status	0.65	0.21	3.02	***0.003***	1.91	1.26	2.92
Current Smoker, Tobacco Status	1.22	0.29	4.21	***<0.001***	3.37	1.91	5.94
Age	0.07	0.08	0.78	0.44	1.07	0.91	1.26
BMI	-0.24	0.11	-2.24	***0.03***	0.79	0.64	0.97
Pack Year	0.44	0.11	4.03	***<0.001***	1.56	1.26	1.94
LAA_{950}	-0.12	0.19	-0.63	0.53	0.89	0.61	1.29
LAC	0.10	0.10	0.98	0.33	1.10	0.90	1.32
Cluster Prominence - Right Upper Thirds - GLCM(-850HUto-1000HU)	1.56	0.21	7.43	***<0.001***	4.77	3.22	7.33
Cluster Prominence - Left Middle Thirds - GLCM(-850HUto-1000HU)	-1.20	0.19	-6.19	***<0.001***	0.30	0.20	0.44
75th percentile Fractal Dimension - Lower Right Lobe - LAA856	0.28	0.14	1.98	***0.048***	1.32	1.00	1.75
3rd Quartile Hole Size - Right Upper Lobe - LAC856	0.53	0.20	2.73	***0.006***	1.71	1.16	2.50
Cluster Prominence - Right Lower Thirds - GLCM(-850HUto-1000HU)	0.34	0.13	2.68	***0.007***	1.41	1.10	1.81

*All CT texture features shown in Table 2 were included in the model.

Are CT texture measurements associated with DL_{CO}, independent of radiologist's visual score of emphysema presence?

Both visual score and QCT provided independent and significant association to DL_{CO}. When CT textures were added into the model, QCT no longer remained significant, while CT textures and visual score provided independent and significant association with DL_{CO}.

Table 5.1. Linear regression model for DL_{CO}

Model: DL_{CO}	Adjusted R^2
Covariates: Age, sex, BMI, pack-years, smoking status	0.47
+ visual score	0.47
+ LAA_{950}	0.47
+ CT texture measurements	0.62
+ visual score, LAA_{950}	0.48
+ visual score, LAA_{950}, CT texture measurements	0.62

Table 5.2. Linear regression model of DL_{CO} with visual score, LAA_{950}, CT texture measurements as predictors

Features	Estimate	Std. Estimate	t-Stat	p-Value
(Intercept)	0.30	0.04	7.33	***<0.001***
Age	-0.35	0.02	-17.83	***<0.001***
Female, Sex	-0.51	0.05	-9.60	***<0.001***
BMI	0.01	0.02	0.32	.75
Pack Year	-0.09	0.02	-3.66	***<0.001***
Ex-Smoker, Tobacco Status	0.02	0.05	0.43	.67
Current Smoker, Tobacco Status	-0.38	0.07	-5.64	***<0.001***
Visible, Emphysema Prevalence Score	-0.12	0.05	-2.48	***.01***
LAA_{950}	0.03	0.05	0.64	.52
Information measure of correlation 2 - Left Lower Thirds - GLCM (100X100)	-0.21	0.04	-5.22	***<0.001***
Max Probability - Left Upper Thirds - GLCM(100X100)	-0.14	0.04	-3.60	***<0.001***
Cluster Prominence - Right Upper Thirds - GLCM(-850HUto-1000HU)	-0.15	0.03	-5.24	***<0.001***
Tissue Volume - Left Lower Thirds	0.10	0.03	3.00	***.003***
Tissue Volume - Left Middle Thirds	0.30	0.03	8.75	***<0.001***
3rd Quartile Hole Size - Right Middle Lobe - LAC950	-0.14	0.03	-4.57	***<0.001***
Unique Hole Size - Right Upper Lobe - LAC856	0.09	0.03	3.21	***.001***
Mode Fractal Dimension - Upper Left Lobe - LAA950	-0.06	0.02	-2.34	***.02***
15th percentile Fractal Dimension - Right Lower Lobe - LAA950	0.09	0.03	2.66	***.008***
75th percentile Fractal Dimension - Middle Right Lobe - LAA856	-0.21	0.05	-4.50	***<0.001***

*All CT texture features shown in Table 2 were included in the model.

Discussion & Conclusions

- CT texture measurements significantly differentiated subjects with COPD, and were significantly associated with lung function measurements
- CT texture measures were significantly associated with qualitative visual emphysema prevalence, and were stronger predictors than standard quantitative measurements, such as LAA_{950} and LAC
- In a multivariable regression model for DL_{CO}, CT texture measurements and qualitative visual score were significant predictors, but standard quantitative CT were no longer significant in presence of CT texture
- Both visual emphysema scoring and CT texture measurements may provide independent and complementary information related to pulmonary function

Future Directions

Further investigation of the use of CT texture features is required, such as:

- Can CT texture features predict longitudinal outcomes, such as FEV_1 or DL_{CO} decline and CT emphysema progression?
- Are CT texture features associated with different emphysema subtypes, such as panlobular, paraseptal and centrilobular emphysema?
- Are CT texture features and emphysema subtype classification independently associated with pulmonary function decline in COPD?

References

1. Lynch D. A. (2014). *COPD (Miami, Fla.)*, *1*(1), 73–82.
2. Lynch, D. A. et al. (2015). *Radiology*, *277*(1), 192–205.
3. Kirby, M. et al. (2017). *Am J Respir Crit Care Med.*, *195*(1).
4. Kim, S. S, et al. (2013). *Radiology*, *266*(2).
5. Mascalchi, M., Camiciottoli, G., & Diciotti, S. (2017). *J Thorac Dis.*, *9*(9), 3319–3345.
6. Zayed, N., & Elnemr, H. A. (2015). *Int J Biomed Imaging.*, *2015*, 267807.
7. Yang, X. et al. (2012). *Medical physics.*, *39*(9), 5732–5739.
8. Chan, A. & Tuszynski, J. (2016). *R.Soc.opensci*, 3: 160558.
9. Lennon, F. E. et al. (2015). *Nat Rev Clin Oncol.*, *12*(11), 664–675.
10. Bourbeau, J. et al. (2014). *COPD*, 11:2, 125-132.
11. Gietema, H.A. et al. (2011). *Acad Radiol.* 18(6):661-71.
12. Lewandowski, Z. & Beyenal, H. (2007). *MATLAB Central File Exchange*.
13. Uppuluri, A.(2008). *MATLAB Central File Exchange*.
14. Haralick, R.M., Shanmugam, K. & Dinstein, I. (1973). *IEEE Trans. on SMC*, *3*(6), 610-621.
15. Soh, L. & C. Tsatsoulis. (1999). *IEEE Trans. on Geoscience and Remote Sensing*, *37*(2), 780-795.
16. D A. Clausi,. (2002). *Can. J. Remote Sensing*, *28*(1), 45-62.
17. Moisy, F. (2008). *MATLAB Central File Exchange*
18. Casero, R. (2011). *GitHub*
19. Tsagris, T. et al. (2019). *R package version 1.4.4*

Acknowledgments

I would like to thank the Faculty of Science Undergraduate Research Opportunities Program (URO) for salary support. I would like to personally thank you for the support and guidance from Dr. M. Kirby, Jason, Daniel, Danesh, and Diana in this research project. In addition, we would like to acknowledge all the researchers, staff and subjects in the CanCOLD study as well as VIDA Diagnostic.

Ryerson University — Faculty of Science