Do computer extracted shape and texture features enable discrimination of granulomas from adenocarcinoma on chest computed tomography?

Christina Williamson

Department of Thoracic and Cardiovascular Surgery, Lahey Hospital and Medical Center, Burlington, Massachusetts, USA
Correspondence to: Christina Williamson, MD. Department of Thoracic and Cardiovascular Surgery, Lahey Hospital and Medical Center, Burlington, Massachusetts, USA. Email: christina.williamson@lahey.org.

Comment on: Orooji M, Alilou M, Rakshit, *et al*. Combination of computer extracted shape and texture features enables discrimination of granulomas from adenocarcinoma on chest computed tomography. J Med Imaging (Bellingham) 2018;5:024501.

Lung cancer screening with low dose computed tomography (CT) performed on multidetector helical CT scanners and the use of high-resolution chest CT scans to evaluate respiratory symptoms have resulted in an increase in the discovery of indeterminate pulmonary nodules. Conventional visual methods used by radiologists including shape, size, contour, density, and attenuation have a relatively high sensitivity (0.89) and moderate specificity (0.70) in differentiating benign from malignant nodules in a large meta-analysis (1). The use of [18F]-fluorodeoxyglucose positron emission tomography (FDG-PET) has further facilitated the differentiation of benign from malignant nodules by enabling metabolic assessment. However, FDG-PET scans although felt to be highly sensitive in detecting malignancy in nodules over 8 mm have a high false positive rate particularly in benign inflammatory lesions such as granulomas (2). Furthermore, normal PET/CT scans are not found to be reliable indicators of the absence of malignancy in patients with a high probability of lung cancer (3). A confounding factor is the regional variation in fungal disease in the U.S. that gives rise to granulomas that can be indistinguishable from lung cancers on both chest CT scan and FDG-PET scan. Aggressive tissue acquisition is often required via bronchoscopy, navigational bronchoscopy, CT guided biopsy, or surgical excisional biopsy to confirm the diagnosis of malignancy especially in high risk individuals.

The use of computer learning or artificial intelligence (AI) for image analysis of chest CT scans has the potential to provide additional objective and discriminating information to support clinical decision making. The use of computer tools for automated pattern recognition and image analysis, particularly texture analysis, is not a new concept (4). It has been investigated in CT and magnetic resonance imaging (MRI) in the lung and other organs including brain, breast, prostate, kidney, bone, and liver (5-9). Texture features found on chest CT images for prediction of malignant pulmonary nodules has been reported (10). The sensitivity and specificity of texture analysis has been studied in the differentiation of primary lung cancer and granulomatous nodules as well (11).

The article by Orooji and his colleagues is posing a different question as to whether texture and shape analysis can discriminate between a subset of non-small cell lung cancer namely adenocarcinoma and granuloma (12). Their approach is unique in that they incorporate both shape and texture features in their development of a training set of 139 pulmonary nodules (70 adenocarcinomas and 69 granulomas). They then constructed machine learning classifiers using the shape and texture features that were robust in feature discriminability to distinguish the two pathologies. They also investigated the sensitivity of radiomic feature expression across different acquisition sites and scanners as well as variations in slice thickness. The most discriminating subset of features was culled by sequential forward feature selection. The top six features were used to train and lock down linear discriminant analysis (LDA), quadratic discriminant analysis (QDA),

as well as support vector machine (SVM) classifiers. These three classifiers were then used on the training set of 56 pulmonary nodules (34 adenocarcinomas and 22 granulomas) to assign a probability of being an adenocarcinoma. The performance of each of these classifiers was individually evaluated using the area under the receiver operating curve (AUC). The best AUC corresponded to the SMV classifier on the training set.

Texture features outperformed shape features in discriminating adenocarcinoma from granuloma due to informative heterogeneity in adenocarcinoma. Malignant lesions tend to demonstrate greater heterogeneity due to more chaotic microarchitecture likely related to abnormal tumor angiogenesis and cellular infiltration (13). However, shape features, particularly convexity, eccentricity, and extend features were still strongly discriminative of granulomas and adenocarcinoma in their study. The inclusion of shape along with texture features improved the predictive performance of the SVM classifier with an AUC of 92.9% on the training set and 77.8% on the validation set. Their model yielded a positive predictive value of 72% and a 0% false negative rate.

The major advantage of this study is the incorporation of both shape and texture features in the classifier and radiomic analysis. This is perhaps the only report of using both parameters. The authors acknowledge that including data from only two institutions raises that question of whether their classifier is as robust when utilized in multiple different sites. They controlled very nicely for the variable of manual segmentation across the region of interest (RoI). Two experienced thoracic radiologists re-segmented the same RoI in a randomly picked subset of 10 adenocarcinomas and 10 granulomas and were blinded to results of the segmentation of each reader. It is questionable whether this process can be reproduced and would be as robust at multiple different sites. I also found it curious that when the authors swapped the training and testing sets the QDA classifier became the top ranked classifier compared to the LDA and SVM classifiers with an AUC of 82.5% on the validation set. I would suspect that this may be a result of the tyranny of small numbers with the testing set.

One of the radiographic appearances of an adenocarcinoma discovered on chest CT scan is as a part solid and part ground glass opacity or GGO nodule. As this study only evaluated solid pulmonary nodules it does not address this particular presentation of adenocarcinoma of the lung. These part solid nodules are less likely to be confused with granulomatous disease, however.

The major flaw in this study was the original design that included only adenocarcinomas and granulomas in their study cohort. Other pulmonary malignancies such as squamous cell, large cell, and small cell carcinomas as well as carcinoid tumors not to mention other benign conditions can present as indeterminate pulmonary nodules and would be in the clinician's differential. The study by Dennie *et al.* included other lung cancers besides adenocarcinoma and found an 88% sensitivity and 92% specificity for detection of lung cancer using quantitative CT texture analysis (11).

This is a retrospective study and would also need to be validated across multiple centers in a prospective fashion to be of clinical usefulness. It is interesting to note that when comparing the performance of the classifier with an experienced thoracic radiologist as well as a pulmonary fellow the classifier only marginally outperformed the two readers.

I commend Orooji and his colleagues for the technical and scientific rigor with which their study was performed. It undoubtedly refines our understanding of pattern recognition and image analysis particularly with respect to texture and shape features of pulmonary nodules found on chest CT scans.

For AI to be clinically useful in informing our decision making in managing indeterminate pulmonary nodules it would need to accurately refine our current risk stratification models. It would have to better define those nodules that require aggressive invasive tissue diagnosis that are suspicious for lung cancer and those that merit further image monitoring. Published and validated clinical lung cancer prediction models include the Mayo Clinic, Veterans Affairs (VA), solitary pulmonary nodule (SPN), and Thoracic Research Evaluation and Treatment (TREAT) models. Existing predictive models for lung cancer incorporate demographic, clinical, and radiologic features to better inform clinicians regarding referral for invasive biopsy (14-18). The TREAT model in particular had an AUC 0.89 (95% CI: 0.79–0.92) in a validation dataset and appears to have better diagnostic accuracy than the Mayo Clinic model (18).

The use of AI for preoperative diagnosis of pulmonary nodules has been studied in surgical patients undergoing resection for suspected pulmonary malignancy. Using preoperative clinical and radiographic parameters, 100 consecutive patients were prospectively evaluated with a diagnostic model that predicted whether the lesion was benign or malignant. The computer prediction agreed with the final tissue diagnosis in 95 of 100 patients

and demonstrated a sensitivity of 96% and specificity of 89% (19). These were patients being operated on for suspected cancer and over half had tumors greater than 3 cm. This is not representative of the typical patient with an indeterminate pulmonary nodule. Computer learning and the application of artificial intelligence can be incorporated into lung cancer prediction models by providing more refined automated pattern recognition and image analysis of chest CT scans. Gene profiling via nasal swabs as well as proteomics are currently being evaluated and could also be combined with robust image analysis using AI to refine clinical decision making and be even more selective in recommending invasive tissue sampling (20-22).

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Journal of Medical Artificial Intelligence*. The article did not undergo external peer review.

Conflicts of Interest: The author has completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jmai.2018.08.03). The author has no conflicts of interest to declare.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Zhang CY, Yu HL, Li X, et al. Diagnostic value of computed tomography scanning in differentiating malignant from benign solitary pulmonary nodules: a meta-analysis. Tumour Biol 2014;35:8551-8.

2. Gould MK, Maclean CC, Kuschner WG, et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a metaanalysis. JAMA 2001;285:914-24.

3. Maiga AW, Deppen SA, Mercaldo SF, et al. Assessment of Fluorodeoxyglucose F18-Labeled Positron Emission Tomography for Diagnosis of High-Risk Lung Nodules. JAMA Surg 2018;153:329-34.

4. Tourassi GD. Journey toward computer-aided diagnosis: role of image texture analysis. Radiology 1999;213:317-20.

5. Sharma N, Ray AK, Sharma S, et al. Segmentation and classification of medical images using texture-primitive features: Application of BAM-type artificial neural network. J Med Phys 2008;33:119-26.

6. Gibbs P, Turnbull LW. Textural analysis of contrastenhanced MR images of the breast. Magn Reson Med 2003;50:92-8.

7. Ganeshan B, Miles KA, Young RC, et al. Texture analysis in non-contrast enhanced CT: impact of malignancy on texture in apparently disease-free areas of the liver. Eur J Radiol 2009;70:101-10.

8. Ganeshan B, Burnand K, Young R, et al. Dynamic contrast-enhanced texture analysis of the liver: initial assessment in colorectal cancer. Invest Radiol 2011;46:160-8.

9. Schieda N, Thornhill RE, Al-Subhi M, et al. Diagnosis of Sarcomatoid Renal Cell Carcinoma With CT: Evaluation by Qualitative Imaging Features and Texture Analysis. AJR Am J Roentgenol 2015;204:1013-23.

10. Wang H, Guo XH, Jia ZW, et al. Multilevel binomial logistic prediction model for malignant pulmonary nodules based on texture features of CT image. Eur J Radiol 2010;74:124-9 .

11. Dennie C, Thornhill R, Sethi-Virmani V, et al. Role of quantitative computed tomography texture analysis in the differentiation of primary lung cancer and granulomatous nodules. Quant Imaging Med Surg 2016;6:6-15.

12. Orooji M, Alilou M, Rakshit S, et al. Combination of computer extracted shape and texture features enables discrimination of granulomas from adenocarcinoma on chest computed tomography. J Med Imaging (Bellingham) 2018;5:024501.

13. Ganeshan B, et al. Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. Radiology 2013;266:326-36.

14. Swensen SJ, Silverstein MD, Ilstrup DM, et al. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med 1997;157:849-55.

15. Schultz EM, Sanders GD, Trotter PR, et al. Validation of two models to estimate the probability of malignancy in patients with solitary pulmonary nodules. Thorax 2008;63:335-41.

16. Gurney JW, Lyddon DM, McKay JA. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. Part II. Application. Radiology 1993;186:415-22.

17. Division of Cancer Control and Population Sciences. Lung Cancer Prediction Models. Available online: http://epi.grants.cancer.gov/cancer_risk_prediction/lung.html

18. Deppen SA, Blume JD, Aldrich MC, et al. Predicting lung cancer prior to surgical resection in patients with lung nodules. J Thorac Oncol 2014;9:1477-84.

19. Edwards FH, Schaefer PS, Cohen AJ, et al. Use of artificial intelligence for the preoperative diagnosis of pulmonary lesions. Ann Thorac Surg 1989;48:556-9.

20. AEGIS Study Team. Shared Gene Expression Alterations in Nasal and Bronchial Epithelium for Lung Cancer Detection. J Natl Cancer Inst 2017;109.

21. Sozzi G, Boeri M, Rossi M, et al. Clinical utility of a plasma-based miRNA signature classifier within computed tomography lung cancer screening: a correlative MILD trial study. J Clin Oncol 2014;32:768-73.

22. Vachani A, Hammoud Z, Springmeyer S, et al. Clinical Utility of a Plasma Protein Classifier for Indeterminate Lung Nodules. Lung 2015;193:1023-7.