Role of quantitative computed tomography texture analysis in the differentiation of primary lung cancer and granulomatous nodules

Carole Dennie¹, Rebecca Thornhill¹, Vineeta Sethi-Virmani¹, Carolina A. Souza¹, Hamid Bayanati¹, Ashish Gupta¹, Donna Maziak²

¹Department of Medical Imaging, ²Department of Surgery, Division of Thoracic Surgery, The Ottawa Hospital, University of Ottawa, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

Background: Texture analysis is a computer tool that enables quantification of gray-level patterns, pixel interrelationships, and spectral properties of an image. It can enhance visual methods of image analysis. Primary lung cancer and granulomatous nodules have identical CT imaging features. The purpose of this study was to assess the sensitivity and specificity of CT texture analysis in differentiating lung cancer and granulomas.

Methods: This retrospective study evaluated 55 patients with primary lung cancer and granulomatous nodules who had contrast-enhanced (CE) and/or non-contrast-enhanced (NCE) CT within 3 months of biopsy. Textural features were extracted from 61 nodules. Mann-Whitney U tests were used to compare values for nodules. Receiver operating characteristic (ROC) curves were constructed and the area under the curve (AUC) calculated with histopathology as outcome. Combinations of features were entered as predictors in logistic regression models and optimal threshold criteria were used to estimate sensitivity and specificity.

Results: The model generated by sum of squares, sum difference, and sum entropy features for NCE CT yielded 88% sensitivity and 92% specificity (AUC =0.90±0.06, P<0.0001). For nodules with fluorodeoxyglucose positron emission tomography (FDG-PET)/CT, sensitivity for detection of lung cancer was 79.2% (CI: 57.8–92.9%), specificity was 38.5% (CI: 13.9–68.4%) and accuracy was 64.8%.

Conclusions: Quantitative CT texture analysis has the potential to differentiate primary lung cancer and granulomatous lesions.

Keywords: Applied imaging technology; chest imaging; nuclear imaging; oncologic imaging

Submitted Dec 20, 2015. Accepted for publication Jan 06, 2016.
doi: 10.3978/j.issn.2223-4292.2016.02.01
View this article at: http://dx.doi.org/10.3978/j.issn.2223-4292.2016.02.01

Introduction

A solitary pulmonary nodule (SPN) is defined as a rounded opacity measuring up to 3.0 cm and surrounded by lung parenchyma (1). With the increasing use of multidetector computed tomography (MDCT) to investigate respiratory symptoms, the detection of a solitary nodule is a frequent incidental finding on chest imaging. In the CT screening setting of high-risk populations, up to 75% of screened individuals have a nodule detected over a multiyear screening process (2). At least one nodule is detected in up to 51% of initial screening CT’s (3) and approximately 20% of these nodules will require further investigation (4,5) even though over 95% of these will ultimately prove to be benign on the basis of further imaging, long-term follow-up, or invasive procedures (4–6).

Conventional visual methods fail to differentiate malignant from benign nodules in many cases. The radiologist generally relies on CT imaging features such as the size, shape, contour and attenuation of the nodule to
classify the majority into benign or malignant categories (7). Despite this, a significant proportion remains indeterminate requiring follow-up or triggering invasive diagnostic investigations (8). Invasive procedures have limitations including cost, training expertise and potential for serious complications (3,8).

The introduction of nuclear [18F]-fluorodeoxyglucose positron emission tomography (FDG-PET) imaging has enabled the metabolic assessment of lung nodules and may facilitate differentiating malignant lesions. Although FDG-PET/CT has a very high sensitivity for the detection of malignant lung nodules measuring over 8 mm, it has important limitations. Specifically uptake of glucose-analogue FDG in both malignant and benign inflammatory lesions increases the potential for false positive results (9). Granulomatous lesions especially necrotizing granulomas, endemic in many parts of North America, are often FDG-avid on PET/CT and can be especially problematic for radiologists as they appear as spiculated or lobulated solid lesions without a sub-solid or ground-glass component on CT and therefore mimic invasive carcinomas of the lung (10).

In the era of low-dose CT screening for lung cancer, the search for non-invasive tools to differentiate benign from malignant nodules has intensified. Computer tools for automated pattern recognition and image analysis can provide objective information to support clinical decision-making. Texture analysis is one such tool that uses a wide range of techniques that enable quantification of the gray-level patterns, pixel interrelationships, and spectral properties of an image (11-13). The role of texture analysis in CT and magnetic resonance imaging (MRI) has been explored in the brain (11,12), lung (14,15), breast (16), prostate (17), skeletal muscles (18), liver (19,20) kidney (21,22) and bones (23).

As lung nodules are complex in structure and contain microscopic textural features imperceptible to the human eye, we sought to assess the value of texture analysis in differentiating primary lung cancer and benign granulomatous lesions. We specifically focused on granulomatous lesions as these are solid and spiculated or lobulated on CT and prove the most challenging to differentiate from invasive lung cancer based on visual characteristics alone.

**Materials and methods**

This was a retrospective case-control study approved by our Research Ethics Board. The need for informed consent was waived. Patients were identified by searching the pathology database from a single academic tertiary care institution for pathologically proven granulomatous lesions and solid primary lung cancers measuring between 1–3 cm identified between March 2010 and April 2013. The histopathologic analyses were performed by pathologists within the same academic institution with subspecialty training in thoracic pathology. Included patients had a definitive diagnosis obtained by image-guided biopsy with cytological and/or histological analysis and had a contrast-enhanced (CE) and/or non-contrast-enhanced (NCE) CT of the thorax performed within 3 months of the biopsy. The lung nodules were all solid and spiculated or lobulated, measured ≤ 3 cm and did not display calcification or fat.

CT studies were obtained on one of three 64-detector scanners (Toshiba Aquilion-64, Toshiba Medical Systems, Otawara, Japan, LightSpeed VCT and Discovery 750 HD, General Electric, Milwaukee, Wisconsin, USA) with a breath-held helical acquisition of the entire thorax, 120 kV, 100–200 mAs, pitch 0.75–1.0 and collimation of 0.5 or 0.625 mm.

All imaging data were reconstructed with a medium-sharp reconstruction algorithm and a slice thickness of 0.63–5.0 mm, FOV 500 mm, matrix 512x512 mm. CE CTs were performed using 40 cc Omnipaque 300 (GE Healthcare, Milwaukee, WI, USA) at a rate of 2 cc/sec and a delay of 25–30 seconds. CT images were reviewed using a picture archiving and communication system (PACS) (Horizon Medical Imaging, McKesson Corporation, San Francisco, CA, USA) on lung window settings (W: 1,500, L: −550).

We divided the nodules into two categories based on whether the CT had been performed with or without intravenous (IV) contrast.

FDG-PET studies were all performed on the same system (Philips Gemini Dual Exp PET/CT scanner), using the same protocol: z-axis image from the base of the skull to mid-thigh, 60 minutes after IV injection of 6 MBq/kg of FDG. All patients had blood glucose <8 mmol/L and were fasting for over 8 hours. Low mA, 6.0 mm thick CT images were acquired for attenuation correction and a Hermes Hybrid Viewer® workstation was used for imaging interpretation. A standardized uptake value (SUV) >2.5 was considered a PET-positive result.

**Image analysis**

An axial CT image at the level of the largest diameter of the nodule was selected for texture analysis by a thoracic radiologist (VS) with 8 years of cross-sectional imaging
experience. A contour was manually drawn along the outer margin of the nodule using a polygonal region of interest (ROI) tool and exported to an independent workstation and de-identified. Each nodule was manually segmented by the same radiologist using Image J® (National Institutes of Health, USA, http://rsbweb.nih.gov) (Figure 1).

Prior to analysis, CT image intensities were normalized between \( \mu \pm 3\sigma \) where \( \mu \) was the mean value of gray levels inside the ROI and \( \mu \) the standard deviation. Gray levels between \([\mu - 3\sigma]\) and \([\mu + 3\sigma]\) were then decimated to 64 gray levels. This normalization procedure has been shown to minimize inter-scanner effects in MRI texture analysis (24) and is presumed to also reduce inter-scanner effects in other modalities such as CT. Gray-level co-occurrence matrix (GLCM) features were extracted using MaZda version 4.6 (25) by a physicist (RT) blinded to the final histopathologic diagnosis of each nodule. We selected textural features related to linear gray-level dependencies (“f3,” or correlation), the sum of squares (“f4”), the sum variance (“f7”), the extent of gray-level randomness (“f8,” sum entropy), as well as difference variance (“f10”) (11). These are textural features that have previously demonstrated potential for identifying malignant renal (21,22) and breast tumors (26). Definitions and formulas for all textural features are provided in Table 1.

**Statistical analysis**

Mann-Whitney U tests were used to compare the value of each textural feature to differentiate granulomatous nodules from primary lung cancers (MedCalc Software version 12.1.4.0, Mariakerke/Belgium). We constructed receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) for each textural feature, with histopathologic diagnosis as outcome. The standard error (SE) of the area under the ROC curve (AUC) was evaluated according to the method of De Long (27). We subsequently entered combinations of textural features as predictors in logistic regression models and computed ROC curves. The same features were used to train multiple support vector machine (SVM) classifiers. A SVM with a non-linear radial basis function was implemented in The Unscrambler® X (v.10.1, CAMO Software), which uses the LIBSVM library (28). For SVM classification, the principle is to construct a hyperplane that best separates the data points into benign and malignant cases. The hyperplane is then applied to classify new cases. We evaluated classification generalizability and performance by means of 10-fold cross-validation procedures (29).

**Results**

A total of 55 patients met inclusion criteria (20 men, 35 women). The mean age for the 24 patients with granulomatous lesions was 67.3±9.7 years (range, 46–86 years) and 67.8±13.1 years (range, 46–89 years) (P=0.934) for the 31 patients with lung cancer. The granulomatous lesions had a mean diameter of 1.7±0.5 cm (range, 1.0–2.9 cm) compared
to 2.0±0.6 cm (range, 1.0–2.9 cm) for the 31 lung cancers (P=0.041). There were 17 necrotizing granulomas and seven non-necrotizing granulomatous lesions. There were 31 patients with lung cancer (25 adenocarcinomas, 4 squamous cell cancers and 1 NSCLC not specified) and 1 small cell lung cancer.

Six patients had both CE and NCE CT of the same nodule (3 granulomas and 3 lung cancers). This generated 61 CT images with a nodule for analysis (27 granulomas and 34 lung cancers).

Thirty-seven nodules were evaluated with FDG-PET/CT (13 granulomas and 24 lung cancers) within 3 months of the biopsy.

Medians and interquartile ranges for GLCM correlation (f3), sum of squares (f4), sum variance (f7), sum entropy (f8), and difference variance (f10) are depicted for both CE and NCE CT in Figure 2. For CE CT, there were no significant differences between granulomatous lesions and lung cancers for any of the individual textural features examined (P>0.05 for each). For NCE CT, the median f4, f7, and f8 were significantly greater in the lung cancer group compared to the granuloma group [median (IQR) f4 =17.6 (14.2–32.5) vs. f4 =10.4 (8.7–15.3), P=0.046; f7 =44.6 (33.4–80.6) vs. f7 =25.6 (18.7–46.0), P=0.03; f8 =1.35 (1.31–1.43) vs. f8 =1.24 (1.15–1.30), P=0.005]. Trends toward increased f3 and f10 in the NCE lung cancer group were also revealed (f3 =0.28 (0.15–0.45) vs. f3 =0.17 (0.01–0.30), P=0.10; f10 =10.07 (8.30–21.89) vs. f10 =6.21 (4.58–9.10), P=0.07).

The area under the ROC curve for each feature, as well as for combinations of features, is provided in Tables 1 and 2, respectively. For CE CT, the AUC was not significantly greater than 0.5 for any of the individual features. Conversely, the AUCs corresponding to the f4, f7, and f8 extracted from NCE images were significantly greater than 0.5 (P<0.05 for each). The ROC curves corresponding to each combination of features for both CE and NCE CT are provided in Figures 3 and 4. The logistic regression model produced from combining all five textural features yielded a sensitivity of 81% and specificity of 79% for identifying primary lung cancers (AUC =0.79±0.11, P=0.002) for CE CT. The model generated by the same five features using NCE images correctly identified primary lung cancers with 75% sensitivity and 100% specificity (AUC =0.91±0.06, P<0.0001). However, when only three features were combined for CE CT (f4 + f7 + f8), the logistic regression model demonstrated 38% sensitivity and 86% specificity (AUC =0.60±0.11, P=0.34; Figure 3), whereas the model generated by this same combination of features for NCE CT yielded 88% sensitivity and 92% specificity (AUC =0.90±0.06, P<0.0001; Figure 4). It was not possible to assess the significance of differences in the AUC between CE and NCE CT, as the number of patients who had nodules assessed using both techniques was too small.

The results of each 10-fold cross-validation procedure

<p>| Table 1 Individual texture features for CE and unenhanced CT |
|---------------------------------|---------------------------------|---------------------------------|</p>
<table>
<thead>
<tr>
<th>Feature</th>
<th>CE CT AUC (SE)</th>
<th>Criterion Se (%)</th>
<th>Sp (%)</th>
<th>CE CT AUC (SE)</th>
<th>Criterion Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>f3: Correlation</td>
<td>0.51 (0.11)</td>
<td>0.94 ≤0.33</td>
<td>56</td>
<td>0.68 (0.11)</td>
<td>0.09 &gt;0.19</td>
<td>69</td>
</tr>
<tr>
<td>f4: Sum of squares</td>
<td>0.51 (0.11)</td>
<td>0.94 ≤27</td>
<td>88</td>
<td>0.72 (0.11)</td>
<td>0.048 &gt;11.3</td>
<td>88</td>
</tr>
<tr>
<td>f7: Sum variance</td>
<td>0.50 (0.11)</td>
<td>0.97 &gt;33</td>
<td>69</td>
<td>0.74 (0.11)</td>
<td>0.03 &gt;23.5</td>
<td>100</td>
</tr>
<tr>
<td>f8: Sum entropy</td>
<td>0.54 (0.11)</td>
<td>0.75 &gt;1.42</td>
<td>38</td>
<td>0.81 (0.11)</td>
<td>0.003 &gt;1.28</td>
<td>94</td>
</tr>
<tr>
<td>f10: Difference variance</td>
<td>0.52 (0.11)</td>
<td>0.84 &gt;9.01</td>
<td>63</td>
<td>0.70 (0.12)</td>
<td>0.08 &gt;7.64</td>
<td>81</td>
</tr>
</tbody>
</table>

*where P(i, j) indicates the joint probability of two pixels having particular co-occurring values i, j = 1, 2, ..., Ng, R indicates the total number of neighboring pixel pairs, and Ng the number of distinct gray levels. μx, μy, σx, and σy indicate means and standard deviations of the row and column sums of the co-occurrence matrix (11). CE, contrast-enhanced; Se, sensitivity; Sp, specificity; SE, standard error.
are also summarized in Table 2 (SVM accuracy). Cross-validation of the SVM classifier created by combining all five textural features extracted from CE CTs indicated an accuracy of 60%, whereas the classifier generated by the same features extracted from NCE CTs yielded an accuracy of 64%. When only three or four features were combined for CE CTs, the average SVM accuracy was 63% for $f_4$, $f_7$, and $f_8$, and 60% for the combination of $f_3$, $f_4$, $f_7$, $f_8$, and $f_{10}$. The average SVM accuracy yielded by either of these combinations for NCE CTs was 64%.

Of the 13 granulomatous lesions evaluated with FDG-PET/CT, 8 (61.5%) had a false positive result. Of the 24 lung cancers evaluated with FDG-PET/CT, 19 (79.1%) had a true positive result. Overall, in the subgroup of nodules evaluated with FDG-PET/CT, sensitivity for the detection of primary lung cancer was 79.2% (CI: 57.8–92.9%), specificity was 38.5% (CI: 13.9–68.4%) and accuracy was 64.8%.

**Discussion**

In our study, we evaluated the role of texture analysis, an applied imaging technology, in differentiating granulomatous

---

**Figure 2** Box and whisker plots depicting medians, interquartile ranges and extrema for GLCM correlation ($f_3$, A), sum of squares ($f_4$, B), sum variance ($f_7$, C), sum entropy ($f_8$, D), and difference variance ($f_{10}$, E) extracted from contrast- and non-contrast-enhanced CT images of benign (granulomatous) and malignant (lung cancer) nodules. Comparisons between group medians were assessed using two-tailed Mann-Whitney U tests. GLCM, gray-level co-occurrence matrix.
lesions from invasive lung cancers. We also analyzed and compared the results of this technique on CE and NCE CT. The model generated by using five textural features on NCE CT identified primary lung cancers with a sensitivity of 75% and specificity of 100% (AUC = 0.90, P < 0.0001) and an accuracy of 79%. The three textural features found to be most relevant for lesion characterization were sum of squares, sum difference, and sum entropy. Using only these three features, sensitivity of texture analysis on NCE CT increased to 88% with a small decrease in specificity (92%). While both sum of squares and sum variance features reflect the degree of deviation from the mean gray level present in the nodule (12,30), sum entropy reflects the degree of gray level randomness in the texture. If there is no predictable gray level pattern, then the lesion has high entropy, whereas if there are repeating or predictable patterns the lesion has

<table>
<thead>
<tr>
<th>Model</th>
<th>CE CT</th>
<th></th>
<th>Unenhanced CT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC</td>
<td>SE (AUC)</td>
<td>P</td>
<td>Se (%)</td>
</tr>
<tr>
<td>f7, f8</td>
<td>0.60</td>
<td>0.11</td>
<td>0.36</td>
<td>38</td>
</tr>
<tr>
<td>f4, f7, f8</td>
<td>0.60</td>
<td>0.10</td>
<td>0.34</td>
<td>38</td>
</tr>
<tr>
<td>f4, f7, f8, f10</td>
<td>0.71</td>
<td>0.10</td>
<td>0.03</td>
<td>63</td>
</tr>
<tr>
<td>f3, f4, f7, f8, f10</td>
<td>0.79</td>
<td>0.09</td>
<td>0.002</td>
<td>81</td>
</tr>
</tbody>
</table>

CE, contrast-enhanced; AUC, area under the curve; Se, sensitivity; Sp, specificity; SVM, support vector machine; SE, standard error.

Figure 3 Receiver operating characteristic (ROC) curves (solid lines) and 95% confidence intervals (dotted lines) corresponding to non-contrast enhanced CT texture features for the identification of malignant (lung cancer) nodules. The area under the ROC curve (AUC) for each of (A-D) is provided in Table 2. AUC, area under the curve.
The entropy-related textural feature was significantly higher in primary lung cancers, presumably reflecting the more complex and more inhomogeneous internal structure of malignant lesions when compared to granulomatous lesions. Lesion heterogeneity is a known feature of malignancy, likely related to abnormal tumor angiogenesis and cellular infiltration (31,32).

The results obtained in our study are comparable to previously published texture analysis studies in lung neoplasms. Wang et al. studied texture analysis on NCE CT in 2,171 benign and malignant nodules in 185 patients and found entropy and sum entropy to be different in malignant and benign pulmonary nodules (P<0.05) (33). High entropy in malignant breast lesions has also been demonstrated in a recent study by Gibbs et al. using MRI texture analysis. In this study, diagnostic accuracy was 0.81±0.07 in the differentiation of malignant from benign breast lesions (16).

Cavouras et al. looked at 51 histologically confirmed nodules on NCE CT using texture analysis combined with CT density matrix analysis. Overall classification accuracy was 90.2% in distinguishing between benign and malignant lung nodules (34). Finally, McNitt-Gray et al. looked at 32 lung nodules (14 benign, 14 malignant) on NCE CT and four textural features yielded an area under the ROC curve (AUC) of 0.992. Sum entropy was one of these features (15).

Using the same three textural features on CE CT as those used for NCE CT, lung cancers were identified with a much lower sensitivity (38% vs. 88%). The reason for better demonstration of texture on NCE CT is not clear. Because the majority of the CE and NCE CTs in our study were obtained in different patients, we cannot rule out the possibility that variations in textural features might reflect different tumor characteristics rather than a direct effect of IV contrast. However, an MR study of benign cystadenomas and borderline tumors by Dujardin et al. had similar results.

Figure 4 Receiver operating characteristic (ROC) curves (solid lines) and 95% confidence intervals (dotted lines) corresponding to contrast enhanced CT texture features for the identification of malignant (lung cancer) nodules. The area under the ROC curve (AUC) for each of (A-D) is provided in Table 2. AUC, area under the curve.
and the authors postulated that the presence of contrast may obscure textural differences between benign and malignant lesions (35). Chae et al. specifically excluded CE CT in their study of texture analysis in the differentiation of pre-invasive from invasive lung adenocarcinomas but did not state the reason for this exclusion (36). Ganeshan et al. also only included NE CT in their study of texture in non-small cell lung cancer and the relationship with glucose metabolism (37). In contrast, a recent publication by the same authors showed good correlation between histological markers and texture in non-small cell lung cancer on both CE and NCE CT performed in the same patients (31). Benign nodules were not included in their study.

In our cohort, texture analysis on NCE CT was more sensitive, specific and accurate than FDG-PET/CT in differentiating primary lung cancers from granulomatous lesions. In the subgroup of nodules evaluated with FDG-PET/CT, the sensitivity for the detection of primary lung cancer was 79.2% with a specificity of 38.5% and accuracy of 64.8%. Therefore, texture analysis may be superior to PET particularly in the differentiation of primary lung cancer from a granulomatous lesion.

Our study has several limitations. It was retrospective with a small number of patients, therefore with inherent selection bias. CT examinations were performed using different technical parameters and this may have resulted in variability in CT attenuation with a potential effect on the estimation of texture features. However, this limitation should have been mitigated by the normalization procedure undertaken prior to texture analysis outlined above. In addition, a minority of patients had nodules assessed with both CE and NCE CT precluding definite conclusions on whether IV contrast affects CT texture analysis. This requires further investigation. FDG-PET/CT was also performed in only 37 nodules. Lastly, texture features of each nodule were derived from manual segmentation performed by a single radiologist and this may have been influenced by subjective bias on the part of the observer. However, manual segmentation is the current reference and previous texture analysis studies have reported a high degree of intra- and inter-observer reproducibility (21,38).

**Conclusions**

Our results suggest that quantitative CT texture analysis has the potential to differentiate primary lung cancer from granulomatous lesions. It performed better on NCE CT than CE CT and may also have an advantage over FDG-PET/CT. Texture analysis is not expected to replace tissue diagnosis but may be added to the radiologist’s arsenal used to characterize spiculated or lobulated solid lung nodules especially in areas where granulomatous infections are endemic. It may also be particularly relevant in patients who are poor surgical candidates or at risk for invasive diagnostic procedures when textural features suggest a benign etiology. In these situations, clinicians may choose observation rather than proceeding to more invasive options. Although encouraging, the usefulness of textural analysis in the differentiation of granulomatous lesions and primary lung cancer needs to be validated in a larger population. In addition, a head to head comparison of textural features on CE and NCE CT’s in the same patients would be beneficial.

**Acknowledgements**

None.

**Footnote**

Conflicts of Interest: The authors have no conflicts of interest to declare.

**References**

5. Swensen SJ, Jett JR, Hartman TE, Midthun DE,


28. Chang CC, Lin CJ. LIBSVM: a library for support vector


