

Diffusion-weighted Magnetic Resonance Imaging and Receiver Operating Characteristic (ROC) Analysis for Lung Cancer Diagnosis

Yahui Peng

School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China

Yongxia Lei, Xinchun Li

First Affiliated Hospital of Zhuangzhou Medical School, Zhuangzhou, China

Houjin Chen

School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China

ABSTRACT: Diffusion-weighted magnetic resonance (DWI-MR) technique is a promising imaging method for lung cancer diagnosis. The purpose of this study is to investigate whether DWI-MR can differentiate lung cancer from infectious nodule accurately. Approved by institutional review board, this study included DWI-MR images of 38 patients who had solitary lung nodules in chest CT examination. Three different non-zero b -values, 600, 800, and 1000 s/mm^2 , were used for DWI-MR imaging and apparent diffusion coefficient (ADC) was calculated for lung cancer detection. Histopathological analysis of biopsy or anti-inflammation treatment results were used as the gold standard. Area under the receiver operating characteristic curve (AUC) was used as the figure of merit for diagnostic accuracy assessment. Results indicated that ADC for lung cancer was statistically significantly smaller than that for infectious nodules. The AUC values of ADC were 0.92 ± 0.07 , 0.82 ± 0.09 , and 0.79 ± 0.10 , for b -values of 600, 800, and 1000 s/mm^2 , respectively. We conclude that the DWI-MR imaging can differentiate lung cancer from infectious nodules accurately.

KEYWORD: Lung cancer; magnetic resonance imaging; diffusion-weighted imaging; receiver operating characteristic analysis; diagnostic accuracy assessment.

1 INTRODUCTION

1.1 Lung cancer and magnetic resonance imaging

Lung cancer is a major medical concern in China and globally (1). In recent years, a strong tendency can be observed in China that the incident rate, especially in young people, is gradually but constantly increasing. Because lung cancer can be a fatal disease, accurate detection of the disease, preferably in its early stage, is critically important to reduce the death rate and to improve the prognosis. Commonly used imaging diagnostic techniques include chest X-ray (2), computed tomography (CT) (3), and positron emission tomography (PET) (4). However, each method has limitations and better imaging techniques are still under development. Recently, new magnetic resonance (MR) imaging methods have been applied to lung cancer diagnosis (5). One of the MR imaging technique is diffusion-weighted imaging (DWI). It has been used for cancer detection in breast, prostate, liver, and other organs. But DWI for lung cancer differentiation has not been studied extensively.

1.2 Receiver operating characteristic analysis

Diagnosis of a disease is a binary decision, i.e., clinicians need to differentiate whether or not a certain abnormality exists in a given patient. Obviously, when a wrong decision is made, the consequence can be costly because the action based on the wrong decision may hurt a healthy person or deteriorate the condition of a sick patient. Therefore, it is important to evaluate how well the diagnostic decision can accurately reflect the "true" state of the disease. Receiver operating characteristic (ROC) analysis is such a tool that can help clinicians evaluate the accuracy of their diagnoses of a disease, and many other binary decisions (6, 7). ROC analysis has been widely utilized in radiology, psychology, and other fields. However, in many medical research reports where ROC analysis should be used, this method was presented in an ambiguous way or not always used properly. Besides ROC analysis, other commonly used accuracy assessment methods include percentage of correctness and the pair of sensitivity and specificity. However, these methods have limitations and their usage should be restricted to specific situations where ROC analysis cannot be

easily applied. Differences between these methods and ROC analysis will be discussed briefly later on.

In this paper, we present results of ROC analysis on images of lung cancer with DWI-MR imaging technique, focused more on the analysis of the diagnostic accuracy. The purpose of the paper is to investigate whether DWI-MR imaging technique can detect lung cancer accurately. In addition, using lung cancer detection as an example, we explain how ROC curves should be used for the assessment of diagnostic accuracy.

2 MATERIALS AND METHODS

2.1 Patient population

This prospective study was approved by the institutional review board at the First Affiliated Hospital of Guangzhou Medical School. From September 2013 to February 2014, we collected 38 patients who had solitary lung nodules in the chest CT examination. The average patient age was 58.8 ± 10.93 years. Among them, 30 were diagnosed with lung cancer, confirmed by histopathology, and 8 were benign abnormalities, confirmed by either histopathology (n=5) or by nodule absorption/disappearance after anti-infection treatment (n = 3). MR examination was performed for all the patients before biopsies or treatments. DWI-MR was performed on a 3.0-Tesla Archieva MR scanner (Archieva, Philips Healthcare, Eindhoven, the Netherlands) for each patient.

2.2 MR data analysis

DWI-MR signal intensity reflects the degree of restricted random diffusion motion, a.k.a. Brownian motion, of water molecules inside the human body (8). In a pure water body, a water molecule in the center of the volume may appear in a position far from its original position in a long period of time due to diffusion. However, when in the human body, a water molecule may not be able to move far from its original position in a given time due to the barriers imposed by cells or other structures. When the density of the cells is increased, the diffusion of water molecules is more restricted because there are more barriers. Quantitatively, the degree of restricted diffusion motion of water molecules, defined as apparent diffusion coefficient (ADC), can be calculated as follows,

$$ADC = \frac{\ln S_1 - \ln S_0}{b_0 - b_1} \quad (1)$$

where $\ln S_0$ and $\ln S_1$ are the natural logarithm of the DWI-MR signal intensities at the diffusion weighting factors b_0 and b_1 , respectively (9). The "b-value" represents the diffusion weighting, which

can be adjusted before the MR scan is performed. Because cancerous tissue often has increased cell density, the calculated ADC map can be very useful in detecting the cancer regions since the restricted diffusion areas normally have hypointensive signals.

In radiological practice, b_0 is often set to be zero so only one non-zero b-value is needed to calculate ADC map. When more than two b-values are used, the linear least-squares method can be used to calculate ADC values. In this study, three non-zero b-values were used to acquire DWI-MR images, including 600, 800, and 1000 s/mm².

Table 1. Two-way contingency table for the calculation of true positive fraction (TPF) or sensitivity and false positive fraction (FPF) or 1-specificity

Classification Result	True state		Subtotal
	Disease*	No disease [§]	
Sick	A	B	A+B
Healthy	C	D	C+D
Subtotal	A+C	B+D	A+B+C+D

* True positive fraction (TPF) = A/(A+C).

§ False positive fraction (FPF) = B/(B+D).

2.3 Histological analysis

Lung biopsy was conducted for 35 patients. The biopsied tissue was processed with immunohistochemical stain, CD31, for micro vessel density (MVD) calculation, as well as the routine hematoxylin and eosin (H&E) stain. The H&E stain was used for lung cancer differentiation, and the results were used as the ground truth in this study. Average MVD in the CD31 stain was obtained over five different areas for each patient. MVD has been related to the pathological change of the blood perfusion property, a strong indicator of malignancy in many organs.

2.4 ROC analysis

In this section, we will introduce briefly how ROC analysis is conducted and what basic principles of ROC analysis should be understood, so that clinicians or medical researchers can use the method properly. More detailed information can be found in the references cited in the following text.

ROC curves are plotted in a two-dimensional space, the x axis of which is false-positive fraction (FPF, or 1 - specificity) and the y axis true-positive fraction (TPF, or sensitivity). The definitions of FPF and TPF can be found in Table 1. Empirically, an ROC curve can be constructed by connecting all neighboring (1 - specificity, sensitivity) points, which can be called operating points conventionally (6). An empirical ROC is shown in Figure 1, using MVD data as an example. This kind of curve is often called empirical ROC

curve. The variable on which the specificity and sensitivity are calculated is the decision variable.

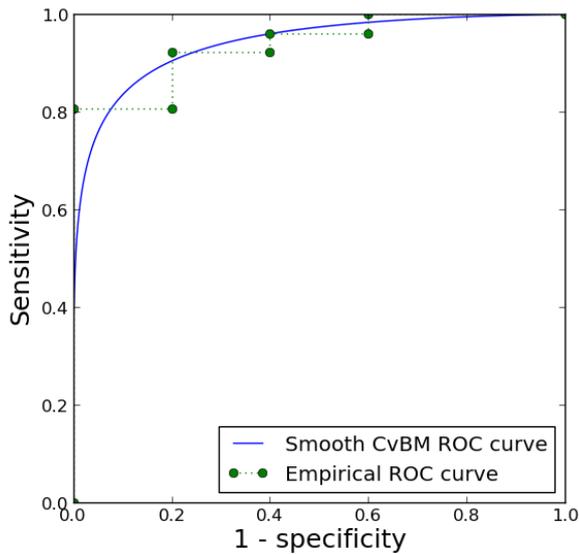


Figure 1. An example empirical and its corresponding smooth ROC curves. The MVD data is used to generate the curves.

Smooth ROC curves can be obtained by fitting ROC models to the operating points based on a set of certain assumptions. The most-commonly used model is the conventional binormal model (CvBM) (7). In CvBM, we assume that the normalized histograms of the decision variable for cancer and benign cases can be converted into two normal distributions (therefore binormal) by a monotonic transform function, which may or may not be known explicitly. Each of the operating points can be generated by changing the threshold of calling a case positive, as for the empirical ROC curve plotting. Now, based on the binormal model, we can further convert the operating points into normal-deviate coordinates. Mathematically we can prove that the operating points should appear around a straight line under the CvBM assumptions. Converting the straight line back into the (1-specificity, sensitivity) coordinates, a smooth ROC can be obtained.

An important aspect of ROC analysis is how to compare ROC curves, i.e., when two ROC curves are obtained from two datasets, one needs to determine whether or not the difference of the ROC curves is caused by statistical noise only. Normally, AUC is used for the analysis of statistical significance, but other ROC metrics can be used as well (10). It is worth to note that the comparing methods are different for correlated and independent datasets, since the correlation between the cases needs to be considered in the statistical analysis.

ROC analysis has several advantages over other accuracy metrics for binary classification problems. In medical literature, people often use "accuracy", which should be called the percentage of correctness (PC). Numerically, $PC = (\text{number of correct calls}) / (\text{number of all calls}) * 100\%$. An obvious deficit

of PC is that it is strongly influenced by the prevalence of a disease. For example, for a disease of 1% prevalence, simply calling every case "non-disease" will yield 99% PC but it will not help accurately tell difference between the disease and non-disease states at all. Therefore, PC should not be used in most medical research scenarios.

Another commonly used accuracy metric is the pair of sensitivity and specificity. The sensitivity and specificity can be considered as the percentage of correctness defined for the disease cases and the non-disease cases, respectively. Since two classes are considered separately, it is not influenced by the prevalence of disease any more. However, the calculation of sensitivity and specificity relies on a threshold. Cases whose decision variable is greater than the threshold will be considered in the disease state, otherwise in the non-disease state. The disadvantage of this metric is its (implicit) dependency of the threshold. Because different studies may use different thresholds, it is usually difficult to compare studies using sensitivity and specificity.

ROC curve includes all possible pairs of sensitivity and specificity. Therefore, all possible thresholds have been included in the curve. Usually, area under the ROC curve (AUC) is used to summarize the accuracy of binary classification results (11). And the AUC value can also be used for comparison of two ROC curves. In situations where only a range of sensitivities is of interest, horizontal partial AUC value can be used (vertical partial AUC value for specificity) (12). Other metrics can be used for comparison are the TPF value at a given FPF and the FPF value at a given TPF.

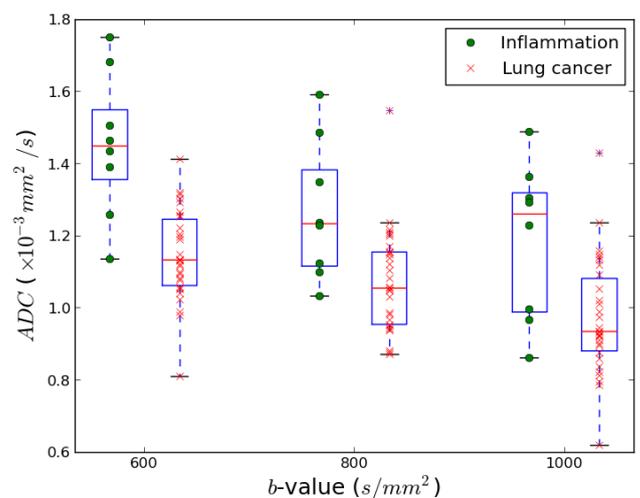


Figure 2. The boxplots of ADC values of DWI-MR images acquired with different b-values. The dots and crosses on top of the boxplots represent the locations of the data points.

2.5 Statistical analysis

Beside ROC analysis, we used student's T test to

compare ADC values between patients with and without lung cancer. P-values less than or equal to 0.05 was considered statistically significant.

3 RESULTS

The prevalence of lung cancer was 78.9% (30/38) in our dataset. Boxplots in Figure 2 show the difference between the patients with and without lung cancer in ADC value. Statistically significant differences were observed: p value was < 0.0001 , 0.003, and 0.004 for b-value of 600, 800, and 1000 s/mm^2 , respectively.

The empirical ROC and smooth CvBM ROC curves of ADC values calculated from different b-values are shown in Figure 3. The AUC values of the ROC curves are 0.92 ± 0.07 , 0.82 ± 0.09 , and 0.79 ± 0.10 , for b-values of 600, 800, and 1000 s/mm^2 , respectively. The differences in AUC value between the highest and the other two ROC curves were statistically significant (both p-values were 0.02). Therefore, DWI-MR with b-value of 600 s/mm^2 produced the best ADC values in differentiating lung cancer from infectious nodules.

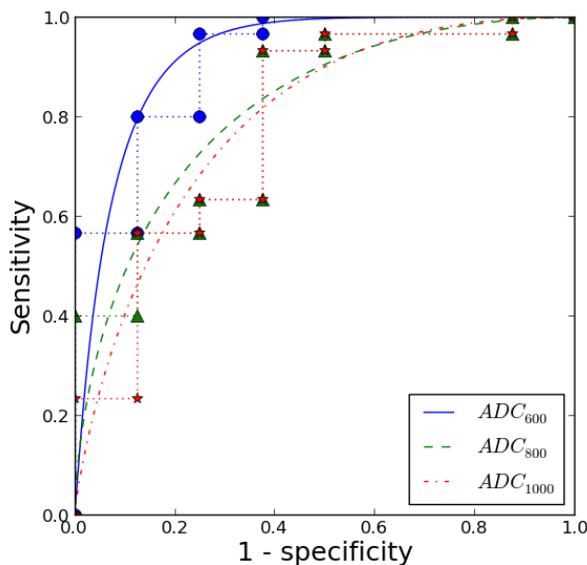


Figure 3. The empirical and smooth CvBM ROC curves for ADC values of DWI-MR images acquired with different b-values.

4 DISCUSSION

ROC is a clear and convenient assessment method for diagnostic accuracy. In this paper, we introduced the basic idea of ROC analysis and focused on how to construct an ROC curve, either empirically or using the CvBM. Fortunately, the algorithms for ROC curve plotting have been developed and validated over the past decades. Medical researchers can download and use the developed

software freely to conduct ROC analysis conveniently (13).

In this paper, we use DWI-MR images for lung cancer detection as an example to demonstrate how ROC analysis may help assess the accuracy of disease detection or diagnosis. The "accuracy," or more specifically, the percentage of correctness, should not be used for the assessment of diagnostic accuracy.

5 ACKNOWLEDGMENT

This study was supported in part by Beijing Jiaotong University Talent Fund 2013RC028 and Research Fund for the Doctoral Program of Higher Education of China 20110009110001.

REFERENCES

- [1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin.* 2012;62(1):10-29.
- [2] Erasmus JJ, Connolly JE, McAdams HP, Roggli VL. Solitary pulmonary nodules: Part I. Morphologic evaluation for differentiation of benign and malignant lesions. *Radiographics.* 2000;20(1):43-58.
- [3] Jeong YJ, Lee KS, Jeong SY, et al. Solitary pulmonary nodule: Characterization with combined wash-in and washout features at dynamic multi-detector row CT. *Radiology.* 2005;237(2):675-683.
- [4] Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions - A meta-analysis. *Jama-Journal of the American Medical Association.* 2001;285(7):914-924.
- [5] Ohno Y, Sugimura K, Hatabu H. MR imaging of lung cancer. *European Journal of Radiology.* 2002;44(3):172-181.
- [6] Metz CE. Basic principles of ROC analysis. *Semin Nucl Med.* 1978;8(4):283-298.
- [7] Metz CE. ROC methodology in radiologic imaging. *Invest Radiol.* 1986;21(9):720-733.
- [8] Qayyum A. Diffusion-weighted imaging in the abdomen and pelvis: concepts and applications. *Radiographics.* 2009;29(6):1797-1810.
- [9] Peng YH, Jiang YL, Antic T, et al. Apparent diffusion coefficient for prostate cancer imaging: Impact of b values. *American Journal of Roentgenology.* 2014;202(3):W247-W253.
- [10] Metz CE, Wang P-L, Kronman H. A new approach for testing the significance of differences between ROC curves measured from correlated data. In: Deconinck F, ed. *Information processing in medical imaging: The Hague: Martinus Nijhoff, 1984; p. 432-445.*
- [11] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36.
- [12] Jiang Y, Metz CE, Nishikawa RM, Doi K. Diagnostic performance index of a partial area to summarize an ROC curve in high-sensitivity regions. *Radiology.* 1995;197:357-357.
- [13] Metz CE. ROC software. [cited 2012 June, 25th]; Available from: <http://metz-roc.uchicago.edu>.