Research paper

Cite as: Levi G, Rocchetti C, Inciardi RM, Bezzi M, Pini L, Tantucci C, Marchetti G: Standardized bilateral thoracic ultrasound image comparison as a tool for the diagnosis of pneumothorax: a pilot exploratory study. J Ultrason 2025; 25: 24. doi: 10.15557/JoU.2025.0024.

Submitted: 22.05.2025
Accepted: 27.07.2025
Published: 30.09.2025

Standardized bilateral thoracic ultrasound image comparison as a tool for the diagnosis of pneumothorax: a pilot exploratory study

Guido Levi[1,2], Chiara Rocchetti[1], Riccardo Maria Inciardi[3], Michela Bezzi[1], Laura Pini[2], Claudio Tantucci[2], Giampietro Marchetti[1]

[1] *Pulmonology Unit, ASST Spedali Civili di Brescia, Italy*
[2] *Department of Clinical and Experimental Sciences, University of Brescia, Italy*
[3] *Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health, University of Brescia, Italy*

Corresponding author: Guido Levi; e-mail: guido.levi@asst-spedalicivili.it

DOI: 10.15557/JoU.2025.0024

Keywords
pneumothorax; diagnostic imaging; lung ultrasound; point-of-care ultrasound; diagnostic techniques and procedures; software tools

Abstract

Aim: Pneumothorax is a potentially life-threatening condition whose diagnosis can be challenging. Ultrasound chest examination is generally fast and user-friendly, but in non-expert hands or with uncooperative patients, it may still be difficult and time-consuming. Adding another tool to support the suspicion of pneumothorax might be useful, potentially enhancing the diagnostic accuracy of standard ultrasound chest examination. We evaluated the feasibility of standardized bilateral ultrasound image comparison as a potential new tool for pneumothorax diagnosis. **Materials and methods:** We enrolled 60 subjects (30 with pneumothorax and 30 controls) and collected bilateral ultrasound images of their chests (each image contained one frame from the left lung and one from the right lung). Ten physicians (eight blinded to diagnosis) divided into five groups according to expertise evaluated the images for potential grayscale differences and/or horizontal artifacts between the two frames. All images were then analyzed with image analysis software for grayscale pixel assessment (one sub-analysis for the entire area under the pleural line, one for a 100-pixel-wide rectangle under the pleural line). **Results:** All clinicians achieved good results in terms of diagnostic accuracy and inter-operator reliability, even those unexperienced in ultrasound. Mean, range, and median grayscale pixel ratio between the pneumothorax side and the healthy side in a single patient proved to be the most reliable parameters, reaching excellent sensitivity and specificity. Combining these parameters proved to be an excellent diagnostic tool (ROC area under curve = 1.00, *p*-value = 0.02). **Conclusions:** Standardized bilateral thoracic ultrasound image comparison may be a potential new tool for the diagnosis of pneumothorax.

Introduction

Pneumothorax (PTX) is defined as the abnormal presence of air in the pleural cavity and represents a potentially life-threatening condition. The gold standard for its diagnosis is the chest computed tomography (CT) scan[1], which provides accurate, multiplanar imaging of thoracic structures and allows detection of even small PTXs that may be missed on standard chest X-rays. However, CT scans are relatively expensive, not suitable for bedside use, and expose patients to a non-negligible dose of ionizing radiation[2].

The use of chest ultrasound (US) for PTX diagnosis is well-established. US offers several advantages, including rapid execution at the patient's bedside[3,4] and suitability for critical care settings, while avoiding exposure to ionizing radiation. Key ultrasonographic signs of PTX include the "lung point"[5], absence of lung sliding (a dynamic horizontal movement of the hyperechoic pleural line during breathing in a normal lung)[6], and absence of B-lines[7]. The addition of standard M-mode imaging (e.g., the "stratosphere sign" or "barcode sign") to B-mode can enhance the visibility of these features[8]. Among these, the lung point is the most specific dynamic sign, with specificity close to 100%; its location roughly corresponds to the radiological size and position of the PTX, and it shows good sensitivity (up to 75% for radiographically occult PTX)[5]. Furthermore, US elastography has been explored as a potential tool for PTX diagnosis, including identification of the Elasto-lung point[9]. Importantly, combining multiple sonographic signs increases the likelihood of accurately diagnosing PTX, improving both sensitivity and specificity in clinical practice[10]. Nonetheless, recognizing PTX with US can be challenging and time-consuming, particularly in

critically ill patients who are unable to cooperate or reposition (e.g., major trauma patients)[11], or in those with hypoventilation. Even in cooperative patients, chest US typically visualizes only approximately 70% of the pleural surface. Detection of the lung point can also be difficult in patients with bullous lung disease or COPD, in those with morbid obesity[12], in cases of hydropneumothorax[13–15], or when a "double lung point" is present. In these situations, full exploration of the chest wall may be limited, and image quality may be significantly reduced. Moreover, the diagnostic accuracy of US is heavily dependent on the operator's experience[16–18], a limitation also confirmed by a recent systematic review in trauma patients, despite US demonstrating superior performance compared to chest X-ray for PTX detection[19].

Given these limitations, an additional tool to support the suspicion of PTX could be valuable, potentially increasing the sensitivity of standard US.

In our experience, thoracic US images in PTX cases (without a visible lung point) appear to exhibit higher grayscale intensity compared to the contralateral healthy lung; additionally, horizontal artifacts (A-lines) tend to be more prominent and better defined. To our knowledge, no previous study has systematically evaluated the feasibility of subjective and/or objective bilateral comparison of thoracic US images as a potential diagnostic approach for PTX. Therefore, we conducted a case-control study to assess the utility of this comparison as a diagnostic tool. Our study highlights differences in grayscale and A-line patterns between healthy lungs and those affected by PTX, integrating both subjective assessment and objective image analysis.

Methods

This was a prospective, case-control pilot study conducted in 2020. A total of 60 subjects were enrolled from the Pulmonology Department of the ASST Spedali Civili, located in Brescia, Italy. The study protocol adhered to the ethical principles outlined in the 1975 Declaration of Helsinki. The study was approved by the internal Institutional Review Board (IRB), and written informed consent was obtained from all participants.

The case group included 30 consecutive adult patients (aged ≥18 years) with a confirmed diagnosis of PTX (spontaneous, traumatic, or iatrogenic) based on chest CT or intraoperative findings during thoracoscopy. Notably, thoracoscopy involves intentional insufflation of air into the pleural cavity to induce lung collapse and allow visualization of the internal chest wall for biopsy. Following the procedure, a chest drain is placed through the surgical incision to evacuate intrapleural air and facilitate gradual lung re-expansion.

The control group comprised 30 hemodynamically stable patients, matched in number, with no evidence of PTX on chest CT performed for other clinical indications (e.g., lung cancer staging, evaluation of hemoptysis, etc.).

Patients with obesity (body mass index >30 kg/m^2), pulmonary emphysema, interstitial lung disease, major thoracic trauma, or evidence of anterior pleural abnormalities (e.g., thickening, adhesions, or malignancy such as mesothelioma) were excluded to minimize potential interference with US image quality. Additionally, due to the ongoing COVID-19 pandemic during the study period, individuals with confirmed or suspected SARS-CoV-2 infection were excluded. Patients with tension PTX were also not included, given the urgency and life-threatening nature of this condition, which precludes standardized imaging protocols.

All subjects were evaluated in the semi-supine position using the same US machine (Mindray M9, Mindray Medical International Limited, Shenzhen, China) equipped with a 7.5 MHz linear probe. The lung preset was applied with standardized settings (depth: 5 cm, gain: 30) to optimize visualization of pleural and subpleural structures and ensure consistency across examinations. The anterior chest wall was examined bilaterally, with all images acquired at the third intercostal space at a 90° angle and a fixed distance from the sternum.

Static US images were used to facilitate standardized comparisons across subjects and to eliminate variability associated with real-time imaging. For patients with PTX, two static images were acquired: one of the affected lung (deliberately excluding the lung point) and one of the contralateral healthy lung. For control subjects, static images were obtained from both lungs. Consequently, each patient contributed two representative frames: one from the left lung and one from the right (Fig. 1A, upper panel), resulting in a total of three images per patient.

All US examinations were performed by a clinician with at least two years of experience in chest ultrasonography. Demographic data, including age, sex, weight, and height, were also recorded.

Subsequently, all images were independently reviewed by two expert pulmonologists trained in chest US. For each subject, the best-defined image was selected for analysis. Both objective and subjective evaluations were then carried out on the selected images.

First, a subjective analysis was conducted. Ten physicians, divided into five groups of two, were individually convened. The first group comprised two pulmonologists who were not blinded to the final diagnosis. The second group included two different pulmonologists, the third group two cardiologists experienced in cardiac US, the fourth group two emergency medicine physicians, and the fifth group two physicians without US experience (one neurologist and one maxillofacial surgeon). All groups, with the exception of the first pair, were blinded to the final diagnosis.

Each physician independently evaluated the static US images presented in random order. For each case, they assessed whether any difference between the two lung images could be observed in terms of horizontal artifacts (A-lines) and/or grayscale intensity beneath the pleural line. Since PTX involves the presence of air, it was hypothesized that the affected side would show a higher grayscale value. A "yes" response for either feature (grayscale or A-lines) was considered indicative of suspected PTX. All physicians provided their assessment independently and were unaware of the responses of their peers.

Following the subjective evaluation, an objective analysis was performed using two image analysis software programs: ImageJ (National Institutes of Health, Bethesda, MD, USA)[20] and Adobe Photoshop (Adobe Systems Incorporated, San Jose, CA, USA). Given the pilot nature of the study, employing both platforms aimed to improve the robustness and reliability of the results through cross-

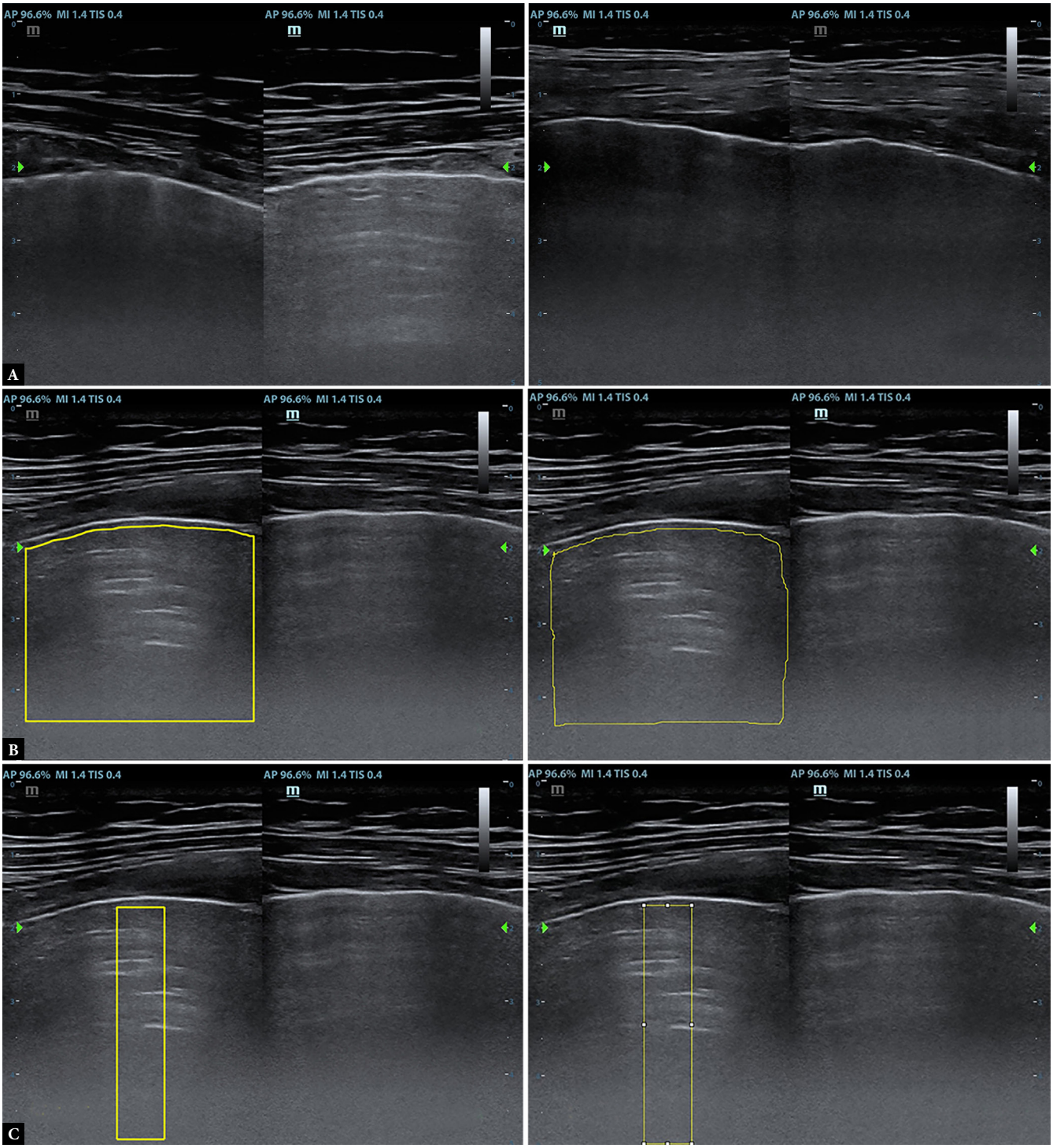

Fig. 1. *Ultrasound image analysis in pneumothorax and control subjects.* **A.** *Upper panel: Examples of sampled lung ultrasound images. Left frame: pneumothorax patient (PTX on the right side, note the higher grayscale intensity and prominent horizontal artifacts); right frame: control patient.* **B.** *Middle panel: Example of "global" sub-analysis. Left frame: Adobe Photoshop output; right frame: ImageJ output. Pneumothorax is on the left.* **C.** *Lower panel: Example of "targeted" sub-analysis. Left frame: Adobe Photoshop output; right frame: ImageJ output. Pneumothorax is on the left*

validation. Importantly, both programs were used solely for image analysis; no image editing or alterations were made prior to analysis.

These software tools were chosen due to their widespread use in medical research: ImageJ is a well-established open-source platform designed for quantitative biomedical image analysis, while Adobe Photoshop offers advanced image evaluation functionalities used in clinical imaging studies[21,22]. Their combined use allowed for a comprehensive and reproducible approach to quantifying grayscale differences and horizontal artifacts.

Each selected US frame underwent two sub-analyses: a "global" and a "targeted" analysis.

In the "global" analysis, comparable regions beneath the pleural line were manually selected in both lungs. Regions of interest (ROIs) were delineated using a mouse in ImageJ and a digital drawing pad in Adobe Photoshop (Fig. 1 B, middle panel), ensuring visual symmetry in size and location between the left and right lungs. Both software platforms extracted quantitative grayscale values within the ROI, ranging from 0 (total black) to 255 (total white). Additional parameters, including standard deviation, range, median (Adobe Photoshop only), and mode (ImageJ only), were also recorded. For all patients, both absolute grayscale values and grayscale ratios were calculated: in PTX cases, the ratio was the affected lung over the contralateral healthy lung; in controls, the higher value was placed in the numerator to ensure consistency.

The "targeted" analysis followed a similar methodology, but focused on a standardized 100-pixel-wide rectangular region selected at the center of each image, beneath the pleural line (Fig. 1 C, lower panel). This approach allowed a more localized assessment, minimizing variability and potentially enhancing sensitivity for detecting subtle grayscale or artifact differences.

The Mann-Whitney U test was used to assess significant differences in pixel evaluations between groups. In cases where distributions were non-normal, the Kolmogorov-Smirnov test was applied. To evaluate the diagnostic accuracy of the potential imaging tests, ROC curve analysis was conducted.

To quantify the relationship between the most promising pixel ratios (mean, range, and median) and the target outcome, odds ratios (ORs) were calculated using logistic regression models. Both unadjusted models (which included only the pixel ratios as predictors) and adjusted models (which accounted for potential confounders such as age, sex, and BMI) were applied to evaluate the impact of these variables.

For the subjective analysis, the Kruskal-Wallis test was used to compare diagnostic parameters across physician groups. Inter-rater reliability for categorical variables was assessed using Cohen's kappa coefficient to evaluate agreement between the physicians' evaluations.

Statistical analysis was conducted using GraphPad Prism 8.0 (GraphPad Software, CA, U.S.A.) and STATA (StataCorp LLC, TX, U.S.A.). Continuous variables are presented as mean ± standard deviation, while categorical variables are expressed as percentages. A p-value of less than 0.05 was considered statistically significant for all tests.

Results

A total of 60 subjects were enrolled, comprising 30 cases and 30 controls. In the case group, the male-to-female ratio was 22/8, while in the control group it was 23/7 (p-value >0.999). No significant differences were observed between the groups in terms of mean age (cases: 62.8 ± 17.1 vs. controls: 61.5 ± 15.4, p-value = 0.762) and BMI (cases: 24.5 ± 3.0 kg/m^2 vs. controls: 23.4 ± 3.4 kg/m^2, p-value = 0.225).

In the case group, iatrogenic PTX (post-thoracoscopy) was the most common type (21/30, 70%). The remaining nine cases were primarily spontaneous (6/30, 20%), with one post-traumatic PTX, one post-bronchoscopy PTX, and one following a right-sided nephrectomy due to kidney cancer. Among the post-thoracoscopy cases, 75% involved pleuro-pulmonary neoplasms (nine mesotheliomas, six lung adenocarcinomas, and one neuroendocrine lung cancer), while the remaining 25% had various other conditions, including pleural tuberculosis (n = 1), breast cancer (n = 1), gastrointestinal cancer (n = 1), and chronic pleuritis (n = 1).

Subjective analysis showed good diagnostic accuracy across all physician groups (at least 70%). No significant differences were found among the five groups, except for borderline diagnostic accuracy (Fig. 2, Tab. 1). Inter-observer reliability within the same group varied greatly, with the unblinded pulmonologist group and the emergency medicine experts showing the highest consistency. Physicians without prior US experience demonstrated better inter-observer reliability regarding horizontal artifacts, while the other groups performed more consistently in evaluating grayscale values (full results available in Tables S1–S4 in the Supplementary material).

In the objective analysis, significant differences were observed in the case group for parameters including mean, minimum, maximum, range, mode (only provided by ImageJ), and median (only provided by Adobe Photoshop). No significant differences were found in the control group. These findings were even more pronounced in the targeted sub-analysis and confirmed when comparing the ratios (i.e., PTX side/healthy side in cases, higher number side/lower number side in controls). Full results can be found in Tables S5–S7 and Figures S1–S16 (Supplementary material).

The mean ratio in the targeted sub-analysis proved to be the most reliable parameter for diagnosis, with an AUC ROC of 0.98 (95% CI: 0.95–1.00, p <0.001). A cut-off value of 1.22 yielded 87% sensitivity, 97% specificity, and a likelihood ratio of 27; similarly, a cut-off of 1.25 in ImageJ showed identical diagnostic performance (87% sensitivity, 97% specificity, likelihood ratio 26). When the cut-off value was set at 1.25 in Adobe Photoshop, specificity reached 100%, making it the optimal threshold.

For Adobe Photoshop, both the range ratio (AUC ROC: 0.89, 95% CI: 0.80–0.99, p <0.001) and median ratio (AUC ROC: 0.95, 95% CI: 0.91–1.00, p <0.001) were useful diagnostic parameters. A cut-off of 1.52 for the range ratio (sensitivity 70%, specificity 97%, likelihood ratio 21) and a cut-off of 1.32 for the median ratio (sensitivity 80%, specificity 97%, likelihood ratio 24) yielded optimal results (Tab. 2, Fig. 3A). Full results are available in Table S8 (Supplementary material).

Finally, when combining these ROC curves, a perfect ROC (AUC: 1.00, p = 0.02) was achieved (Fig. 3B). Odds ratios for these parameters were significant in both unadjusted and adjusted models (for age, sex, and BMI) (Tab. 2).

Almost all results obtained with Adobe Photoshop differed significantly from those obtained with ImageJ (see Tables S9 and S10 and Figures S17 and S18 in the Supplementary material).

Discussion

Our findings support the hypothesis that there is a statistically significant difference when comparing US images of PTX with the

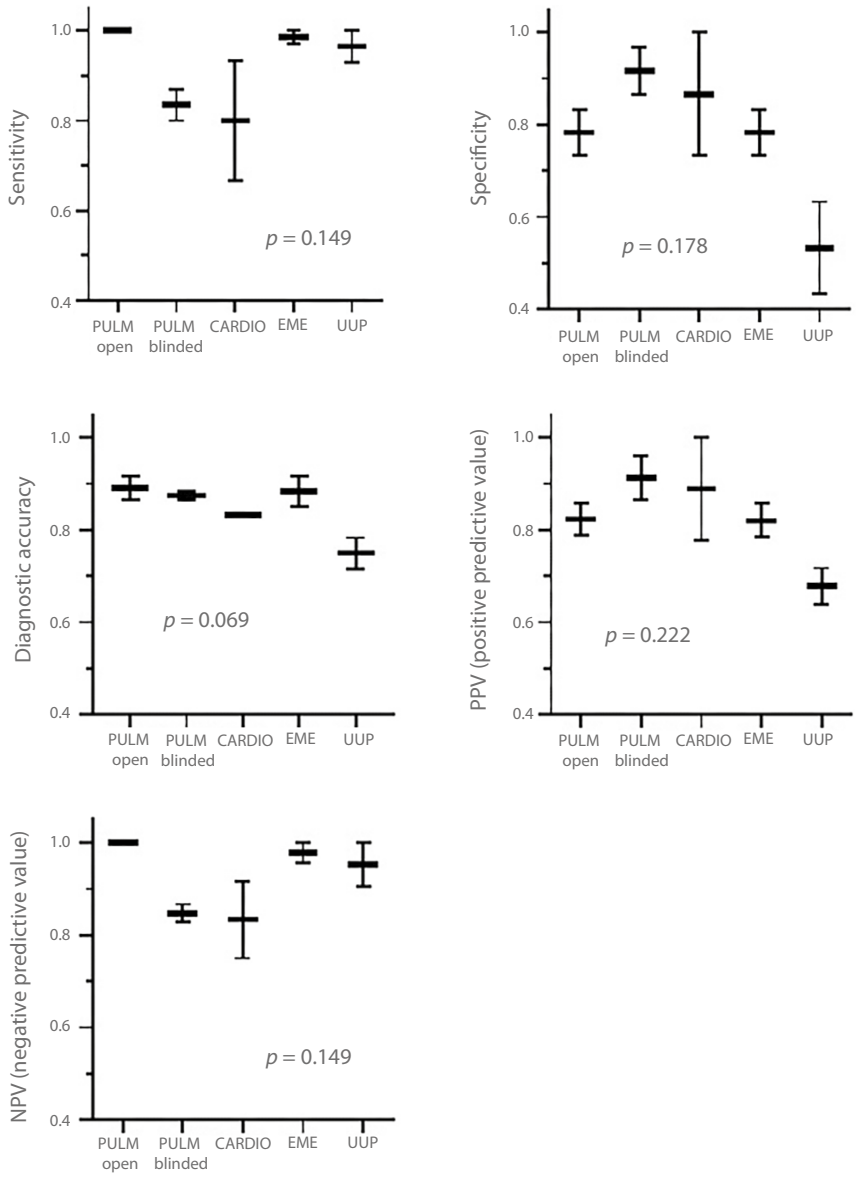

Fig. 2. *Subjective analysis. Differences among the five physicians' groups. PULM – pulmonologists; CARDIO – cardiologists; EME – emergency medicine expert; UUP – US un-experienced physicians; PPV – positive predictive value; NPV – negative predictive value*

healthy contralateral lung in the same patient. Both "global" and "targeted" sub-analyses proved valuable, with the "targeted" analysis being quicker and easier to perform. Among the evaluated parameters, the mean ratio in the targeted analysis was the most reliable, achieving excellent sensitivity and specificity, with a cut-off value of 1.25 (i.e., a mean ratio greater than 1.25 warrants further diagnostic investigation). Additionally, when using Adobe Photoshop, meeting three conditions (mean ratio >1.25, range ratio >1.52, and median ratio >1.32) further confirmed the PTX diagnosis. These parameters remained statistically significant even after adjusting for potential confounders such as age, sex, and BMI.

However, the combination of the ROC curves resulted in a perfect AUC of 1.00 ($p = 0.02$). The rationale behind combining these curves was to provide a comprehensive overview of how each statistical measure (range, mean, and median) contributes to the overall performance of the diagnostic model.

Pixel ratio has previously been employed in various clinical applications, demonstrating its versatility and effectiveness in different medical contexts. For example, Kalkanis *et al*. proposed a quantitative US approach to assess pleural effusion in 62 subjects, sampling two images per effusion (one axial and one coronal). These images were analyzed using free image editing software to calculate the mean echogenicity value, where each pixel is assigned a number from 0 (total black) to 255 (total white) based on its grayscale. A "hypoechogenicity index" was then calculated by determining the ratio between the mean coronal and axial values. A ratio >1 indicated a hypoechogenic pleural effusion, which correlated with pleural fluid analysis[23].

Tab. 1. Subjective analysis results

	TP	FN	TN	FP	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	PPV (95% CI)	NPV (95% CI)
Pulmonologist A (unblinded)	30	0	25	5	100% (0.89–1.00)	83% (0.66–0.93)	92% (0.82.0.96)	86% (0.71–0.94)	100% (0.87–1.00)
Pulmonologist B (unblinded)	30	0	22	8	100% (0.89–1.00)	73% (0.56–0.86)	87% (0.76–0.93)	79% (0.64–0.89)	100% (0.85–1.00)
Pulmonologist C (blinded)	26	4	26	4	87% (0.70–0.95)	87% (0.70–0.95)	87% (0.76–0.93)	87% (0.70–0.95)	87% (0.70–0.95)
Pulmonologist D (blinded)	24	6	29	1	80% (0.63–0.91)	97% (0.83–1.00)	88% (0.86–0.91)	96% (0.81–1.00)	83% (0.67–0.92)
Cardiologist A	20	10	30	0	67% (0.49–0.81)	100% (0.89–1.00)	83% (0.74–0.93)	100% (0.84–1.00)	75% (0.60–0.86)
Cardiologist B	28	2	22	8	93% (0.79–0.99)	73% (0.56–0.86)	83% (0.74–0.93)	78% (0.62–0.88)	92% (0.74–0.99)
EME A	29	1	22	8	97% (0.83–1.00)	73% (0.56–0.86)	85% (0.76–0.94)	78% (0.63–0.89)	96% (0.79–1.00)
EME B	30	0	25	5	100% (0.89–1.00)	83% (0.66–0.93)	92% (0.85–0.99)	86% (0.71–0.94)	100% (0.87–1.00)
MF surgeon	30	0	13	17	100% (0.89–1.00)	43% (0.27–0.61)	72% (0.60–0.83)	64% (0.50–0.76)	100% (0.77–1.00)
Neurologist	28	2	19	11	93% (0.79-0.99)	63% (0.46–0.78)	78% (0.68–0.89)	72% (0.56–0.84)	91% (0.71–0.98)

CI – confidence interval; EME – emergency medicine expert; MF – maxillofacial; TP – true positive; FN – false negative; TN – true negative; FP– false positive; PPV – positive predictive value; NPV – negative predictive value
Please note that all physicians were blinded to the diagnosis, except for the first two pulmonologists.

Tab. 2. "Targeted" grayscale ratio analysis and calculation of odds ratios using Adobe Photoshop and ImageJ software

Metric	Software	PTX Patients (mean ± SD)	Control Subjects (mean ± SD)	*p*-value	AUC ROC	OR (95% CI, unadjusted)	*p*-value	OR (95% CI, adjusted*)	*p*-value
Mean ratio	Adobe Photoshop	1.53 ± 0.26	1.09 ± 0.07	<0.001	0.98	1.027 (1.010–1.045)	0.002	1.038 (1.006–1.071)	0.021
	ImageJ	1.64 ± 0.35	1.11 ± 0.08	<0.001	0.97	–	–	–	–
Range ratio	Adobe Photoshop	1.93 ± 0.60	1.16 ± 0.16	<0.001	0.89	1.772 (1.303–2.410)	<0.001	1.802 (1.251–2.595)	0.002
	ImageJ	1.79 ± 0.52	1.19 ± 0.20	<0.010	0.87	–	–	–	–
Median ratio	Adobe Photoshop	1.63 ± 0.41	1.12 ± 0.10	<0.001	0.95	1.015 (1.007–1.023)	<0.001	1.021 (1.006–1.035)	0.005
Mode ratio	ImageJ	3.04 ± 5.59	2.00 ± 4.00	0.410	–	–	–	–	–

AUC – area under the curve; CI – confidence interval; OR – odds ratio; PTX – pneumothorax; ROC – receiver operating characteristic; BMI – body mass index
* Adjusted for age, sex, and BMI. ORs are calculated per 0.001 unit increase for mean and median ratios, and per 0.1 unit increase for range ratio.

The significant differences observed between the results obtained with Adobe Photoshop and ImageJ can be attributed to the operator-dependent nature of US examination, as well as the differing methods used for selecting areas under the pleural line. In ImageJ, the selection was made manually with a mouse, while Adobe Photoshop utilized a drawing pad, allowing for more precise delineation.

Since each area consists of millions of pixels, each with a range of values (0–255), the method of area selection can significantly influence statistical outcomes. Larger areas tend to smooth out small-scale variations, while smaller, higher-resolution regions capture more detailed variability. Nevertheless, results from both software programs were satisfactory. Future studies should carefully consider the choice of software, as it can have a substantial impact on outcomes. Furthermore, the method of area selection plays a crucial role and may introduce statistically significant differences; thus, future research should ensure that the selection strategy aligns with study objectives and accurately reflects the underlying data.

Regarding the subjective analysis, diagnostic accuracy varied notably both within and between groups. Excluding the unblinded pulmonologists, experience in emergency medicine seemed to be a key factor behind the higher performance and better inter-operator reliability in that group. Conversely, the two physicians without prior US experience demonstrated the lowest diagnostic accuracy and inter-operator reliability, despite showing good sensitivity and specificity. Interestingly, these two physicians exhibited better inter-operator reliability in evaluating horizontal artifacts, while the other groups were more consistent in recognizing grayscale differences. No significant differences were observed among the five groups, except for diagnostic accuracy, which was borderline, likely reflecting the lack of US experience in the fifth group.

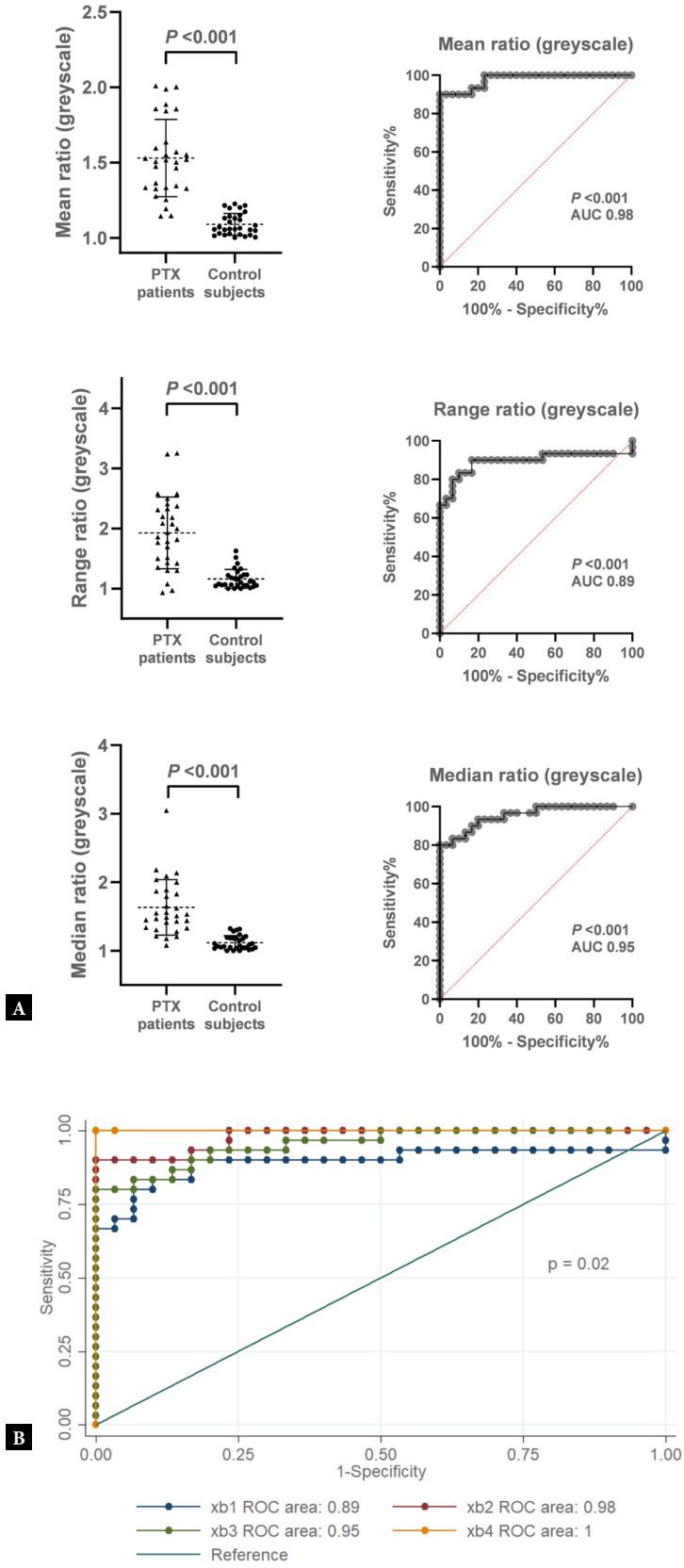

Fig. 3. *"Targeted" sub-analysis using Adobe Photoshop software.* **A.** *Distribution of grayscale mean, range, and median ratios in PTX and control subjects, along with their corresponding ROC curves.* **B.** *Combined ROC curve analysis including mean (xb2), range (xb1), and median (xb3) ratios, demonstrating a near-perfect diagnostic performance when all three metrics are integrated (xb4). PTX – pneumothorax; ROC – receiver operating characteristic; AUC – area under the curve; xb1 – range ratio ROC curve; xb2 – mean ratio ROC curve; xb3 – median ratio ROC curve; xb4 – combined ROC curve.*

These findings support the hypothesis that training physicians to identify PTX should involve subjective comparisons of both lungs, with particular attention to horizontal artifacts and grayscale differences. Increased brightness or grayscale over the lung or A-lines is likely due to posterior enhancement, a phenomenon that occurs when US attenuation decreases in non-solid areas, such as the air trapped in PTX. In suspected cases, even minor differences in these parameters should prompt further investigation, including searching for the lung-point or elasto-lung-point sign or obtaining additional imaging (e.g., chest X-ray or CT scan).

Beyond clinical and anatomical factors influencing grayscale values, several US artifacts may contribute to the observed differences. These include: 1. shadowing and enhancement, which occur when structures with significantly different acoustic impedances cause a decrease or increase in signal strength; 2. reverberation, occurring when sound waves bounce between highly reflective surfaces, such as the pleural line and chest wall; 3. beam hardening, which results from US waves passing through denser structures, such as pleural thickening or adhesions, causing a darker appearance in the affected region; and 4. motion artifacts, caused by patient movement or breathing, which can blur the image.

Understanding these artifacts is crucial for accurately interpreting US data and minimizing potential confounders. By recognizing these artifacts, clinicians can more effectively analyze observed differences in grayscale values. This highlights the importance of consistent imaging techniques and careful consideration of the imaging environment when comparing lung tissue.

Incorporating objective parameters into the diagnostic process may further enhance accuracy, particularly in the context of machine learning and AI-assisted pattern recognition, which have the potential to improve diagnostic precision.

The differences in grayscale values observed in this study offer significant potential for enhancing diagnostic algorithms, particularly in the fields of radiology and imaging analysis. The ability to quantify these differences could contribute to the development of more sensitive and specific algorithms for detecting pathological conditions, such as lung diseases or other abnormalities. By integrating these grayscale variations into automated image analysis systems, physicians may be better able to identify subtle differences in lung tissue that are not easily detectable through traditional visual inspection.

One promising application is in dual-screen modes, which are commonly used by physicians to compare different images or regions of interest in a patient's lungs. By incorporating automated detection of grayscale differences in such setups, diagnostic software could highlight or flag areas where changes in tissue density or texture suggest potential pathology. This capability could enable earlier detection of subtle abnormalities and provide radiologists with a more efficient, evidence-based tool to aid in diagnosis, potentially reducing human error and improving diagnostic consistency. Moreover, automated detection could reduce the time physicians spend analyzing images, supporting faster decision-making, particularly in high-volume clinical settings.

However, a significant limitation arises in cases of bilateral pathology, such as major thoracic trauma (e.g., bilateral PTX) or certain diseases like bilateral pneumonia, where both sides of the lungs may exhibit similar grayscale patterns. In such cases, the algorithm may struggle to differentiate between normal variations and pathological changes, potentially resulting in false negatives or overdiagnosis. This issue is especially important in trauma settings, where bilateral injuries or abnormalities are common. Additionally, the accuracy and reproducibility of grayscale measurements can be influenced by factors such as the quality of imaging equipment, differences in image resolution, or variations in contrast settings. Algorithms must account for these factors to ensure robustness across different clinical environments.

While algorithmic integration of grayscale differences shows considerable promise, there is also a risk of overfitting, where the system may learn to detect subtle differences that are not clinically significant, leading to unnecessary investigations or alerts. Ongoing validation and refinement of these algorithms will be necessary to ensure their robustness across diverse patient populations and clinical settings.

Therefore, while integrating grayscale differences into diagnostic algorithms has significant potential to enhance diagnostic accuracy and efficiency, careful consideration of these limitations is crucial to ensure that the technology is practical, reliable, and clinically useful.

In conclusion, the "targeted" sub-analysis method proved to be the most reliable and efficient approach for consistent image evaluation, minimizing potential errors from manual area tracking. US software could adopt this method, enabling rapid assessment of new diagnostic tools by combining mean, range, and median ratios. The integration of AI or automated software for grayscale comparison in PTX analysis is promising[24]; however, current applications remain limited by small datasets and a lack of external validation[25].

This study has several limitations that must be addressed. First, the sample size was relatively small, and matching cases with controls was not feasible due to the study design. Additionally, only one US machine was used for all image acquisition, which may limit the generalizability of the findings. Second, patients with emphysema, interstitial lung disease, and obesity were excluded, so the findings need validation in a larger and more diverse population, particularly in obese individuals commonly seen in Western countries. Finally, patients with tension PTX and major thoracic trauma were not included, which could affect the applicability of the results in critical care settings. Future studies should aim to incorporate these populations to expand the relevance of the approach in broader clinical contexts.

While nine mesothelioma cases were included, which could have impacted image quality, the location of the tumors was carefully considered, ensuring they were not anterior, thereby minimizing potential image quality issues. Furthermore, at least three static images were captured for each patient, which may have influenced the results. Final image selection was performed by an expert in US imaging. Static images were selected to ensure analytical standardization and minimize inter-observer variability; nonetheless, future studies integrating dynamic imaging (e.g., cine loops) may offer additional diagnostic value and improve generalizability. Despite being performed by trained physicians, image acquisition remains an operator-dependent procedure, as does area tracking. These factors may have contributed to the significant differences observed between software programs. Additionally, the presence of echogenicity

and horizontal artifacts could be influenced by factors beyond PTX, which warrants further exploration.

Regarding the subjective analysis, only 10 physicians were enrolled (two per group), which may explain the lack of significant differences among the five groups, with the exception of positive predictive value.

Conclusions

Simultaneous bilateral static US image comparison, both subjective and objective, between PTX and the healthy contralateral lung appears to be a fast, efficient, and easy-to-use diagnostic approach. This method requires minimal training, even for physicians without prior US experience, making it a promising tool for broad clinical application. Incorporating this diagnostic tool into US software could expedite PTX diagnosis, especially in time-sensitive situations. Moreover, its excellent specificity makes it particularly useful for excluding the diagnosis of PTX. However, further prospective studies are needed to validate its utility and refine its application, particularly in critical care settings.

Conflict of interest

The authors do not report any financial or personal connections with other persons or organizations which might negatively affect the contents of this publication and/or claim authorship rights to this publication.

Author contributions

Original concept of study: GL, GM. Writing of manuscript: GL. Analysis and interpretation of data: GL, RMI, MB, LP, CT, GM. Final acceptation of manuscript: GL, MB. Collection, recording and/or compilation of data: GL, CR, RMI. Critical review of manuscript: GL, CR, RMI, MB, LP, CT, GM.

References

1. Tschopp JM, Bintcliffe O, Astoul P, Canalis E, Driesen P, Janssen J *et al.*: ERS task force statement: diagnosis and treatment of primary spontaneous pneumothorax. Eur Respir J 2015; 46: 321–335. doi: 10.1183/09031936.00219214.
2. Husain LF, Hagopian L, Wayman D, Baker WE, Carmody KA: Sonographic diagnosis of pneumothorax. J Emerg Trauma Shock 2012; 5: 76–81. doi: 10.4103/0974-2700.93116.
3. Metzler EK, Everløff N, Juul AD, Laursen CB, Graumann O, Pietersen PI: Value of focused lung ultrasound in diagnosing and monitoring pneumothorax after CT-guided lung biopsy. J Ultrasound 2025; 28: 11–18. doi: 10.1007/s40477-025-00998-w.
4. Sieber S, Garbe J, Böhm S, Eisenmann S: Pneumothorax detection with thoracic ultrasound as the method of choice in interventional pulmonology – a retrospective single-center analysis and experience. BMC Pulm Med 2023; 23: 227. doi: 10.1186/s12890-023-02511-7.
5. Brims FJ, Davies HE, Lee YC: Respiratory chest pain: diagnosis and treatment. Med Clin North Am 2010; 94: 217–232. doi:10.1016/j.mcna.2010.01.003.
6. Sperandeo M, Filabozzi P, Varriale A, Carnevale V, Piattelli ML, Sperandeo G *et al.*: Role of thoracic ultrasound in the assessment of pleural and pulmonary diseases. J Ultrasound 2008; 11: 39–46. doi: 10.1016/j.jus.2008.02.001.
7. Asrani A, Kaewlai R, Digumarthy S, Gilman M, Shepard JA: Urgent findings on portable chest radiography: what the radiologist should know – review. AJR Am J Roentgenol 2011; 196(6 Suppl): S45–S61. doi: 10.2214/AJR.09.7170.
8. Fraser RS, Müller NL, Colman N, Pare PD: Diagnosis of Diseases of the Chest. 4th ed. Philadelphia: WB Saunders, 1991; p. 2741–2750.
9. Bandelli GP, Levi G, Quadri F, Marchetti GP: "Elasto-lung point": A new tool for the sonographic confirmation of pneumothorax. Clin Respir J 2020; 14: 758–762. doi: 10.1111/crj.13193.
10. Berry L, Rehnberg L, Groves P, Knight M, Stewart M, Dushianthan A: Lung ultrasound in critical care: a narrative review. Diagnostics (Basel) 2025; 15: 755. doi: 10.3390/diagnostics15060755.
11. Quarato CMI, Mirijello A, Bocchino M, Feragalli B, Lacedonia D, Rea G *et al.*: Low diagnostic accuracy of transthoracic ultrasound for the assessment of spontaneous pneumothorax in the emergency setting: a multicentric study. J Clin Med 2024; 13: 4861. doi: 10.3390/jcm13164861.
12. Soldati G, Testa A, Sher S, Pignataro G, La Sala M, Silveri NG: Occult traumatic pneumothorax: diagnostic accuracy of lung ultrasonography in the emergency department. Chest 2008; 133: 204–211. doi: 10.1378/chest.07-1595.
13. Lichtenstein DA, Mezière G, Lascols N, Biderman P, Courret JP, Gepner A *et al.*: Ultrasound diagnosis of occult pneumothorax. Crit Care Med 2005; 33: 1231–1238. doi: 10.1097/01.ccm.0000164542.86954.b4.
14. Slater A, Goodwin M, Anderson KE, Gleeson FV: COPD can mimic the appearance of pneumothorax on thoracic ultrasound. Chest 2006; 129: 545–550. doi: 10.1378/chest.129.3.545.
15. Ball CG, Kirkpatrick AW, Feliciano DV: The occult pneumothorax: what have we learned? Can J Surg 2009; 52: E173–E179.
16. Ding W, Shen Y, Yang J, He X, Zhang M: Diagnosis of pneumothorax by radiography and ultrasonography: a meta-analysis. Chest 2011; 140: 859–866. doi: 10.1378/chest.10-2946.
17. Chung MJ, Goo JM, Im JG, Cho JM, Cho SB, Kim SJ: Value of high-resolution ultrasound in detecting a pneumothorax. Eur Radiol 2005; 15: 930–935. doi: 10.1007/s00330-004-2518-7.
18. Malla M, Purna Shrestha A, Prasad Shrestha S, Shrestha R: Diagnostic accuracy of bedside lung ultrasound in detecting traumatic pneumothorax by novice physicians in the Emergency Department of a Tertiary Care Hospital of Nepal. Emerg Med Int 2024; 2024: 9956637. doi: 10.1155/2024/9956637.
19. Chan KK, Joo DA, McRae AD, Takwoingi Y, Premji ZA, Lang E *et al.*: Chest ultrasonography versus supine chest radiography for diagnosis of pneumothorax in trauma patients in the emergency department. Cochrane Database Syst Rev 2020; 7: CD013031. doi: 10.1002/14651858.CD013031.pub2.
20. ImageJ [computer program]. Version 1.42q; Bethesda (MD): U.S. National Institutes of Health. Available from: https://imagej.net/ij/ [access: 04.05.2025].
21. Russ JC: The Image Processing Handbook. 7th ed. Boca Raton (FL): CRC Press; 2016.
22. Schneider CA, Rasband WS, Eliceiri KW: NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012; 9: 671–675. doi: 10.1038/nmeth.2089.
23. Kalkanis A, Varsamas C, Gourgoulianis K: Correlation of pleural effusions' grayscale sonographic parameters with fluid's analysis results. J Thorac Dis 2017; 9: 543–546. doi: 10.21037/jtd.2017.03.31.
24. Yang C, Zhao H, Wang A, Li J, Gao J: Comparison of lung ultrasound assisted by artificial intelligence to radiology examination in pneumothorax. J Clin Ultrasound 2024; 52: 1051–1055. doi: 10.1002/jcu.23756.
25. Kossoff J, Duncan S, Acharya J, Davis D: Automated Analysis of ultrasound for the diagnosis of pneumothorax: a systematic review. Cureus 2024; 16: e72896. doi: 10.7759/cureus.72896.