DERIVATION OF A TEST STATISTIC FOR EMPHYSEMA QUANTIFICATION
Gonzalo Vegas-Sanchez-Ferrero1,2, George Washko1, Farbod N. Rahaghi1
Maria J. Ledesma-Carbayo2, R. San José Estépar1

1 Applied Chest Imaging Lab., Brigham and Women’s Hospital, Harvard Medical School, Boston, USA
2 Biomedical Image Technologies, Universidad Politécnica de Madrid & CIBER-BBN, Madrid, Spain

ABSTRACT
Density masking is the de-facto quantitative imaging phenotype for emphysema that is widely used by the clinical community. Density masking defines the burden of emphysema by a fixed threshold, usually between -910 HU and -950 HU, that has been experimentally validated with histology. In this work, we formalized emphysema quantification by means of statistical inference. We show that a non-central Gamma is a good approximation for the local distribution of image intensities for normal and emphysema tissue. We then propose a test statistic in terms of the sample mean of a truncated non-central Gamma random variable. Our results show that this approach is well-suited for the detection of emphysema and superior to standard density masking. The statistical method was tested in a dataset of 1337 samples obtained from 9 different scanner models in subjects with COPD. Results showed an increase of 17% when compared to the density masking approach, and an overall accuracy of 94.09%.

Index Terms— Emphysema quantification, statistical test, non-central Gamma, truncated random variable.

1. INTRODUCTION
Chronic Obstructive Pulmonary Disease (COPD) is projected to be the 3rd leading cause of death worldwide by 2020 [1, 2]. Thus, an accurate diagnosis, quantification, and assessment of the progression of COPD is of paramount importance. For this purpose, density masking is currently the method of choice for noninvasive assessment of the anatomical changes caused by emphysema [3], one of the pathological components of COPD.

Pulmonary emphysema is defined as an abnormal and permanent enlargement of the airspaces distal to the terminal bronchioles, accompanied by destruction of their walls, without obvious fibrosis [4]. Emphysema is anatomically described at the level of the secondary pulmonary lobe, the smallest functional unit of the lung that has a nominal size of 25 mm³. The characterization of the different stages and types of emphysema follows different patterns according to the enlargement of airspaces, leading to lower intensity levels in the image within the secondary pulmonary lobe. Density masking provides a detection by a deterministic thresholding of the intensity levels in the CT image (in Hounsfield units) within the lung region.

Although several attempts have been made to include spatial texture, structure and location [5, 6, 7], density masking has become the de-facto phenotype for clinical studies [8]. Several thresholds between -910 HU and -950 HU have been proposed based on the correlation with pathological samples of diseases [9].

A fixed threshold has the caveat that is dependent on image formation parameters like reconstruction kernels, slice thickness and dose [10]. Different experimental methods have been proposed to compensate these effects [11, 12, 13]. More recently, a Hidden Markov models have been proposed to adapt to the local statistics of the signal while exploiting spatial coherence [14, 15]. In all these cases, weak assumptions are made about the underlying distribution of intensities without any formal formulation. Furthermore, density masking does not employ the notion that emphysema is defined within the secondary lobe unit.

In this work, we proposed a new formulation for emphysema quantification based on statistical inference within the secondary pulmonary lobe. Our approach is based on a non-central Gamma distribution assumption for local intensities that we empirically validate. Then, emphysema quantification is cast in terms of a hypothesis testing problem based on the NOAT test statistic. The proposed test statistic is inspired by the density masking philosophy, that has shown good historical correlations, and includes assumptions about the partial volume effect between airspaces within the acinus. The test statistic follows a truncated Gamma distribution whose PDF and CDF are analytically derived in here. These derivations allow us to propose a constructive method for the determination of the threshold to accept/reject the test at a given significance level.
2. STATISTICAL CHARACTERIZATION OF EMPHYSEMA IN SECONDARY PULMONARY LOBE

We propose to characterize statistically the behavior of normal and emphysematous tissue with a family of parametric distributions that include the qualitative behavior commonly shown in CT scanners with a clear skewness to the higher attenuation levels, its origin is located in $-1000$ HU and approximates 0 with a derivative of higher order than 1. In Fig. 1 the histogram of normal pulmonary tissue is depicted clearly showing these characteristics. This qualitative behavior can be described by means of the non-central Gamma (nc-$\Gamma$) distribution, whose PDF is defined as follows:

$$f_X(x; \alpha, \beta, \delta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} (x - \delta)^{\alpha-1} e^{-\frac{x-\delta}{\beta}}; \quad x \geq \delta$$  \hspace{1cm} (1)

where $\delta$ is the location parameter, $\alpha$ accounts for the shape and $\beta$ for the scale and $\Gamma(\cdot)$ is the Euler’s Gamma function.

The statistical characterization of the secondary pulmonary lobe assumes that there are two different kinds of tissues: normal tissue (NT) accounting for the description of conventional parenchyma, and panlobular emphysema (PL), which implies the complete destruction of the entire acinus in the secondary pulmonary lobule causing an increase of airspaces and their volume. We will consider that any degree of detectable emphysema with CT scans is a combination of the statistical response of normal tissue and airspaces and, thus, it can be described by a combination of these two cases.

Empirical validation. In order to see the performance of the model for describing both cases, a set of 148 regions of Panlobular emphysematous tissue and 370 regions of Normal Tissue were fitted. The regions were manually selected by a professional pulmonologist in patches of size $24.18 \times 24.18$ mm$^2$. These ROIs were acquired from 16 sites hosting CT scanners by 3 vendors and a total of 9 specific models in subjects with COPD. This is a representative sample of a general population and a wide variate of scanners. In Fig. 1 we show the fitting of the proposed model for the whole set of regions. Note that the statistical behavior is described by the model for each of the two kinds of tissue considered. The parameters obtained for each of the tissues are: $\Theta_0 = \{\alpha_0, \beta_0, \delta_0\} = \{20.95, 2.18, -1050\}$ for NT tissue, and $\Theta_1 = \{\alpha_1, \beta_1, \delta_1\} = \{11.81, 2.58, -1050\}$.\[^1\]

Emphysema detection. The characterization of emphysema by a statistical behavior resulting from a combination of NT and PS tissues lead to the definition of the most likely threshold to distinguish between both kinds of tissues. This can be obtained by simply equalizing the classification error for NT type to the classification error for PL type as $F_X(t|\Theta_1) = 1 - F_X(t|\Theta_0)$.

So, for the analyzed dataset, this threshold is located at $t = -940.74$ HU. Note that this value is in the range normally used in density masking for $t \in [-950, -910]$ [3].

So far, the definition of this threshold by considering the attenuation response of airspaces and normal tissue differs little from the perspective of the density masking. However, the combination of the probabilistic model proposed for normal and emphysematous tissue with this optimal threshold will allow us to propose a statistical test that can evaluate those pixels below this threshold to increase the classification performance within an inferential and hypothesis testing framework. The next section is devoted to the derivation of the statistical test.

3. STATISTICAL TEST FOR EMPHYSEMA

3.1. Definition of the statistic

For the development of the statistical test we consider the same hypothesis assumed in the mask density method: emphysema implies a decrease of the attenuation levels due to the destruction of lung tissue. With this hypothesis, we want to define a statistical test that accounts for samples below the threshold that differentiates normal tissue and destruction of tissue, so the statistical test will consider samples below the threshold between normal and PL tissue calculated in the previous section. Additionally, the partial volume effect of combinations of airspaces and normal parenchyma will affect the attenuation levels of samples below this threshold. This fact allows us to define the mean of truncated samples as the ideal statistic for testing emphysema, where the truncation is set to the threshold $t$, i.e. from a set of samples $X = \{x_i\}_{i=1}^N$ we consider $Y = \{y \in X : y < t\}$ and the statistic:

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^{n} y_i, \text{ with } y_i \in Y$$  \hspace{1cm} (2)

where $n$ is the number of samples in $Y$. So, if we assume that samples $X$ come from normal parenchyma, the resulting samples would follow a truncated nc-$\Gamma$ distribution with

[^1]: Note that the location parameter $\delta$ is set to $-1050$ HU because noisy response of the attenuation may cause more negative attenuation values than $-1000$ HU.
PDF:
\[
f_Y(y; \Theta_0) = \frac{(y - \delta_0)^{\alpha_0 - 1} e^{-\frac{y - \delta_0}{\beta_0}}}{\Gamma(\alpha_0)\beta_0^{\alpha_0}} F_X(t, \Theta_0),
\]
where \( y \in [0, t] \), \( \Theta_0 = \{\alpha_0, \beta_0, \delta_0\} \) and \( F_X(\cdot, \Theta_0) \) is the cumulative distribution function (CDF) of the \( \text{nc-\Gamma} \) distribution describing the original random variable \( X \) already introduced.

3.2. Derivation of probabilistic distribution of the statistic

In order to define a statistical test for hypotheses involving the sample mean of truncated \( \text{nc-\Gamma} \) variates, the PDF of the sample mean should be estimated. Many efficient methods can be applied for the estimation of the sample mean distribution of \( \text{nc-\Gamma} \) variates. In this work we choose the Edgeworth expansion which approximates the distribution around the standard Gaussian distribution by means of combinations of Hermite polynomials whose coefficients depend on the moments of the random variable. This approach has been demonstrated to be an asymptotic expansion of the PDF and has good convergence properties [16].

In our case, let \( Y_n = \frac{\sum y_i - E(Y)}{\sqrt{\text{var}(Y)}} \) be a variate with CDF \( F_n(\cdot) \) where \( Y_1, \ldots, Y_n \) are independent and identically distributed (IID) random variables distributed as in eq. (3). The Central Limit Theorem states that \( F_n(z) \to \Phi(z) \) for every \( z \) where \( \Phi(z) \) is the CDF of a standard Normal \( \mathcal{N}(0, 1) \). The bound of this approximation is provided by the Berry-Essen theorem as \( |F_n(z) - \Phi(z)| = O\left(\frac{1}{\sqrt{n}}\right) \) uniformly in \( z \) when \( Y_n \) has three finite moments [17]. When this approximation is modified by suitable constants and polynomials, \( q_r(\cdot) \), a better bound of the approximation can be derived:

\[
F_n(z) = \Phi(z) + \sum_{r=1}^{k} q_r(z) n^{-r/2} + o(n^{-r/2})
\]
uniformly in \( z \), which is the so called Edgeworth expansion for \( Y_n \) [17] and order \( k \). The coefficients in eq. (4) depend on the moments of \( Y_n \) and, thus, the Edgeworth expansion is especially convenient for our formulation since the central moments with respect to \( \delta_0 \) can be calculated as follows:

First, the \( r \)-th central moment, \( \mu_r = E\{(Y - \delta_0)^r\} \), of a truncated \( \text{nc-\Gamma} \) variate is:

\[
\mu_r = \int_0^t (y - \delta_0)^r e^{-\frac{y - \delta_0}{\beta_0}} \frac{\Gamma(\alpha_0)\beta_0^{\alpha_0}}{\Gamma(\alpha_0 + r)\beta_0^{\alpha_0}} F_X(t, \Theta_0) \, dy
= \frac{\Gamma(\alpha + r)\beta^\alpha}{\Gamma(\alpha)} F_X(t, \Theta_0) - \mu_0^{r+1, \beta_0} \frac{\Gamma(\alpha + r)\beta^\alpha}{\Gamma(\alpha)} F_X(t, \Theta_0).
\]

Note that these moments converge to the ones of the conventional \( \text{nc-\Gamma} \) as \( t \to \infty \).

Now the moments of the sample mean with respect \( \delta_0 \), \( m_r = E\left\{ \frac{1}{n} \sum (y_i - \delta_0)^r \right\} \), can be exactly calculated by the theorem of moments of sample variates [18] as:

\[
m_r = \frac{1}{n^r} \sum_{s=1}^{r} \binom{r}{s} \sigma(p)^{s} \prod_{j=1}^{s} \mu_{E_{j,1}}
\]
where \( \mathcal{P}_s \) is the set of solutions of the integer partition problem for \( n \) into \( s \) terms, \( \sigma(p) \) is a multinomial coefficient, and \( E_{j,1} \) accounts for the permutations without repetition of \( s \) elements with multiplicities the elements of \( p \) (see [18] for more details).

Finally, both the PDF and CDF of the typified sample mean \( Y_n \) of truncated \( \text{nc-\Gamma} \) variates is derived by the Edgeworth expansion as:

\[
F_n(z) = \Phi(z) - \Phi(z) \sum_{r=1}^{k} \frac{h_r(z)}{n^{r/2}}; f_n(z) = \phi(z) \left(1 + \sum_{r=1}^{k} \frac{\tilde{h}_r(z)}{n^{r/2}}\right)
\]
where \( \Phi \) and \( \phi \) are the CDF and PDF of a standard Normal \( \mathcal{N}(0, 1) \) respectively. The polynomials \( h_r, \tilde{h}_r \) can be easily obtained as linear combinations of Hermite Polynomials [17].

From eqs. (7) one can easily obtain both the CDF and PDF of \( Y_n \) by simply applying the transformation \( Y_n = Z_n \sqrt{\frac{m_2 - m_1^2}{m_1}} + m_1 \).

In Fig. 2 we show the PDFs and CDFs of the probability distribution of sample mean of truncated \( \text{nc-\Gamma} \) for an increasing number of samples \( N = n \).

3.3. Significance levels for one-tailed test

With the calculation of the PDF and CDF of the sample mean, we can develop the statistical test for an arbitrary statistical significance. We state the null hypothesis \( H_0 \) as “The truncated mean of samples below the threshold \( t \) come from normal parenchyma”, whereas the alternative hypothesis, \( H_1 \) is “The truncated mean of samples below the threshold come from abnormal parenchyma”.

In order to find the critical values of the test statistic, one can compare the CDF of the truncated \( \text{nc-\Gamma} \) to the CDF of the test statistic, such as the standard Normal distribution.
from emphysema”. In this context, the threshold accounting
for accepting/rejecting the test can be directly calculated from
de CDF obtained by the Edgeworth expansion for the desired
significance. In Fig. 3 the acceptance threshold for different
significance values ($t_{0.05, n}$, $t_{0.025, n}$ and $t_{0.001, n}$) is shown for
an increasing number of samples.

The threshold was exactly calculated for $N \leq 20$, while
the Gaussian approximation is used for $N > 20$ since the
sample distribution quickly converges to the Gaussian distri-
bution for higher number of samples.

4. PERFORMANCE TEST

In this section we test the performance of the hypothesis test
by comparing the conventional threshold of $-950$ HU to the
proposed test. For this purpose, we consider a dataset of
1337 lung regions of interest (ROIs) of size $24.18 \times 24.18$
mm² segmented and classified by a pulmonologist as nor-
mal parenchyma (NT) and emphysematous tissue including
paraseptal (PS), panlobular (PL), emphysema and centrilob-
ular for three different severity levels (CL1-3) from mild to
severe distributed as: 370 NT, 184 PS, 148 PL, 170 CL1, 287
CL2, 178 CL3. These ROIS were also acquired from 16 sites
hosting CT scanners by 3 vendors and a total of 9 specific
models in subjects with COPD. The classification follows a
leave-one-out methodology as explained en sections 2 and 3.

The test was performed to all the ROIs with a statistical
significance 0.05, while the conventional approach assumed a
threshold for $-950$ HU considering that the sample has em-
physema 5% of the samples in the ROI are below this thresh-
old. The classification performance is shown in table 1 where
the superiority of the proposed approach is clear, showing
a good balance between sensitivity (97.64%) and specificity
(85.93%) compared to the conventional approach. This sup-
poses an increase of 7% of the accuracy and an increase of
16% in the specificity, confirming the suitability of the prob-
abilistic model of emphysema/parenchyma and the statistic

5. CONCLUSION

In this paper we propose a statistical test to evaluate the ex-
istence of emphysematous tissue in the secondary pulmonary
lobe. This is achieved by proposing a family of statistical dis-
tributions (the nc-$\Gamma$ distribution) to describe the attenuation of
conventional tissue and airspaces caused by panlobular em-
physema. This family of distributions accurately describes
the statistical behavior of both kinds of tissues and leads to
the definition of an optimal threshold for considering samples
in the secondary pulmonary lobe as potential emphysema.

A statistical test is then proposed by defining a statistic
that accounts for the average attenuation of samples under
the optimal threshold that differentiates between normal tis-
sue and panlobular tissue. This statistic is modeled as the
sample mean of truncated nc-$\Gamma$ distributions. The PDF and
CDF of the statistic are derived and the statistic threshold is
obtained for different confidence intervals.

Interestingly, the proposed statistic shows an asymptotic
convergence to $-951.8$ HU, which is very close to the one
proposed for density masking based on histological analy-
sis of lung lobes[9]. This threshold is optimal in terms of
the statistic for large regions ($n \rightarrow \infty$). However, when a
smaller region of the lung is considered, our model suggests
that lower thresholds ($\leq -950$) should be used to maintain
the same significance level.

The classification was evaluated in a dataset from a mul-
ticenter study with 1,337 samples from different CT scanner
models and vendors. We chose this dataset to provide a
general evaluation of the statistical test. The accuracy
reached was 94.09% with a sensitivity 97.64% and Speci-
ficity 85.93%. The statistical test supposes an increment of
7.5% in total accuracy and 16% in specificity when compared
to the density masking approach.

The model proposed in this study can be combined to
other approaches such as Hidden Markov models to take ad-
vantage of the local statistics with spatial coherence. For this
purpose, the family of distributions purposes may serve to
pave the way to well grounded potentials of the Random Field
definition. Our future work will focus in this direction. Alter-
atively, this work opens the door to propose additional test
statistics that consider the severity of disease in terms of the
percentage of secondary lobule that is affected.
6. REFERENCES


