Clinical Study of Artificial Intelligence in Imaging Diagnosis of False Positive Lesions of Pulmonary Nodules

He Sun¹, Jiaheng Wei¹, Junfu Wang², Zhanyue Pang²,³ and Liangming Zhu⁴,⁵

¹School of Clinical Medicine, Weifang Medical University, Weifang, Shandong, China
²Department of Thoracic Surgery, Jinan Central Hospital, Jinan, Shandong, China
³Department of Thoracic Surgery, Jinan Central Hospital, Shandong University, Jinan, Shandong, China
⁴Department of Thoracic Surgery, School of Medicine, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, China

ABSTRACT

Objective: To determine the accuracy of diagnosis of pulmonary nodules using artificial intelligence method.

Study Design: Observational study.

Place and Duration of the Study: Department of Thoracic Surgery, Jinan Central Hospital, Jinan, China, from January 2020 to May 2021.

Methodology: An analysis of clinical characteristics exhibited by 32 patients initially diagnosed with malignant tumours through imaging (LDCT) and artificial intelligence (AI), was reclassified as having benign lesions following surgical intervention. Quantitative parameters were assessed, including CT mean value, kurtosis, skewness, solid ratio, and the ratio of length to short diameter, within a cohort of 32 benign patients juxtaposed with 58 patients diagnosed with lung cancer during the same time frame. The AI-derived parameters were subjected to Mann-Whitney U non-parametric test.

Results: A total of 32 benign pulmonary lesions were evaluated that were initially misdiagnosed as malignant prior to surgery. These lesions displayed an average length of (18.56 ± 12.16) mm, with the majority characterised as solid (68.8%). Notably, a substantial proportion of these lesions exhibited imaging features akin to malignant growths. The AI-derived quantitative parameters of the 32 benign cases and the 58 malignant cases revealed statistical significance in average CT value and solid ratio. However, statistical significance was not established for kurtosis, skewness, or the ratio of length to short diameter. The area under the Receiver Operating Characteristic (ROC) curve for average CT value and solid ratio stood at 0.71 and 0.705, respectively.

Conclusion: Among the cases initially misdiagnosed as malignant yet subsequently identified as benign, a notable number of these instances were solid nodules, often resembling malignant lesions in imaging characteristics. There was moderate discriminatory capacity for average CT value and solid ratio, rendering them valuable tools for distinguishing between benign and malignant lesions within this particular cohort. This underscores their high diagnostic significance.

Key Words: Artificial intelligence, Benign lesions of lung, Lung cancer, Quantitative parameters, Postoperative.


INTRODUCTION

Lung cancer holds the highest mortality rate among malignant tumours globally, as supported by the statistical data. In China, the incidence and fatality rate of lung cancer are also prominently elevated among malignancies. In the realm of medical advancements, artificial intelligence (AI) has rapidly emerged as a transformative technology, finding widespread utility. Researchers, both domestic and international, have harnessed AI for analysis and exploration of pulmonary nodules. Of these technologies, Convolutional Neural Networks (CNNs), as distinguished deep learning algorithms, have garnered substantial attention for their role in diagnosing pulmonary nodules.

Notably, studies reveal that 2D and 3D CNNs boast a detection sensitivity of approximately 95% for pulmonary nodules. Computed tomography (CT) assumes a pivotal role in the early detection of lung cancer through nodule screening, thereby significantly mitigating the mortality rate. According to the relevant literature, the current detection rate of chest CT for pulmonary nodules has reached 21-33%. In the developing countries’ areas with high incidence of tuberculosis, chest CT can detect more pulmonary lesions.

The contemporary proliferation and rapid advancements in low dose computed tomography (LDCT) have substantially bolstered pulmonary nodule detection rates. LDCT exhibits heightened sensitivity in the early lung cancer diagnosis; however, it is characterised by a notable false-positive rate, diminished specificity, and susceptibility to physicians’ subjectivity. AI, in its capacity to rapidly differentiate between benign and malignant pulmonary lesions across an extensive array of CT images, emerges as a crucial tool. The management of benign nodules typically involves vigilant observation. For lesions defined as intermediate in nature, bronchoscopy and lung biopsy are convention-
ally advised. Nevertheless, certain nodules remain diagnostically elusive due to factors such as location and size, compelling patients to seek resolution through video-assisted thoracoscopy, driven by their comprehension of the tumorous condition and ensuing psychological stress. Notably, patients confronted with indeterminate pulmonary lesions often opt for surgical intervention to ascertain lesion nature, thereby imposing a multifaceted toll on their well-being, mental equilibrium, and society at large. Addressing the imperative of minimising the frequency of surgical interventions for benign cases assumes critical significance.

In this study, the aim was to determine the average CT value and solid ratio for improving the accuracy of diagnosis of pulmonary nodules based on artificial intelligence method, which has contributed to develop a novel tool to classify pulmonary nodules.

**METHODOLOGY**

A total of 244 patients including 35 malignant tumours and 209 lung cancer were collected from Jinan Central Hospital in this study. The exclusion criteria was patients who received pulmonary lesion resection from January 1, 2020 to May 31, 2021, the thickness of chest CT layer (1mm or 1.25mm) and the postoperative pathological conditions such as preoperative bronchoscopy, lung puncture biopsy, lymph node biopsy, the history of other tumours and chemoradiotherapy, and findings of squamous cell carcinoma or small cell lung cancer. Finally, 32 benign cases and 58 malignant cases were finally included in this study.

The CT equipment models, GE750HD (USA), Siemens SOMATIOM Definition AS (Germany) were used for the study.

The CT scanning parameters were tube rotation time 0.5s, standard soft tissue reconstruction algorithm, slice thickness 1mm or 1.25mm, reconstruction interval 1mm or 1.25mm x DFOV 200mm-410mm; tube voltage 120V, tube current 160mA, lung window: 800HU / 750HU; and mediastinal window: 400HU / 40HU. Utilising the capabilities of the Infervision artificial intelligence assisted diagnosis software, hinged on deep neural network technology for auxiliary analysis of medical images, the quantitative parameters of patient CT images underwent analysis across the stages of anomaly detection (lesion identification), differential diagnosis (intelligent localization), segmentation measurement (quantitative analysis), image registration (intelligent tracking), and generation of intelligent reports.

The dataset was statistically processed and visualised using SPSS version 27 software. Expressing measurement data as mean ± standard deviation (mean ± SD), while adoption rates of count data (%) were employed. The normality of data distribution was assessed using the S-W test. Non-normally distributed continuous variable data were represented using the median and interquartile range M (Q25, Q75). The comparison between the two groups was executed through the Mann-Whitney U test. Parameters exhibiting statistical significance underwent further analysis via ROC curve calculations. A statistical significance threshold of p < 0.05 was adopted.

### RESULTS

This investigation encompassed 32 patients who were initially diagnosed with malignant lesions preoperatively, only to be reclassified as having benign lesions postoperatively. Among the participants, the male-to-female ratio stood at 18:14, with an average age of 58.87 ± 8.44 years. The average long diameter measured 18.56 ± 12.16 mm, with lesions ≥ 8mm constituting 93.7% of cases. Notably, 22 cases (68.8%) presented as solid lesions. The observations revealed the presence of lobulation signs in 21 cases (65.6%), spiculation signs in 17 cases (53.1%), pleural traction signs in 10 cases (31.3%), vascular convergence signs in 3 cases (9.4%), and vacuole signs in 1 case (3.1%). Among them, 20 patients (62.5%) were found to have pulmonary lesions due to health examination. All the 32 patients with benign lesions underwent intraoperative rapid pathological examination, of which 1 case was diagnosed as atypical hyperplasia during operation and pulmonary inflammatory lesions were indicated by routine pathological report after operation. The coincidence rate between rapid frozen pathological section and routine pathological section was 96.9%. Postoperative routine pathological types showed that the top three were inflammatory lesions in 13 cases (40.6%), atypical hyperplasia in 11 cases (34.4%) and hamartoma in 6 cases (18.8%, Table 1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter of pulmonary lesions (mm)</td>
<td>20-30</td>
</tr>
<tr>
<td>Lung signs</td>
<td>3(9.4%)</td>
</tr>
<tr>
<td>Density classification</td>
<td>5(15.6%)</td>
</tr>
<tr>
<td>Pathological results</td>
<td>22(68.8%)</td>
</tr>
<tr>
<td>Partial solid nodule</td>
<td>7(21.9%)</td>
</tr>
<tr>
<td>Other</td>
<td>2(6.3%)</td>
</tr>
<tr>
<td>No special signs</td>
<td>11(34.4%)</td>
</tr>
<tr>
<td>Pleural traction sign</td>
<td>10(31.3%)</td>
</tr>
<tr>
<td>Vessel convergence sign</td>
<td>3(9.4%)</td>
</tr>
<tr>
<td>Lobulation sign</td>
<td>21(65.6%)</td>
</tr>
<tr>
<td>Spicule sign</td>
<td>17(53.1%)</td>
</tr>
<tr>
<td>Vacuole sign</td>
<td>1(3.1%)</td>
</tr>
<tr>
<td>Pure ground glass nodule</td>
<td>3(9.4%)</td>
</tr>
<tr>
<td>Inflammatory lesions</td>
<td>13(40.6%)</td>
</tr>
<tr>
<td>Atypical hyperplasia</td>
<td>11(34.4%)</td>
</tr>
<tr>
<td>Hamartoma</td>
<td>6(18.8%)</td>
</tr>
<tr>
<td>Detected by physical examination</td>
<td>Yes 20(62.5%)</td>
</tr>
</tbody>
</table>

**Table I: Clinical and imaging features of lung benign lesions.**
Clinical study of artificial intelligence in imaging diagnosis of false positive lesions of pulmonary nodules

Table II: Analysis of AI quantitative parameters by statistical method.

<table>
<thead>
<tr>
<th>Quantitative parameters median (IQR)</th>
<th>Benign lesion (n=32)</th>
<th>Malignant lesion (n=58)</th>
<th>p-value*</th>
<th>AUC (p-value)</th>
<th>Asymptomatic 95% Confidence Interval Lower-Upper</th>
<th>ppv</th>
<th>npv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average CT value</td>
<td>-95 (-360.75, 33)</td>
<td>-458 (-572, -66.75)</td>
<td>0.001*</td>
<td>0.71 (0.001)</td>
<td>0.6-0.819</td>
<td>87.4%</td>
<td>76.9%</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.92 (-1.16, -0.37)</td>
<td>-0.8 (-1.11, -0.11)</td>
<td>0.451*</td>
<td>0.452 (0.451)</td>
<td>0.324-0.580</td>
<td>32.5%</td>
<td>26.3%</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.1 (-0.02, 0.59)</td>
<td>0.35 (0.06, 0.67)</td>
<td>0.052*</td>
<td>0.376 (0.052)</td>
<td>0.246-0.505</td>
<td>30.6%</td>
<td>25.7%</td>
</tr>
<tr>
<td>Solid ratio</td>
<td>76.56 (15.46, 97.38)</td>
<td>12.08 (1.15, 77.86)</td>
<td>0.001*</td>
<td>0.705 (0.001)</td>
<td>0.598-0.812</td>
<td>84.6%</td>
<td>73.5%</td>
</tr>
<tr>
<td>Ratio of long diameter to short diameter</td>
<td>1.28 (1.17, 1.4)</td>
<td>1.33 (1.17, 1.51)</td>
<td>0.466*</td>
<td>0.453 (0.466)</td>
<td>0.331-0.576</td>
<td>33.5%</td>
<td>29.7%</td>
</tr>
</tbody>
</table>

* Mann-Whitney U test; the difference was statistically significant (p <0.05).


Infervision artificial intelligence software was employed to determine AI-derived quantitative parameters (CT mean value, kurtosis, skewness, solid ratio, ratio of length to short diameter) across 32 patients with benign lesions and 58 patients with malignant lesions. The outcome of the Mann-Whitney U test demonstrated that in comparison to the malignant lesion group, the benign lesion group exhibited higher average CT values and a greater proportion of solid lesions, both achieving statistical significance (p <0.05). Conversely, statistical significance was not observed for kurtosis, skewness, and the ratio of length to diameter (p >0.05). Detailed results can be found in Table II.

Figure 2: Result of Infervision artificial intelligence software for benign lesions.

Figure 3: Result of Infervision artificial intelligence software for malignant lesions.

Representative CT images of benign and malignant lesion cases are illustrated in Figure 1, while artificial intelligence software-aided analyses of benign and malignant lesions are depicted in Figures 2 and 3, respectively.

As shown in Table II, ROC analysis showed AUC of average CT value and solid ratio was 0.71 (p = 0.001), and 0.705 (p = 0.001), respectively. The best critical values of CT average and solid ratio were -502.5 HU (sensitivity = 90.6%, specificity = 46.6%) and 7.96% (sensitivity = 93.8%, specificity = 44.8%).
DISCUSSION

The current study delved into an examination of the clinical and imaging characteristics of 32 patients initially diagnosed with malignancy, only to be reclassified postoperatively as having benign lesions during the period spanning from January 2020 to May 2021 within the study hospital. Most patients were identified through the routine health examinations, underscoring a heightened awareness of health among individuals. It is imperative to improve the precision of identifying potentially malignant benign lesions, thereby alleviating unwarranted patient burdens. As per prevailing guidelines, where malignancy risk ranges from 3 to 68% for uncertain nodules, pursuing CT-guided percutaneous lung biopsy or bronchoscopic biopsy to obtain definitive diagnoses represents a feasible approach, effectively reducing the prevalence of unnecessary surgical interventions.

The evaluation encompassed a total of 32 pathologically established benign lesions, with a noteworthy 93.7% exhibiting sizes exceeding 8cm. Among these, 9.4% were pure ground glass nodules, while 90.7% presented as mixed ground glass and solid lesions. Simultaneously, malignancy-associated imaging attributes like the spiculation sign, vacuole sign, lobulation sign, vascular convergence sign, and pleural traction sign were evident. Importantly, a substantial proportion of suspected malignant benign lesions manifested imaging features resembling malignancies. The existing literature has extensively explored the attributes of pulmonary nodules. The authors can augment assessments by scrutinising characteristics such as calcification and fat density indicative of benign lesions, thereby enhancing the evaluation of suspicious lesions. Additionally, this study found a notable 96.9% agreement between rapid frozen pathological assessments and conventional pathology reports. This underscores the significance of rapid frozen pathology in expediting operations and refining the scope of necessary lung resections, affirming its indispensable role. Furthermore, this research revealed that the majority of benign lesions treated within our hospital were inflammatory in nature, followed by hamartoma and atypical hyperplasia. Clinicians routinely implement empirical anti-inflammatory treatments for suspicious lesions. However, some patients may receive non-standardised drug regimens, and specific infectious lesions might not resolve promptly, leading to diagnostic ambiguity. To address this, consistent monitoring and standardised antibiotic usage are advocated. Concurrently, targeted laboratory tests for specific infections can be instrumental in enhancing the accuracy of assessments for suspicious lesions.

Prior research has indicated that the average and maximum diameters observed through CT scans act as independent risk factors of notable value in discerning between benign and malignant pulmonary lesions. Concurrently, density attributes such as kurtosis and skewness of pulmonary nodules have also displayed discriminative potential between the two categories, albeit often in conjunction with other parameters. A significant contrast in solid components had been observed between malignant and benign pulmonary nodules. Building upon these findings, this study employed Infervision artificial intelligence software to determine AI-derived quantitative parameters (average CT value, kurtosis, skewness, solid ratio, ratio of length to short diameter). A comparison was drawn between 32 patients with benign lesions and 58 patients with pathologically confirmed malignant lesions. Remarkably, the average CT value and solid ratio emerged as the two parameters exhibiting statistical significance. In imaging features, because there are different degrees of morphological similarity between benign and malignant lesions, kurtosis, skewness and the ratio of length to short diameter could not provide a reference for the differential diagnosis of malignant benign lesions. For suspected malignant benign lesions, the average value and solid ratio of CT could still provide some suggestions for the differentiation of benign and malignant lesions. The average value of CT represented the mean value of CT value of pulmonary nodules, and the solid proportion represented the proportion of solid components in pulmonary nodules. These two indicators represented the density characteristics of pulmonary nodules. Through the analysis of ROC curve, the best critical values of CT average and solid ratio were -502.5 HU (sensitivity = 90.6%, specificity = 46.6%) and 7.96% (sensitivity = 93.8%, specificity = 44.8%). Because the sensitivity of these two kinds of parameters is more than 90%, it can be found that the rate of missed diagnosis is low. When the average value of CT is more than -502.5 HU, it is more inclined to suspected malignant benign lesions, when the average value of CT is less than -502.5 HU, it is more inclined to malignant lesions; when the solid proportion is more than 7.96%, it is more inclined to suspected malignant benign lesions, and when the average CT is less than 7.96%, it is more inclined to malignant lesions. Most of these benign lesions are inflammation and hamartoma with more solid components. The higher the solid component, the higher the mean CT value. The related literature suggests that most of the benign nodules of the lung are solid nodules and most of the malignant nodules are partial solid nodules. In this study, it was found that the average CT and solid ratio of these benign nodules were higher than those of malignant nodules, which was consistent with the conclusion of the proportion of solid components in benign and malignant nodules. The average area of CT in ROC curve is 0.71 and the area of solid ratio is 0.705, indicating that the average value of CT and solid ratio are of high predictive value, which can be used to distinguish benign lesions suspected of malignancy from malignant lesions.
The limitations of this study stemmed from the fact that it was a retrospective case analysis. Being a single-centre study, the sample size and index parameters are less. The authors hope to increase more samples in the future, take multi-centric, prospective research, at the same time, with the development of AI intelligent software, as that will lead to more indicators for follow-up researches. Despite the aforementioned limitations, the authors are confident that this study will help to better identify benign lesions suspected of malignancy and increase people's understanding of lesions that are difficult to define. It will be helpful to the practice of clinicians and radiologists.

CONCLUSION
The presence of benign nodules displaying imaging attributes akin to malignant lesions, presents diagnostic challenges. Upon juxtaposing postoperative pathological findings of benign and malignant groups, the ROC curve analysis of average CT value and proportion of solid lesions demonstrated a moderate area under the curve, signifying their substantial diagnostic significance. These outcomes hold the potential to guide efforts aimed at mitigating the proportion of surgical interventions in cases of benign pulmonary lesions.

ETHICAL APPROVAL:
The study was approved by the Ethics Committee of Jinan Central Hospital, Shandong, China (No. 2022-244-01, dated 2022.10.07).

PATIENTS’ CONSENT:
Informed consents were obtained from the patients for performing the tests and to publish the obtained data.

COMPETING INTEREST:
The authors declared no competing interest.

AUTHORS’ CONTRIBUTION:
HS: Design, acquisition and analysis of data, and writing of the manuscript.
JW, JW, ZP: Interpretation and discussion of results.
LZ: Proofreading and final approval of the final manuscript. All authors approved the final version of the manuscript to be published.

REFERENCES

