
http://dx.doi.org/10.1117/12.844013

Copyright 2010 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.
Automated Quantification of Pulmonary Emphysema from Computed Tomography Scans: Comparison of Variation and Correlation of Common Measures in a Large Cohort

Brad M. Keller, Anthony P. Reeves, David F. Yankelevitz, and Claudia I. Henschke

Department of Biomedical Engineering, Cornell University, Ithaca, NY, USA;
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA;
Department of Radiology, Weill Medical College of Cornell University, New York, NY, USA

ABSTRACT

The purpose of this work was to retrospectively investigate the variation of standard indices of pulmonary emphysema from helical computed tomographic (CT) scans as related to inspiration differences over a 1 year interval and determine the strength of the relationship between these measures in a large cohort. 626 patients that had 2 scans taken at an interval of 9 months to 15 months (µ: 381 days, σ: 31 days) were selected for this work. All scans were acquired at a 1.25mm slice thickness using a low dose protocol. For each scan, the emphysema index (EI), fractal dimension (FD), mean lung density (MLD), and 15th percentile of the histogram (HIST) were computed. The absolute and relative changes for each measure were computed and the empirical 95% confidence interval was reported both in non-normalized and normalized scales. Spearman correlation coefficients are computed between the relative change in each measure and relative change in inspiration between each scan-pair, as well as between each pair-wise combination of the four measures. EI varied on a range of -10.5 to 10.5 on a non-normalized scale and -15 to 15 on a normalized scale, with FD and MLD showing slightly larger but comparable spreads, and HIST having a much larger variation. MLD was found to show the strongest correlation to inspiration change (r=0.85, p<0.001), and EI, FD, and HIST to have moderately strong correlation (r = 0.61-0.74, p<0.001). Finally, HIST showed very strong correlation to EI (r = 0.92, p<0.001), while FD showed the least strong relationship to EI (r = 0.82, p<0.001). This work shows that emphysema index and fractal dimension have the least variability overall of the commonly used measures of emphysema and that they offer the most unique quantification of emphysema relative to each other.

Keywords: X-ray Computed Tomography, Computer-aided diagnosis, Emphysema, Histogram analysis, Fractal dimension, Variation analysis

1. INTRODUCTION

The introduction of high-resolution, multi-row detector CT has allowed radiologists to view the anatomical basis of emphysema from CT scans. Given that emphysema is defined as the destruction and breakdown of the alveolar air sacs in the lung, emphysematous regions are visually described as being regions of lung parenchyma that are of a significantly low density. This allows for a qualitative scoring of the extent to which an individual has emphysema present in the lungs. Computer-based scoring systems have been developed that extend this concept to allow for quantitative evaluation of emphysema from CT scans. Quantitative measures have been shown to have good relation to underlying morphometry.1-4 The majority of methods focusing on the use of density information as the primary index, either through relative area low-density parenchyma, such as the emphysema index,5,6 or distribution of low-attenuation region sizes as in the fractal dimension.7,8 This work focuses on the four most commonly used density-based scores reported in the literature: the emphysema index,9,10 the fractal dimension,11,12 the n-th percentile of the histogram,10,13 and the mean lung density.14,15

Long term studies provide an opportunity to retroactively look at population distributions of emphysema scores.16 Recently, there has been concern that the variation of these measures over time would limit the usefulness in measuring disease progression. Measure variation can be attributed to multiple sources, including varying dose,17 inspiration levels,18 and altered scanner settings.19 It is therefore important to quantify measure
distributions, to understand overall and underlying measure variability. Previous work has looked at the effect of altering scan acquisition parameter settings on reported emphysema metrics,\(^{20}\) however little work has gone into evaluating the effect of inspiration on multiple emphysema measurements or into evaluating the relationship between various measures. Shaker et al. have shown there is a relationship between volume change and metric change,\(^{21}\) but did not show the relationship for multiple measures concurrently. Information such as the correlation between inspiration and measurement change as well as inter-measurement correlation is therefore useful as it identifies which measures offer the most unique information and how much redundancy there is between measures.

The aim of this study was to evaluate emphysema measures in a large cohort and determine measure distribution and variation over a standard scan interval of approximately 1 year. In addition, it was also sought to establish what measures were offered similar information when compared with other scores and which measures were most affected by inspiration level changes.

2. METHODS

This research evaluates the distribution and variation of established quantitative measures for emphysema from low-dose, whole-lung CT scans. The relationship of the measures to lung volume is also analyzed. Finally, to quantify the amount of related information provided by the emphysema quantification scores, inter-measure correlations are computed as well. The primary measures of interest are the emphysema index, mean lung density, histogram percentiles, and fractal dimension; all of these have been promoted as measures for quantification of the underlying anatomical basis for emphysema. The inter-scan variability of these measures over a one year time period is examined in which the actual change due to the progression of emphysema is expected to be small relative to the random variation present in the measures.

As previously shown by Gietema et al., a major concern with the use of these emphysema measures has been their clinical lack of repeatability as a result of large inter-scan variations.\(^{16}\) Four primary measures commonly used for the quantification of the anatomical basis of emphysema from CT data were investigated in this work. To minimize bias due to the low density volume in the major airways on the density-based metrics, the major airway structures (trachea, main bronchi, etc...) were removed from consideration using a segmentation method as described by Lee et. al.\(^ {22}\)

The emphysema index is the classic measure of pulmonary emphysema and is the relative volume of the lung that falls below a given density threshold. The emphysema index is also commonly called the percent low attenuation area (\(\%\text{LAA}\)). Given Given L as a set of pixels belonging to the lung field contained within a CT

Figure 1. Sample emphysema index computed at -910 HU from a whole lung CT scan. Left: Standard axial CT slice. Center: Density mask (white regions) denoting emphysematous regions defined by attenuation area of less than -910 HU superimposed. Right: Overlap of emphysema regions on original scan. It should be noted that emphysema index, histogram percentiles and the fractal dimension are all based on the computation of the the density mask.
image, the emphysema index (EI) can be calculated as

$$EI = \frac{|\{p_i : p_i \epsilon L, I(p_i) < T\}|}{|L|}$$

where $I(p_i)$ is the density has also been shown that the mean lung density cany value of the pixel $p_i$ in the CT image being analyzed and $T$ is the density level for which the emphysema index is being calculated. In this study, $T$ is set to -910 H.U. as this level is believed to encompass all regions in the lung parenchyma with some level of emphysema present. An sample calculation on an axial CT slice is given in figure 1.

The second measure of interest is the Nth percentile of the histogram. This returns the density value that would return an emphysema index of N. Thus, given $L$ as the set of pixels belonging to the lung field contained within a CT image, the $N^{th}$ percentile point $T$ can be calculated as

$$T = \sup \left\{ \tau : \frac{|\{p_j : p_j \epsilon L, I(p_j) < \tau\}|}{|L|} \leq N \right\}$$

where $\tau$ is any density level that allows for satisfaction of the given conditions, $I(p_j)$ is the density value of the pixel $p_j$ in the CT image being analyzed, and $N$ is bounded [0-1]. In this work, N is set to be the 15th percentile in order to be at a comparable level to the emphysema index.

The third measure is the fractal dimension. This returns a value that is indicative of the distribution of emphysema region sizes within the lung and can be computed by fitting a cumulative frequency distribution of low-attenuation-area sizes, $Y$, to a power law of size $X$ of the form

$$Y = K \times X^{-D}$$

due to the fractal properties of the lung airways.\textsuperscript{7} The exponent D indicates the level of disease severity and has been shown to be mostly related to emphysematous change and, in addition, can help catagorize emphysema severity in persons with asthma.\textsuperscript{23}

The final measure of interest is the mean lung density. Given a set of pixels belonging to the lung field contained within a CT image, $L$, the mean lung density of a CT image can be calculated as

$$MLD = \sum \left\{ \frac{I(p_i)}{|L|} : p_i \epsilon L \right\}$$

where $I(p_i)$ is the density value of the pixel $p_i$ in the CT image being analyzed.

Given that the measures investigated in this study are on different scales relative to one another, normalization of the practical range was used to allow for direct comparison.\textsuperscript{20} For example, the standard reporting of emphysema index is as percentage score ranging from 0% in the healthy case to approximately 70% in the most severe cases clinical cases, although the true range of the measure is 0%-100%. In order for accurate comparisons between these various measures, we normalize the measures to be on the same scale in order to bring them into agreement. This was done by empirically determining the effective range of the measures as seen in the dataset, and bringing them inline with a 0 to 100 scale. This is accomplished by calculating minimum and maximum values for each of the measures, and scaling that range to 0-100.

To understand how these measures vary for a screening population, we first computed the distribution of the emphysema measures, including the mean and empirical 95% confidence interval for both the non-normalized and normalized measures. To assess variability of these measures, we compute relative and absolute differences between the measures for each scan pair. These differences are also reported using original and normalized score. Spearman’s rank correlation coefficient is also calculated to determine how much unique information is provided by each measure relative to the others with regards to detecting disease state change. Finally, in order to determine how much variation can be attributed to volume differences in a cohort, Spearman correlation coefficients are computed for change of each measure versus lung inspiration volume. Pearson correlation is also computed to determine the strength of linear dependance, if any. The dependence of the measures on lung volume would indicate how much the measure variability can be accounted for by inspiration volume differences and, therefore, not directly related to the anatomical basis of the disease.
Table 1. Measurement Distributions for Emphysema Index (EI), Mean Lung Density (MLD), 15th Percentile of the Histogram (HIST), and Fractal Dimension (FD). Top: Distribution of original scores, Bottom: Distribution of normalized scores. Reported statistics are mean (μ), median, standard deviation (σ) and the lower and upper bounds of the empirical 95% confidence interval (95% CI).

<table>
<thead>
<tr>
<th>Measure</th>
<th>EI</th>
<th>MLD</th>
<th>HIST</th>
<th>FD</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>27.05</td>
<td>-0.66.27</td>
<td>-933.07</td>
<td>-1.24</td>
</tr>
<tr>
<td>Median</td>
<td>26.40</td>
<td>-809.90</td>
<td>-935.00</td>
<td>-1.16</td>
</tr>
<tr>
<td>σ</td>
<td>10.13</td>
<td>30.90</td>
<td>20.29</td>
<td>0.42</td>
</tr>
<tr>
<td>95% CI</td>
<td>8.72 to 48.11</td>
<td>-853.10 to -734.40</td>
<td>-966.00 to -884.90</td>
<td>-2.25 to -0.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>EI</th>
<th>MLD</th>
<th>HIST</th>
<th>FD</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>38.64</td>
<td>53.14</td>
<td>53.07</td>
<td>58.59</td>
</tr>
<tr>
<td>Median</td>
<td>37.71</td>
<td>54.95</td>
<td>55.00</td>
<td>61.65</td>
</tr>
<tr>
<td>σ</td>
<td>14.48</td>
<td>15.45</td>
<td>20.29</td>
<td>16.61</td>
</tr>
<tr>
<td>95% CI</td>
<td>12.45 to 68.73</td>
<td>17.20 to 76.55</td>
<td>4.90 to 86.00</td>
<td>17.82 to 82.52</td>
</tr>
</tbody>
</table>

3. DATA

The image data used in this work was derived from a long term CT study at the Weill Medical College of Cornell University. In order to eliminate the variation caused by changes in scanner acquisition settings, only cases where two consecutive scans were acquired with constant settings were used for this work. For this work, 626 cases with 2 scans (1252 scans) taken at an interval time difference between 9 months and 15 months (μ: 381 days, σ: 31 days) were selected. All scans were acquired at a 1.25mm slice thickness with a LightSpeed Ultra MDCT scanner at 120 kVp using a low dose protocol. Recently, low-dose protocols have been shown to offer comparable information to standard dose scans for the purposes of evaluating COPD and emphysema. 24,25 For each scan, the emphysema index, 15th percentile of the histogram, fractal dimension, and mean lung density were computed using in-house developed software.

4. RESULTS

4.1 Emphysema Metric Distributions

For the study cohort of 1252 screening patients analyzed, we found 95% of emphysema indices fell in the range of 8.7 to 48.1, covering the range of mild to severe. All scans were acquired using the same scanner parameter settings. Table 1 gives the distribution of measures for the study cohort for the four most commonly used emphysema measures. In order to allow the most accurate comparison between the measures, both original and normalized scores of the various measures are reported. Using emphysema index as a baseline, both mean lung density and the 15th percentile of the histogram have a slight comparable bias toward severity of emphysema for this cohort when all measures are normalized, while the fractal dimension shows a large degree of additional severity.

4.2 Measurement Variability

626 scan pairs between with an interval time difference between 9 months and 15 months (μ: 381 days, σ: 31 days) using identical scanner parameter settings were analyzed to determine the inter-scan differences in measurements associated with a screening cohort. Table 2 shows the variability of the normalized measures when taking account both relative and absolute change distributions of the measures. When looking at the relative variation of the measures; both emphysema index and the fractal dimension have roughly equivalent distributions, as do the mean lung density and 15th percentile of the histogram, with the latter measures having slightly more variability. However, when looking at the absolute variation between measures, a larger discrepancy in the variability between the measures is observed, with the emphysema index having the least absolute variation and the 15th percentile of the histogram being most variable.
Table 2. Distribution of Score Variation for Normalized Emphysema Measures. Top: Distribution of relative change in normalized score, Bottom: Distribution of absolute change in normalized score. Reported statistics are mean (μ), median, standard deviation (σ) and the lower and upper bounds of the empirical 95% confidence interval (95% CI).

<table>
<thead>
<tr>
<th>Measure</th>
<th>ΔE</th>
<th>ΔMLD</th>
<th>ΔHIST</th>
<th>ΔFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>-0.24</td>
<td>0.42</td>
<td>-1.03</td>
<td>-0.52</td>
</tr>
<tr>
<td>Median</td>
<td>-0.14</td>
<td>0.40</td>
<td>-1.00</td>
<td>-0.78</td>
</tr>
<tr>
<td>σ</td>
<td>7.72</td>
<td>10.12</td>
<td>11.65</td>
<td>9.29</td>
</tr>
<tr>
<td>95% CI</td>
<td>-15.00 to 14.97</td>
<td>-18.56 to 20.39</td>
<td>-26.40 to 20.00</td>
<td>-18.45 to 18.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>ΔE</th>
<th>ΔMLD</th>
<th>ΔHIST</th>
<th>ΔFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>5.75</td>
<td>6.69</td>
<td>8.07</td>
<td>6.51</td>
</tr>
<tr>
<td>Median</td>
<td>4.57</td>
<td>4.55</td>
<td>6</td>
<td>4.8</td>
</tr>
<tr>
<td>σ</td>
<td>5.15</td>
<td>7.60</td>
<td>8.47</td>
<td>6.63</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.14 to 18.40</td>
<td>0.30 to 31.1</td>
<td>0.10 to 34.0</td>
<td>0.18 to 24.37</td>
</tr>
</tbody>
</table>

Table 3. Pairwise Inter-measure Variation Correlation. Moderate to strong correlations can be found when comparing the inter-scan change in one measure to the variation in another measure.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Spearman’s ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔE : ΔMLD</td>
<td>-0.89</td>
</tr>
<tr>
<td>ΔE : ΔHIST</td>
<td>-0.92</td>
</tr>
<tr>
<td>ΔE : ΔFD</td>
<td>0.83</td>
</tr>
<tr>
<td>ΔMLD : ΔHIST</td>
<td>0.80</td>
</tr>
<tr>
<td>ΔMLD : ΔFD</td>
<td>-0.86</td>
</tr>
<tr>
<td>ΔHIST : ΔFD</td>
<td>-0.72</td>
</tr>
</tbody>
</table>

Table 4. Correlation between Inspiration Volume Change and Measurement Change. Moderately strong correlations can be found between all measures and inspiration change, with mean lung density showing the strongest relationship (ρ=−0.87, p<0.001).

<table>
<thead>
<tr>
<th>Measures</th>
<th>Pearson’s r²</th>
<th>Spearman’s ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔVolume : ΔE</td>
<td>0.772</td>
<td>0.743</td>
</tr>
<tr>
<td>ΔVolume : ΔMLD</td>
<td>-0.830</td>
<td>-0.857</td>
</tr>
<tr>
<td>ΔVolume : ΔHIST</td>
<td>-0.664</td>
<td>-0.611</td>
</tr>
<tr>
<td>ΔVolume : ΔFD</td>
<td>0.723</td>
<td>0.714</td>
</tr>
</tbody>
</table>

4.3 Intra-metric Correlation
Table 3 presents the inter-measure correlations in variation between the 4 measures investigated in this study. Using emphysema index variation as the baseline metric of disease progression, we found that the most additional information comes from the fractal dimension (ρ=0.72, p<0.001) and the mean lung density (ρ=0.83, p<0.001). We also found that the 15th percentile of the histogram provides little additional information beyond the emphysema index, as evidenced by a strong correlation (ρ=0.92, p<0.001) between the two measures. The weakest relationship overall was between the fractal dimension and the 15th percentile of the histogram.

4.4 Correlation to Inspiration Volume
For most measures, there is a moderately-strong correlation between variation in inspiration volume and measure difference, as is seen in Table 4. Figure 2 illustrates this relationship between inter-scan change in inspiration volume and associated change in emphysema index. Interestingly, we found that mean lung density had the strongest correlation to volume (ρ=−0.87, p<0.001), which implies that most of the variation can be explained by varied inspiration levels.

5. DISCUSSION
This study provides information regarding the inter-scan variability of several commonly used quantitative measures of emphysema from whole-lung CT scans on a large cohort. Investigating the distribution of the measures allows for the understanding of the relationship between them as they represent a cohort. We found that after
Variation of Emphysema Index vs. Inspiration Volume Variation

Figure 2. Emphysema Index variation as a function of inter-scan inspiration change. Linear regression analysis provides a line-of-best-fit (shown) and a Pearson correlation of $r^2 = 0.596$.

normalization, the mean lung density, the 15th percentile of the histogram, and particularly the fractal dimension, tend to report higher severities of emphysema present in a given scan as compared to emphysema index. This is most notable when viewing the normalized distribution of measures in Table 1. This effect should be noted during longitudinal COPD studies, especially when one of these measures is used as a baseline scoring from CT data.

Several thresholds have been proposed and evaluated for use in the emphysema index for various reasons such as the slice thickness of scans$^{26}$ and dose, and commonly has a range ranging from -970 H.U. to -830 H.U.$^{3,27,28}$ As there is no overall consensus in which density threshold is most appropriate,$^{29}$ -910 HU was used in order to ensure that all possible emphysematous regions of the lung were selected in our quantitative evaluation, and comparable parameters were used in all measures in order to allow valid comparisons to be made between the four measures investigated.

The inter-scan variability of quantitative measures of emphysema is important to understand in longitudinal studies in order to accurately assess true change in a person’s status versus the random variability present in the measure. Previous studies have found that CT scanner settings and dosing, as well as the level of inspiration, can have significant effects on quantitative measurements.$^{17-20}$ However, no study to the author’s knowledge has compared several of these quantitative measures concurrently and on a very large cohort. In that context, we found the emphysema index tended to produce lower inter-scan measurement variation, indicating that it may be the most useful single measurement for gauging patient change. The fractal dimension had a comparable variation, which agrees with Mitsunobu et al. in that the fractal dimension is more correlated with changes in disease state versus other possible causes of variation, such as changes in fibrosis and asthma.$^{23}$ In contrast,
we found that both the mean lung density and 15th percentile of the histogram had almost twice the inter-scan variability, implying that neither measure would be useful in long-term observation. However, since it has been noted that an increase in extent emphysema tends to be associated with hyperinflation, the high correlation of change in mean lung density and change in lung volume agrees with the idea that the mean lung density may still be useful in detecting that aspect of disease progression. The challenge remains in determining how much of the effect is caused by inspiration changes versus true disease progression and would require further investigation.

As has been described in previous works, lung density is affected by changes in inspiration level. It is also commonly known that quantitative emphysema measures from CT scans are dependent on the level of inspiration. Even using spirometry-gated inspiration, there can still be upwards of thirty percent variability in inspiration volumes. Therefore, as there appears to be no apparent advantage in using spirometric gating, this study also looked at the relationship between the measures evaluated and lung volume change in order to determine how much variation can be attributed to volume differences in a cohort. We found that in general that the variation in fractal dimension and histogram percentiles are least related to changes in inspiration level, which would imply that the two measures could be beneficial in measuring emphysema in studies where spirometric gating is unavailable due to the reduction of one known source of measure variation.

The investigation of inter-measure correlation is also relevant to longitudinal studies of emphysema as it gives the relationship between how much additional information can be gained through the use of multiple measures versus the use of a single measure. This can become critical in some instances, as the use of multiple measures to make a singular conclusion could allow for random variations in a particular measure to become misleading in analysis. Therefore it becomes important to note how much overlap there is between measures. Both mean lung density and the fractal dimension give similar levels of additional information when used to supplement the emphysema index. Given the lower inter-scan variability of the fractal dimension, it would seem likely the fractal dimension should be used in place of mean lung density when analyzing emphysema. It should be noted, however, that change in mean lung density more correlated to volume change than the fractal dimension, and therefore could be giving different, yet relevant information about disease state. Given the strong correlation between emphysema index and the 15th percentile, in general, only one should be used for assessment, and the other reported for completeness.

6. CONCLUSION
This work established several inter-measurement relationships for the primary measures of emphysema from CT. It was found that fractal dimension and emphysema index have less inter-scan measurement variability as compared to N-th percentile of the histogram and mean lung density. It was shown that inter-scan variation of the four measures have moderately strong relationships to one another ($|\rho|=0.72-0.92$). Finally, all measures were shown to have moderate correlation to the volume change ($|\rho|=0.61-0.86$), with the weakest being fractal dimension and the strongest relationship being with the mean lung density.

ACKNOWLEDGMENTS
Dr Yankelevitz and Dr. Reeves are co-inventors on patents and other pending patents relating to evaluation of diseases of the chest including measurement of nodules. Some of these, which are owned by Cornell Research Foundation (CRF) are non-exclusively licensed to General Electric and they receive royalties from CRF pursuant to Cornell policy, which in turn is consistent with the Bayh-Dole Act. Dr Henschle is a co-inventor on patents and other pending patents relating to evaluation of diseases of the chest including measurement of nodules, some of which are owned by Cornell Research Foundation (CRF) are non-exclusively licensed to General Electric, but has divested herself of all financial or other interests.

REFERENCES


