Lung Ultrasound Signs to Diagnose and Discriminate Interstitial Syndromes in ICU Patients: A Diagnostic Accuracy Study in Two Cohorts

Micah L.A. Heldeweg (m.heldeweg@amsterdamumc.nl)
Amsterdam UMC Location VUmc

Marry R. Smit
Amsterdam UMC Location AMC

Shelley R. Kramer-Elliott
Amsterdam UMC Location VUmc

Mark E. Haaksma
Amsterdam UMC Location VUmc

Jasper M. Smit
Amsterdam UMC Location VUmc

Laura A. Hagens
Amsterdam UMC Location AMC

Nanon F.L. Heijnen
Maastricht University Medical Centre

Annemijn H. Jonkman
Amsterdam UMC Location VUmc

Frederique Paulus
Amsterdam UMC Location AMC

Marcus J. Schultz
Amsterdam UMC Location AMC

Armand R.J. Girbes
Amsterdam UMC Location VUmc

Leo M.A. Heunks
Amsterdam UMC Location VUmc

Lieuwe D.J. Bos
Amsterdam UMC Location AMC

Pieter R. Tuinman
Amsterdam UMC Location VUmc

Research Article

Keywords: Critical Care, Heart Failure, Lung, Pulmonary Edema, Respiratory Distress Syndrome, Adult, Ultrasonography

Posted Date: January 20th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1261473/v1

Abstract

Background. Diagnosing and discriminating non-cardiogenic interstitial syndrome (NCIS) and cardiogenic pulmonary edema (CPE) is notoriously difficult in a mixed intensive care unit (ICU) population. We hypothesized that a comprehensive lung ultrasound exam can be used to accurately diagnose interstitial syndrome and can discriminate between NCIS and CPE.

Methods. A single center prospective diagnostic accuracy study was used as derivation cohort and a post-hoc analysis of a multi-center prospective observational study as validation cohort. Consecutive adult ICU patients that received a lung ultrasound examination for clinical or study purposes were included. The reference standard was the diagnosis interstitial syndrome (NCIS or CPE) or non-interstitial syndromes (other pulmonary diagnoses and no pulmonary diagnoses) at the moment of examination based on full post-hoc clinical chart review except lung ultrasound. The index test was a comprehensive lung ultrasound examination performed and scored by a researcher blinded to clinical information.

Results. 101 patients were included in the derivation and 125 in validation cohort. In the derivation cohort, patients with interstitial syndrome (n=56) were reliably discriminated from other patients based on the presence of a B-pattern (defined as ≥3 B-lines in one frame) with an accuracy of 94.7% (sens 90.9%, spec 91.1%), while specificity was higher for a bilateral B-pattern (accuracy 86.0%; sens 74.5%, spec 100%). For the discrimination of NCIS (n=29) from CPE (n=27), the presence of bilateral pleural line abnormalities had the highest diagnostic accuracy (94.6%; sens 89.3%, spec 100%) followed by consolidation (76.8%; sens 69.0%, spec 85.2%) and subtle lung sliding (67.9%; sens 62.1%, spec 74.1%). A diagnostic algorithm (BLUISH protocol) using B-pattern and bilateral pleural abnormalities had an accuracy of 0.86 (95%CI 0.77-0.95) for diagnosis and discrimination of interstitial syndromes. In the validation cohort, which included 125 patients with interstitial syndrome, pleural line abnormalities discriminated NCIS (n=101) from CPE (n=24) with a sensitivity of 30% (95%CI 21-40%) and a specificity of 100% (95%CI: 86-100%).

Conclusions. Lung ultrasound can be used to diagnose and discriminate interstitial syndromes in ICU patients with moderate to good accuracy. Pleural line abnormalities are highly specific for NCIS, but sensitivity is limited.

Background

Interstitial syndrome is an important cause of respiratory failure for intensive care unit (ICU) patients. Discriminating between the two most common etiologies of interstitial syndrome, namely non-cardiogenic interstitial syndrome (NCIS) and cardiogenic pulmonary edema (CPE) is notoriously difficult due to their similar clinical manifestations and radiologic appearances, but of clinical importance considering their associated therapeutic and prognostic implications. Despite advances in diagnostic and monitoring instruments, the current gold standard remains post-hoc expert clinical review [1].

Lung ultrasound is an accurate bedside diagnostic tool that can help to differentiate between several causes of respiratory failure [2,3]. In healthy lungs the pleura acts as a specular reflector of ultrasound

beams due to high acoustic impedance disparity with the air-filled lungs beneath. In a diseased lung this impedance disparity is altered, resulting in artefactual or anatomical ultrasound images. Both NCIS and CPE disrupt (regional) acoustic behavior of the pleura, but their distinct pathogenesis alludes to potential ultrasound differences [4]. The 2012 international consensus recommendations suggest a potential role for ultrasound to diagnose and discriminate interstitial syndromes, but advise further research as recommendations are based on one study [5,6]. To date, bar a study using M-Mode, it remains unvalidated [7]. When further substantiated, bedside lung ultrasound may lead to quicker diagnosis and discrimination of interstitial syndromes, which will decrease the need for more invasive and costly diagnostic modalities and facilitate earlier initiation of appropriate therapy [8,9].

This study aims to evaluate the diagnostic accuracy of predefined lung ultrasound signs for the diagnosis of interstitial syndrome and to discriminate NCIS and CPE in ICU patients. We hypothesized that a B-pattern can be used to accurately diagnose interstitial syndrome and consolidations and pleural abnormalities can discriminate between NCIS and CPE.

Methods

Study design

This is a multicenter observational diagnostic accuracy study with a derivation and validation cohort. The protocol for the derivation cohort was approved by the local institutional review board (METc VUmc, registration 2016.002) and necessity to obtain informed consent was waived. The protocol for the validation cohort was approved by its institutional review board (METc AMC, registration W18_311). Written informed consent for use of data was obtained from the patient or the legal representative for the validation cohort. The STARD checklist was used to draft this manuscript (EQUATOR network, 2015).

Participants

The derivation cohort included prospectively collected adult (>18 years) patients admitted to the mixed ICU of the Amsterdam UMC, location VUmc, an academic hospital in The Netherlands, prospectively between 1st of January 2018 and 1st of August 2020. Patients were included when they received a clinically indicated lung ultrasound examination (as determined by the clinical team).

The validation cohort consisted of a post-hoc analysis of an observational, prospective study performed at the mixed ICU's of the Amsterdam UMC, location AMC and the Maastricht UMC+, both academic hospitals located in The Netherlands, between 26th of March 2019 until 26th of February 2021. Patients were included when their expected ventilation duration was >24 hours.

Test methods: index test

For the derivation cohort, all images were acquired or supervised by lung ultrasound certified clinicians, using a Sonosite-EDGE II ultrasound machine. Certification entailed a two-day course and thereafter

supervision by a physician with extensive ultrasound experience (>5 years) until sufficient expertise was reached (a minimum of 30 exams) [10]. All subjects were scanned using a standardized protocol based on the Bedside Lung Ultrasound in Emergency (BLUE) protocol, consisting of two anterior and one posterolateral point, on either side of the thorax [11]. Anterior measurements were performed in B-mode and M-mode using a 10-5 MHz linear transducer with image depth set >6cm in lung setting, whereas posterolateral measurements were performed with a 5-1 MHz cardiac (phased array) transducer in cardiac setting. All patients were scanned in supine position. As per our local clinical protocol patients were scanned in supine position and with probe perpendicular to rib orientation.

All 2D and M-mode images and clips were independently evaluated offline by two investigators (MLAH and SKE) blinded to clinical diagnosis. Any disputes were resolved by a third investigator (MEH). In each anterior zone the following 2D items were evaluated using the linear transducer in accordance with international evidence-based recommendations [6,12]: 1. presence of a B-pattern (≥3 B-lines in a single frame); 2. if a B-pattern was present whether its appearance was nonhomogeneous (discordant B-line appearance with asymmetric distribution) 3. presence of an anterior consolidation defined as a subpleural echo-poor region or one with tissue-like echotexture; 4. lung sliding as present, subtle (identified by clinical “gestalt”), or absent; and 5. pleural line characterization as fragmented (disruption of the pleural line), thickened (pleural line thickness >2 mm), and irregular (coarse pleural line lacking well defined borders). Figure 1 shows the morphology of specific pleural line abnormalities and examples of their appearance in CPE and NCIS. The M-Mode images of each anterior zone were evaluated for the following items [7]: 6. pleural line characterization as normal, fragmented, or sinusoidal; and 7. orientation of the subpleural area as horizontal or vertical. Each posterolateral zone was evaluated using the phased array transducer for 8. the presence of posterolateral alveolar and/or pleural syndrome (PLAPS) (11). When PLAPS was positive, the presence of a 9. pleural effusion >1cm was also recorded.

Lung ultrasound examinations in the validation cohort were performed 1-3 days after start of mechanical ventilation by one of three dedicated researchers (LAH, NFLH & MRS). Four anterior regions were analyzed in the present study, corresponding to the analysis of the derivation cohort. Images were obtained with the linear probe using a transversal approach between the ribs. Pleural line abnormalities were scored in all images and defined as a pleura that is evidently altered through (a combination of) fragmentation, thickening, or irregularity.

Test methods: reference standard

In the derivation cohort the reference standard was the determined by expert consensus. Two investigators (MLAH and SKE) independently annotated patients with NCIS, CPE, other pulmonary diagnoses (‘other’), and no pulmonary diagnoses (‘healthy’) at the moment of examination based on full post-hoc clinical chart review including imaging, laboratory data, physical examination, and other clinical characteristics except lung ultrasound. Any disputes were resolved by a third investigator (PRT). Interstitial syndrome was defined as the composite group of NCIS and CPE. The NCIS group included patients with interstitial lung disease (ILD) based on pulmonologist diagnosis in concert with imaging or

serology, or acute respiratory distress syndrome (ARDS) without clinical evidence of systolic or diastolic cardiac dysfunction and using the Berlin definition [13]. The CPE group includes patients with a clinical diagnosis of decompensated systolic or diastolic heart failure, based on imaging, including echocardiography. The non-interstitial syndrome group was defined as the composite group of 'other' and 'healthy'; other were patients with non-interstitial pulmonary pathology (atelectasis, pneumonia, or pleural effusion) and healthy were patients without pulmonary pathology. Patients with an overlapping diagnosis of NCIS and CPE were excluded from the study. Patients were also excluded if a pneumothorax was present, as it impedes normal view of the (visceral) pleura and M-mode evaluation.

In the validation cohort the NCIS group consisted of patients with ARDS, that was scored by an expert panel using the Berlin definition [13]. Patients with ARDS were only analyzed in the present study when the expert panel was certain about the ARDS diagnosis. CPE was additionally scored by one expert (LDJB). The groups 'other' and 'healthy' were not specified in the validation cohort as data on these diagnoses was not prospectively collected.

Outcomes and analysis

The primary outcome of the study was to identify which predetermined lung ultrasound signs were most accurate to diagnose and discriminate interstitial syndromes. To address this, the following comparisons were made: 1. patients with interstitial syndrome versus the non-interstitial group; and 2. patients with NCIS versus CPE. Novel findings from the derivation cohort were corroborated in the validation cohort. To further evaluate clinical applicability, secondary outcomes were interrater agreement of ultrasound signs and constructing a clinical diagnostic algorithm to arrive at the correct pulmonary diagnosis. In both cohorts, baseline, laboratory, ventilator characteristics, and the sequential organ failure assessment (SOFA) score were collected from the electronic patient database at the time of ultrasound examination. Prevalence of baseline characteristics and lung ultrasound signs across groups were tested for significant differences using Pearson's chi-squared test. All data were analyzed with SPSS for Windows (version 22 IBM) and R studio (version 4.0.3). Continuous variables were presented as means ± standard deviations (±SD), medians and interquartile range [IQR], or numbers (percent %) when appropriate. A Shapiro-Wilk's test, visual inspection of histograms, and Q-Q plots were used to determine data distribution.

Diagnostic accuracy.

Diagnostic accuracy parameters (sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio) of 2D and M-mode ultrasound signs were calculated for diagnosis of interstitial syndrome and for discrimination of NCIS and CPE. Specific pleural line abnormalities (fragmented, thickened, and irregular) were tested for internal consistency across items with a two-way random intraclass correlation model with average measures of absolute agreement to evaluate whether compilation was possible. In case of excellent consistency (intraclass correlation coefficient >0.8 [14]) between specific pleural abnormalities, diagnostic accuracy parameters for the one, two, or three pleural line abnormalities simultaneously, in any ultrasound zone or bilaterally, were calculated.

Interrater agreement.

Agreement for dichotomous and ordinal variables as derived by the two observers was evaluated with a kappa statistic and a spearman's correlation coefficient respectively. In addition, a joint probability of agreement was calculated of all variables as follows: (number of cases in agreement / total number of cases).

Diagnostic algorithm. For the derivation cohort, a hierarchical diagnostic algorithm was built based on substantial reliability (defined as kappa statistic >0.60 or spearman's correlation coefficient >0.70 [15,16]) and substantial accuracy (defined as both a positive and negative likelihood ratio of >4.0 and <0.30 respectively, indicating a clinically useful shift in disease probability of at least 25% [17]). The initial step was diagnosis of interstitial syndromes, and thereafter discrimination of NCIS and CPE, as well as other and healthy patients respectively.

Sample size

A sample size calculation was made using a chi-square contingency table. Based on previous literature and expert opinion the proportion of pleural line abnormalities was estimated to be 0.9 in the NCIS group and 0.1 in the CPE group [7,12]. At least twenty-one patients per group would be required assuming an α of 0.05 and a β of 0.20. Patients were collected until each group reached the required sample size. This study was powered for pleural line abnormalities because previous literature indicated their importance for discrimination, whilst anterior consolidations and B-pattern are already well-established signs of NCIS and CPE respectively. No sample size calculation was performed for the validation cohort.

Post-hoc analysis

To further elucidate differences between the derivation and validation cohort a post-hoc analysis of the NCIS group was performed and reported separately.

Results

In the derivation cohort, a total of 110 patients were included until the sample size for each group was reached. Nine patients were excluded due to pneumothorax (n=5) or simultaneous occurrence of CPE and NCIS (n=4). One-hundred and one patients were finally included and a total of 1010 ultrasound files were examined (404 anterior clips, 202 posterolateral clips, and 404 M-Mode images). Twenty-seven (2.7%) of these files were deemed of insufficient quality, leaving 983 for analysis. Characteristics of patients at time of ultrasound examination are shown in Table 1. Sixty-eight percent was male and mean age was 64(±14.5) years. NCIS patients were more frequently intubated, had a lower median PaO2/FiO2, and a higher median C-reactive protein; CPE patients were found to have a higher median brain natriuretic peptide level.

Table 1. Characteristics of included patients of the derivation cohort at time of ultrasound examination

Diagnosis	Total (n=101)	NCIS (n=29)	CPE (n=27)	Other (n=24)	Healthy (21)	*P*-value
Age	64.0 ± 14.5	60.0 ± 14.7	63.7 ± 16.1	68.6 ± 11.7	64.6 ± 14.2	.145
Gender (male)	69 (67.6%)	19 (65.5%)	16 (59.3%)	18 (75.0%)	15 (71.4%)	.649
Diagnosis		12 ILD 17 ARDS		9 Pneumonia 8 Pleural Effusion 7 Atelectasis	No pulmonary diagnosis	
Intubated	59 (58.4%)	25 (86.2%)	15 (55.6%)	7 (29.2%)	12 (57.1%)	<.001
P/F ratio	196 [172]	134 [94]	185 [164]	206 [94]	323 [143]	<.001
SOFA	8 [7]	9 [5]	9 [8]	7 [10]	6 [12]	.037
CRP	73 [131]	99 [191]	66 [127]	93 [118]	37 [101]	.013
BNP	1771 [9441]	1081 [4582]	8815 [19441]	1627 [160]	270 [1526]	.001
Clips and frames						
Ventral	401	114	108	95	84	-
M-mode	392	113	102	94	83	-
PL-clips	190	58	52	45	35	-

BNP brain natriuretic peptide; CPE cardiogenic pulmonary edema; CRP C-reactive protein; COPD chronic obstructive pulmonary disease; ILD interstitial lung disease; NCIS noncardiogenic interstitial syndrome; P/F ratio between arterial oxygen saturation and oxygen fraction of inspired air; SOFA sequential organ failure assessment; PL-clips; posterolateral clips; Continuous variables were presented as means ± standard deviations (±SD), medians and interquartile range [IQR], or numbers (percent %) when appropriate.

The distribution of ultrasound signs across diagnostic groups is shown in Additional file 1. The prevalence of B-pattern was different for the interstitial compared to the non-interstitial groups. Anterior consolidation, abnormal lung sliding, and pleural abnormalities had a different prevalence for the NCIS group compared to the CPE group. Any ultrasound sign appearing bilaterally had a different prevalence in the interstitial groups compared to the non-interstitial group, except for M-mode subpleural vertical orientation, which was not differently distributed across diagnostic groups. In the validation cohort, 125

patients with interstitial syndrome were included. Baseline characteristics for the validation cohort are shown in Additional file 2.

Diagnosis of interstitial syndrome in derivation cohort

Table 2 shows the diagnostic accuracy parameters of ultrasound signs for interstitial syndrome (versus non-interstitial). Only signs with a sensitivity and specificity of >60% for diagnosis or discrimination are presented (full table of estimates per group is shown in Additional file 3; full table of estimates for interstitial versus noninterstitial, NCIS versus CPE, and other versus healthy is shown in Additional file 4). The intraclass correlation coefficient for degree of resemblance among pleural line abnormalities was 0.87 (95%CI; .84-.90; p<.001) indicating excellent agreement [14]. Therefore, diagnostic accuracy parameters of one, two, or three composite pleural line abnormalities were also calculated and reported. For diagnosis of interstitial syndrome, the presence of a B-pattern had the highest diagnostic accuracy; specificity increased to 100% when observed bilaterally.

Table 2. Diagnostic accuracy of predetermined ultrasound signs for interstitial (versus non-interstitial syndrome) in the derivation cohort

	Interstitial vs non-interstitial				
	Accuracy	Sens	Spec	+LR	-LR
B-pattern					
Any	**91.0 (83.6-95.8)**	**90.9 (80.1-97.0)**	**91.1 (78.8-97.5)**	**10.21 (4.0-26.1)**	**0.10 (0.04-0.23)**
Nonhomogenous	73.2 (59.7-84.2)	72.5 (58.3-84.1)	80.0 (28.4-99.5)	3.13 (0.62-21.1)	0.47 (0.18-0.64)
Bilateral	**86.0 (77.6-92.1)**	**74.5 (61.0-85.3)**	**100 (92.1-100)**	**∞**	**0.26 (0.16-0.40)**
Anterior consolidation (any)	56.4 (46.2-66.3)	42.9 (29.7-56.8)	73.3 (58.1-85.4)	1.60 (0.91-2.85)	0.78 (0.58-1.04)
Lung sliding (subtle)	61.6 (51.3-71.2)	44.6 (31.3-58.5)	83.7 (69.3-93.2)	2.73 (1.31-5.74)	0.66 (0.50-0.87)
Pleural line abnormalities					
Fragmented	69.3 (59.3-78.1)	73.2 (59.7-84.2)	64.4 (48.8-78.1)	2.06 (1.35-3.15)	0.42 (0.26-0.67)
Thickened	68.3 (58.3-77.2)	51.8 (38.0-65.3)	88.9 (76.0-96.3)	4.67 (1.96-11.1)	0.54 (0.41-0.73)
Irregular	68.3 (58.3-77.2)	60.7 (46.8-73.5)	77.3 (62.9-88.8)	2.67 (1.52-4.91)	0.51 (0.35-0.72)
One any zone	66.3 (56.3-75.4)	71.4 (57.8-82.7)	60.0 (44.3-74.3)	1.79 (1.20-2.65)	0.48 (0.30-0.77)
One bilateral	69.7 (59.7-78.5)	52.7 (38.8-66.4)	90.9 (78.3-97.5)	5.80 (2.20-15.3)	0.50 (0.39-0.70)
Two any zone	68.3 (58.3-77.2)	58.9 (45.0-71.9)	80.0 (65.4-90.4)	2.95 (1.58-5.50)	0.51 (0.36-0.73)
Two bilateral	67.7 (57.6-76.8)	45.5 (32.0-59.5)	95.5 (84.5-99.4)	10.11 (2.50-39.9)	0.57 (0.45-0.73)
Three any zone	69.3 (59.3-78.1)	50.0 (36.3-63.6)	93.3 (81.7-98.6)	7.46 (2.44-23.1)	0.54 (0.41-0.70)
Three bilateral	61.6 (51.3-71.2)	30.9 (19.1-44.8)	100 (92.0-100)	∞	0.69 (0.58-0.82)
PLAPS (bilateral)	64.1 (53.3-73.9)	66.7 (52.6-78.9)	61.1 (43.5-76.9)	1.71 (1.09-2.69)	0.55 (0.34-0.86)
M-mode fragmentation	58.7 (48.0-68.9)	54.0 (39.3-68.2)	64.3 (48.0-78.5)	1.51 (0.94-2.44)	0.72 (0.49-1.04)

-LR negative likelihood ratio; +LR positive likelihood ratio; PLAPS posterolateral alveolar and/or pleural syndrome; Sens sensitivity; Spec specificity. Bold numbers were those with both a +LR and –LR of >4.0 and <0.30 respectively, based on a clinically useful shift in disease probability of 25% (17).

Discrimination between NIC and CPE in derivation cohort

Table 3 (and Additional file 4) shows the diagnostic accuracy parameters of ultrasound signs for discrimination between NCIS and CPE. Pleural line abnormalities were found to have the highest diagnostic accuracy (91.0%; 95%CI 83.6-95.8). When compiling abnormalities per zone, justified by their consistency, bilateral presence of at least two pleural line abnormalities had the highest combined diagnostic accuracy for discrimination of NCIS from CPE (94.6%; 95%CI 84.9-98.9%). Specificity of two pleural line abnormalities increases as more zones are positive (figure 2). Additional file 5 shows diagnostic accuracy for the subgroup of mechanically ventilated patients. The diagnostic accuracy in this subgroup remained the highest for B-pattern (94.7%; 95%CI 85.3-98.9) and bilateral presence of at least two pleural line abnormalities (92.3%; 95%CI 79.1-98.4) for diagnosis and discrimination of interstitial syndrome, respectively.

Table 3. Diagnostic accuracy of predetermined ultrasound signs for NCIS (versus CPE) in the derivation cohort

	NCIS vs CPE				
	Accuracy	Sens	Spec	+LR	-LR
B-pattern					
Any zone	49.1 (35.4-62.9)	89.3 (71.8-97.7)	7.4 (0.9-24.3)	0.96 (0.82-1.14)	1.45 (0.26-7.99)
Nonhomogenous	78.4 (34.9-75.6)	100 (86.8-100)	56.0 (34.9-75.6)	2.27 (1.46-3.54)	0
Bilateral	50.9 (37.1-64.7)	75.0 (55.1-89.3)	25.9 (11.1-46.3)	1.01 (0.74-1.38)	0.97 (0.74-1.38)
Anterior consolidation (any)	76.8 (63.6-87.0)	69.0 (49.2-84.7)	85.2 (663-95.8)	4.66 (1.82-11.9)	0.36 (0.21-0.64)
Lung sliding (subtle)	67.9 (54.0-79.7)	62.1 (42.3-79.3)	74.1 (53.7-88.9)	2.39 (1.19-4.81)	0.51 (0.31-0.86)
Pleural line abnormalities					
Fragmented	78.6 (65.6-88.4)	100 (88.1-100)	55.6 (35,3-74.5)	2.25 (1.48-3.43)	0
Thickened	**92.9 (82.7-98.0)**	**93.1 (77.2-99.2)**	**92.6 (75.7-99.1)**	**12.7 (3.30-47.9)**	**0.07 (0.02-0.28)**
Irregular	**91.1 (80.4-97.0)**	**100 (88.1-100)**	**81.5 (61.9-93.7)**	**5.40 (2.45-11.9)**	**0**
One any zone	76.8 (63.6-87.0)	96.6 (82.2-99.9)	55.6 (35.3-74.5)	2.17 (1.42-3.33)	0.06 (0.01-0.44)
One bilateral	**87.3 (75.5-94.7)**	**89.3 (71.8-97.7)**	**85.2 (66.3-95.8)**	**6.03 (2.4-15.0)**	**0.13 (0.04-0.37)**
Two any zone	**92.9 (82.7-98.0)**	**100 (88.1-100)**	**85.2 (66.3-95.8)**	**6.76 (2.73-16.7)**	**0**
Two bilateral	**94.6 (85.0-98.9)**	**89.3 (71.8-97.7)**	**100 (87.2-100)**	∞	**0.11 (0.04-0.31)**
Three any zone	**91.1 (80.4-97.0)**	**89.7 (72.7-97.8)**	**92.6 (75.6-99.1)**	**12.1 (3.17-46.2)**	**0.11 (0.04-0.33)**
Three bilateral	80.0 (67.0-89.6)	60.7 (40.6-78.5)	100 (87.2-100)	∞	0.39 (0.25-0.62)
PLAPS (bilateral)	35.2 (22.7-49.4)	51.7 (32.5-70.6)	16.0 (4.5-36.1)	0.61 (0.42-0.91)	3.02 (1.14-7.99)
M-mode fragmentation	80.0 (66.3-90.0)	81.5 (61.9-93.7)	78.3 (56.3-92.5)	3.76 (1.69-8.31)	0.24 (0.10-0.54)

-LR negative likelihood ratio; +LR positive likelihood ratio; CPE cardiogenic pulmonary edema; NCIS noncardiogenic interstitial syndrome; PLAPS posterolateral alveolar and/or pleural syndrome; Sens sensitivity; Spec specificity. Bold numbers were those with both a +LR and –LR of >4.0 and <0.30 respectively, based on a clinically useful shift in disease probability of 25% (16).

Discrimination between NCIS and CPE in validation cohort

The application of bilateral pleural line abnormalities in the validation cohort showed a sensitivity of 30% (95%CI 21-40%) and a specificity of 100% (95%CI 86-100%). The positive and negative likelihood ratios were 'infinity' and 0.70 (95%CI 0.62-0.80), respectively.

Secondary outcome: interrater agreement

All ultrasound signs had at least a substantial strength of interrater agreement, except for a nonhomogenous B-pattern and M-mode pleural line evaluation which had fair and moderate agreement respectively (Additional file 6).

Secondary outcome: diagnostic algorithm

The high diagnostic accuracy and interrater agreement of ultrasound signs enabled the development of a diagnostic algorithm; presented as the Bedside Lung Ultrasound for Interstitial Syndrome Hierarchy (BLUISH)-protocol (figure 3). First, the presence of any B-pattern was selected to diagnose interstitial syndromes, and the presence of bilateral pleural line abnormalities (at least 2 out of 3 specific pleural abnormalities present in a zone) was subsequently used to discriminate NCIS from CPE for a total accuracy of 0.86 (95%CI 0.79-0.93). Although not part of our primary outcome, diagnostic accuracy parameters to discriminate other non-interstitial pulmonary pathology from healthy lungs using anterior consolidation and positive PLAPS were included (Additional file 4). When restricting the diagnostic algorithm to mechanically ventilated patients its accuracy remained 0.86 (95%CI 0.77-0.95).

Post-hoc analysis

The NCIS group of the derivation cohort contained 12 ILD patients (41.4%) whereas the NCIS of the validation cohort contained exclusively ARDS. A logistic regression analysis on the derivation cohort showed an that bilateral B-pattern had an odds ratio of 0.067 (95%CI .007-.678) for ILD (versus ARDS) and total pleural abnormalities had an odds ratio of 2.99 (95%CI 1.1-8.3) for ILD (versus ARDS).

Discussion

The main findings of this diagnostic accuracy study on lung ultrasound to diagnose and discriminate interstitial syndromes in ICU patients are: 1. B-pattern is the most accurate ultrasound sign to diagnose interstitial syndromes; 2A. Bilateral pleural line abnormalities is the most accurate ultrasound sign to discriminate NCIS from CPE; 2B. In the validation cohort, bilateral pleural line abnormalities had a specificity of 100% for the differentiation of NCIS from CPE, but limited sensitivity; 3. Interrater agreement

for aforementioned ultrasound signs is excellent and substantial, respectively; 4. An ultrasound diagnostic algorithm (the BLUISH protocol) can diagnose and discriminate interstitial syndromes in critically ill patients with a high accuracy of 0.86.

Discriminating NCIS from CPE is notoriously difficult but of major clinical importance. Ultrasound signs to discriminate NCIS from CPE have been included in expert consensus recommendations despite paucity of evidence [6]. The only prior study on this subject has a small and selected population, lacks sample size justification, has a single derivation cohort, and is inapplicable in a complex ICU setting. Similarly, recently described M-mode signs of NCIS lack validation [7]. We comprehensively evaluated and validated the diagnostic accuracy of NCIS-specific ultrasound signs in ICU patients and show that lung ultrasound is an accurate tool in this regard. Evaluating these signs in ultrasound examinations could lead to quicker, bedside, arrival at diagnosis and facilitates timely and appropriate treatment. A patient with CPE is typically treated with diuretics and pre- or afterload reduction, and may require treatment for underlying coronary pathology. NCIS patients require appropriate identification and management of concealed underlying etiology and consideration of systemic therapy such as corticosteroids [18]. As such, prompt utilization of lung ultrasound can reduce uptake of more invasive or costly monitoring tools such as pulmonary artery catheters or chest computed tomography. Even more so, it could provide a framework for an improved clinical definition of ARDS [19].

Based on the most accurate ultrasound signs, we developed a diagnostic algorithm, the BLUISH protocol, for use in ICU patients. The BLUISH protocol contains simple and often used sonographic signs, which will allow for rapid clinical implementation. Its high diagnostic accuracy of 86% is comparable to the accuracy of the landmark BLUE-protocol in emergency department patients. Addition of non-ultrasound, but readily available, clinical information may further improve its accuracy, for example C-reactive protein and brain natriuretic peptide were different between diagnoses.

Inconsistent definitions and methodology are a persistent issue in lung ultrasound literature, especially concerning pleural line abnormalities [20]. The consensus for specific pleural line abnormalities is categorizing it as ‘irregular’, ‘fragmented’, and ‘thickened’, but other terms such as blurred, coarse, and tightening are also frequently used, albeit equally ill-defined [6,21–23]. Consequently, reproducibility and generalizability of studies is limited. The current study’s explicit *a priori* definitions generated substantial interrater agreement, increasing external validity. In addition, specific pleural abnormalities carry distinct diagnostic accuracies; thickening being the most accurate for NCIS, followed by irregular and fragmented. Considering their consistency across items it is reasonable to compile pleural line abnormalities for bedside decisions. After testing many variables it appears that the dichotomy of normal or abnormal pleural line is most appropriate to balance clinical applicability and efficacy of the test. Future investigations should explore whether specific pleural line abnormalities have characteristic spectral signatures and distinct histopathological correlates or belong on a continuum of aeration amongst A-lines, B-lines, and consolidation [4,24].

External validation demonstrates that clinically evident bilateral pleural abnormalities could be an excellent sign to rule in NCIS, with a specificity of 100%. The lower sensitivity found in validation cohort may be due to several differences between the cohorts. First, 41% of included NCIS patients in the derivation cohort were classified as ILD, whilst the validation cohort did not include ILD patients. Post-hoc analyses showed that ILD had a higher propensity for pleural abnormalities and lower propensity for B-pattern when compared to ARDS. Second, scoring of the pleura line was different between the cohorts. In the derivation cohort all images were scored by two raters and disputes concerning pleural abnormalities were resolved by a third rater while in the validation cohort the pleura was scored by a single rater. Lastly, the validation cohort performed transversal lung ultrasound scanning, which increases visible pleural surface when compared to examination perpendicular to the ribs. Despite differences, this finding is very useful for the derivation of refined ARDS definitions and clinical diagnosis of NCIS, but does not allow for ruling out of CPE. This may also be because initial CPE presentations, emerging extremely rapid or existing for an extended period, may convert to an inflammatory syndrome with secondary permeability edema due to endothelial cell dysfunction, but this effect cannot occur vice-versa [25].

Other ultrasound signs of potential interest are spared areas and lung pulse. These measurements were not included in the current study [12]. We did evaluate nonhomogenous quality of B-pattern, and although its diagnostic accuracy parameters were reasonable, agreement between offline reviewers was ‘fair’, making it less clinically useful. These results are in line with other literature finding low agreement when evaluating or classifying the distribution of B-pattern [26,27]. In contrast to results of a recent study, our study shows that use of M-mode on lung ultrasound does not offer an advantage when compared to 2D pleural line examination. Though M-mode ultrasound can be used to detect other thoracic abnormalities (pneumothorax and pleural effusion), its usefulness to discriminate NCIS from CPE appears limited [2].

This study has several limitations. There is a risk of selection bias as the patient population was derived from patients with a clinical indication for bedside ultrasound. Besides a relatively high ILD population, baseline characteristics of our population are comparable to that of the case mix of ICU patients in The Netherlands [28]. The additional validation on an external cohort further diminished this limitation. There are differences in cohort design, such as timing of ultrasound examination, examination protocol, and center-specific population, but these can be considered within the scope of normal between-center variation and therefore further increase external validity. It should be noted that this study evaluates the diagnostic accuracy of singular lung ultrasound signs in a deterministic fashion, whereas clinician-based practice in a complex ICU setting often relies on clinical conglomerates. This should be considered in future research and during clinical application [3]. Finally, patients with an uncertain ARDS diagnosis were excluded from the validation cohort, possibly leading to selection bias. This study also has several strengths. Our study is adequately powered and, even more so, the largest study to date on this subject. Compared to previous research, we increased external validity for a general ICU population by adding ‘other’ and ‘healthy’ patient populations as control groups. Moreover, performed an external validation of previously determined ultrasound signs which had not been done before and constructed the BLUISH protocol based on straightforward ultrasound signs increasing this algorithm’s generalizability.

Conclusion

This diagnostic accuracy study in ICU patients shows that lung ultrasound is a valuable tool to diagnose and discriminate interstitial syndromes. Pleural line abnormalities, anterior consolidations, and subtle lung sliding have high diagnostic accuracy for NCIS and substantial to excellent interrater agreement. Our novel diagnostic algorithm (BLUISH protocol), containing a B-pattern and bilateral pleural line abnormalities, has an accuracy of 86% for diagnosing and discriminating the cause of interstitial syndrome. External validation of these findings demonstrates that pleural line abnormalities have a perfect specificity for NCIS, but limited sensitivity to rule out CPE.

Abbreviations

ARDS: Acute Respiratory Distress Syndrome

BLUE: Bedside Lung Ultrasound for Emergency

BLUISH: Beside Lung Ultrasonography for Interstitial Syndrome Hierarchy

CPE: Cardiogenic Pulmonary Edema

ICC: Intraclass Correlation Coefficient

ILD: Interstitial Lung Disease

IQR: Interquartile Range

NCIS: Non-cardiogenic interstitial syndrome

PLAPS: Postero Lateral Alveolar and/or Pleural Syndrome

SD: Standard Deviation

Declarations

Ethics approval and consent to participate

The protocol for the derivation cohort was approved by the local institutional review board (METc VUmc, registration 2016.002) and necessity to obtain informed consent was waived. The protocol for the validation cohort was approved by its institutional review board (METc AMC, registration W18_311). Written informed consent for use of data was obtained from the patient or the legal representative for the validation cohort.

Consent for publication Consent for publication was waived by the local ethics board.

Availability of data and material The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests The authors declare that they have no competing interests

Funding No funding was received for this study.

Authors' contributions:

MLAH, SKE, MEH, JS, AJ, FP, MS, AG, LH, and PT were responsible for the conception and design of the work. MLAH, SKE, MEH, and PT were responsible for acquisition and analysis of the data for the derivation cohort. MS, LH, NH, LB, were responsible for the acquisition and analysis of the data for the validation cohort. MLAH and SKE were responsible for building the database. MLAH and PT were responsible for drafting the manuscript and all authors provided critical revisions for it. All authors read and approved the final manuscript and ensured that questions related to the accuracy or integrity of any part of the work were investigated and resolved.

Acknowledgements

We thank assistant professor and biostatistician Peter van de Ven (Amsterdam Cardiovascular Sciences research institute, the Netherlands) for review of, and advice on, the statistical analysis.

References

1. Schmickl CN, Shahjehan K, Li G, et al. Decision support tool for early differential diagnosis of acute lung injury and cardiogenic pulmonary edema in medical critically ill patients. Chest. 2012;141(1):43–50.
2. Mojoli F, Bouhemad B, Mongodi S, Lichtenstein D. Lung ultrasound for critically ill patients. Am J Respir Crit Care Med. 2019 Mar 15;199(6):701-714.
3. Smit JM, Haaksma ME, Winkler MH, et al. Lung ultrasound in a tertiary intensive care unit population: a diagnostic accuracy study. Crit Care. 2021 Sep 17;25(1):339.
4. Soldati G, Smargiassi A, Demi L, Inchingolo R. Artifactual Lung Ultrasonography: It Is a Matter of Traps, Order, and Disorder. Appl Sci 2020; 10(5):1570
5. Soldati G, Demi M, Demi L. Ultrasound patterns of pulmonary edema. Ann Transl Med. 2019;7(S1): S16.
6. Volpicelli G, Elbarbary M, Blaivas M, et al. International evidence-based recommendations for point-of-care lung ultrasound. Intensive Care Med. 2012;38(4):577–91.
7. Singh AK, Mayo PH, Koenig S, Talwar A, Narasimhan M. The Use of M-Mode Ultrasonography to Differentiate the Causes of B Lines. Chest. 2018;153(3):689–96
8. Brogi E, Bignami E, Sidoti A, et al. Could the use of bedside lung ultrasound reduce the number of chest x-rays in the intensive care unit? Cardiovasc Ultrasound. 2017;15(1):23.

9. Rodriguez-Fanjul J, Jordan I, Balaguer M, Batista-Muñoz A, Ramon M, Bobillo-Perez S. Early surfactant replacement guided by lung ultrasound in preterm newborns with RDS: the ULTRASURF randomised controlled trial. Eur J Pediatr. 2020 Dec 1;179(12):1913–20
10. Touw HRW, Tuinman PR, Gelissen HPMM, Lust E, Elbers PWG. Lung ultrasound: Routine practice for the next generation of internists. Neth J Med. 2015 Mar;73(3):100–7.
11. Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure the BLUE protocol. Chest. 2008 Jul;134(1):117–25.
12. Copetti R, Soldati G, Copetti P. Chest sonography: A useful tool to differentiate acute cardiogenic pulmonary edema from acute respiratory distress syndrome. Cardiovasc Ultrasound. 2008 Apr 29;6:16.
13. Ranieri VM, Rubenfeld GD, Thompson BT, et al. Acute respiratory distress syndrome: The Berlin definition. JAMA - J Am Med Assoc. 2012 Jun 13;307(23):2526–33.
14. Cicchetti D V. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. Psychol Assess. 1994;6(4):284–90.
15. Schober P, Boer C, Schwarte LA. Correlation Coefficients. Anesth Analg. 2018;126(5):1763–8.
16. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159.
17. McGee S. Simplifying likelihood ratios. J Gen Intern Med. 2002;17(8):647–50.
18. Ware LB, Matthay MA. Acute Pulmonary Edema. N Engl J Med. 2005;353(26):2788–96.
19. Riviello ED, Kiviri W, Twagirumugabe T, et al. Hospital incidence and outcomes of the acute respiratory distress syndrome using the Kigali modification of the Berlin definition. Am J Respir Crit Care Med. 2016 Jan 1;193(1):52–9.
20. Neuteboom OB, Heldeweg ML, Pisani L, et al. Assessing Extravascular Lung Water in Critically Ill Patients Using Lung Ultrasound: A Systematic Review on Methodological Aspects in Diagnostic Accuracy Studies. Ultrasound Med Biol. 2020 Jul;46(7):1557-1564
21. Manolescu D, Oancea C, Timar B, et al. Ultrasound mapping of lung changes in idiopathic pulmonary fibrosis. Clin Respir J. 2020 Jan 1;14(1):54–63.
22. Saraogi A. Lung ultrasound: Present and future. Lung India. 2015 May-Jun;32(3):250-7.
23. Buda N, Piskunowicz M, Porzezińska M, Kosiak W, Zdrojewski Z. Lung ultrasound in the evaluation of interstitial lung diseases in the course of connective tissue diseases. Eur Resp J. 2017 50:PA1520.
24. Lichtenstein DA. Current Misconceptions in Lung Ultrasound: A Short Guide for Experts. Chest. 2019 Jul 1;156(1):21–5.
25. Huang D, Ma H, Xiao Z, et al. Diagnostic value of cardiopulmonary ultrasound in elderly patients with acute respiratory distress syndrome. BMC Pulm Med. 2018;18(1):1–9.
26. Rimoldi SF, Yuzefpolskaya M, Allemann Y, Messerli F. Flash pulmonary edema. Prog Cardiovasc Dis. 2009 Nov-Dec;52(3):249-59.

27. Hasan AA, Makhlouf HA. B-lines: Transthoracic chest ultrasound signs useful in assessment of interstitial lung diseases. Ann Thorac Med. 2014;9(2):99–103.
28. Haaksma ME, Smit JM, Heldeweg MLA, Pisani L, Elbers P, Tuinman PR. Lung ultrasound and B-lines: B careful! Intensive Care Med. 2020 Mar;46(3):544-545
29. de Lange DW, Dusseljee J, Brinkman S, et al. Severity of illness and outcome in ICU patients in the Netherlands: Results from the NICE registry 2006-2007. Neth J Crit Care. 2009;13(1):16–22.

Figures

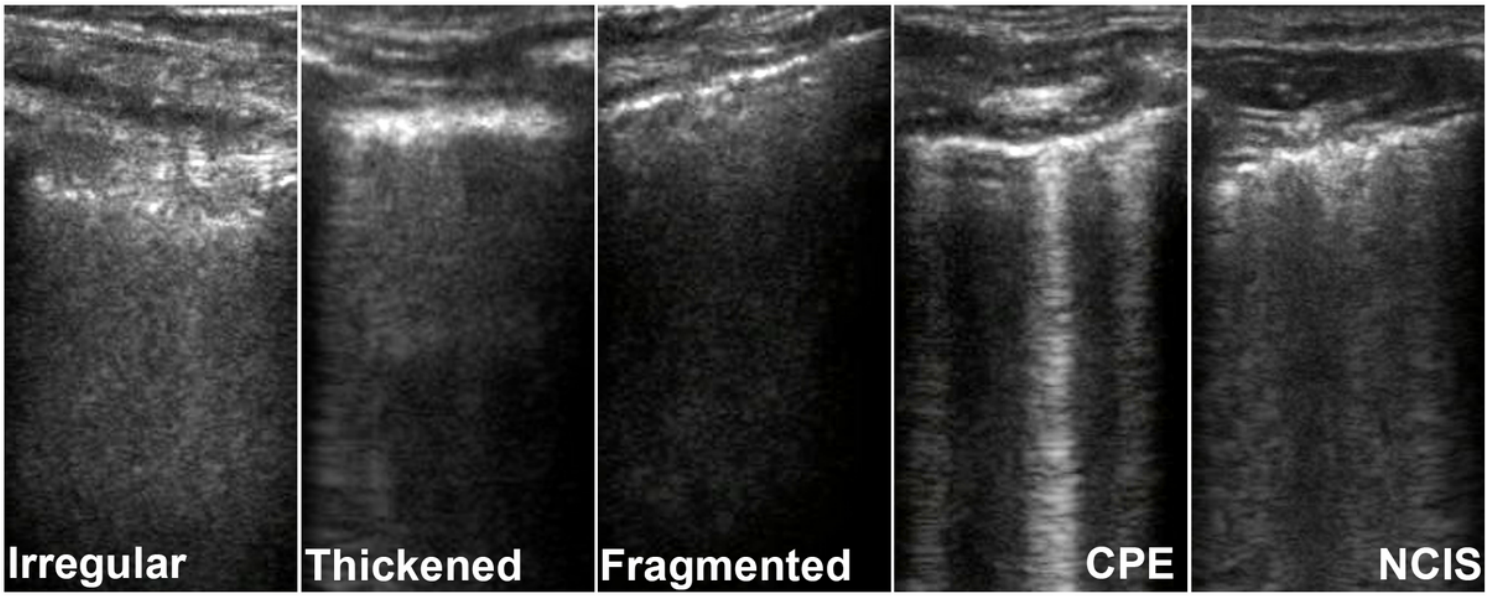

Figure 1

From left to right: examples of irregular, thickened, and fragmented pleural line. On the right, examples of patients with cardiogenic pulmonary edema and noncardiogenic interstitial syndrome, respectively.

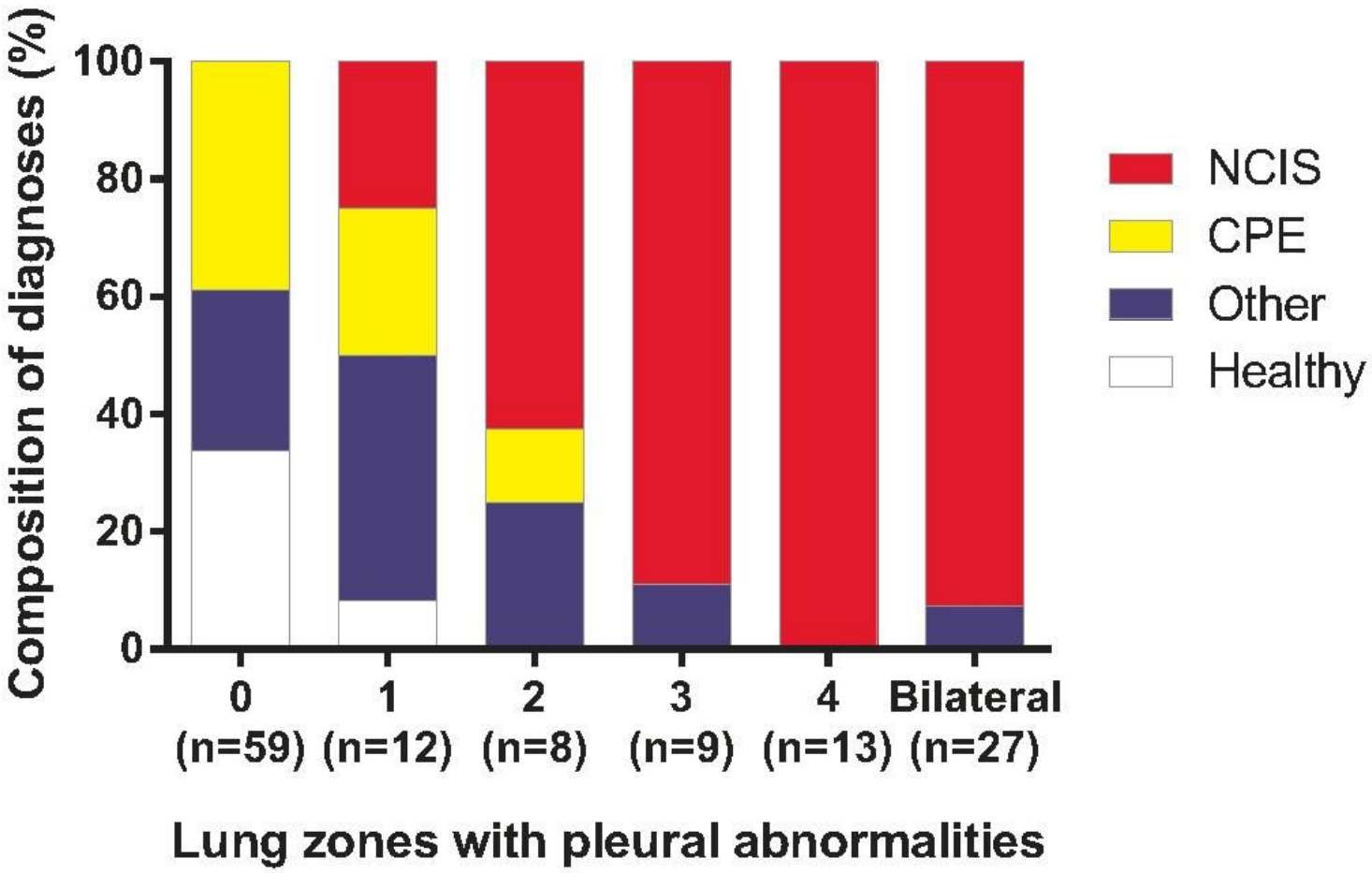

Figure 2

Composition of diagnoses for increasing lung zones with a pleural line abnormality

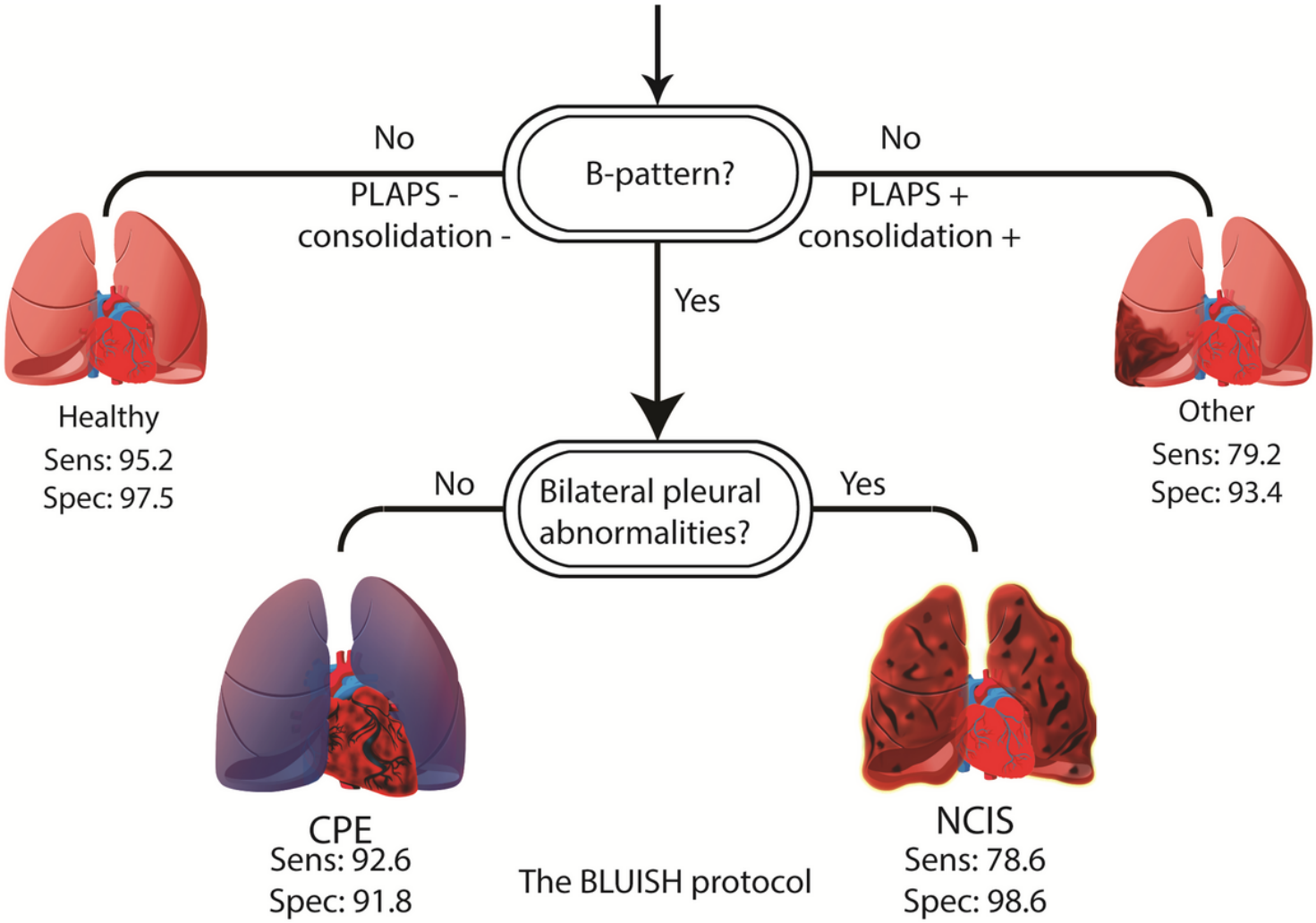

Figure 3

Diagnostic algorithm for the diagnosis and discrimination of interstitial syndrome, the BLUISH protocol, with a total accuracy of 0.86. Bilateral pleural abnormalities signifies the presence of at least two pleural abnormalities (fragmented, thickened, or irregular) per hemithorax.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.docx
- Additionalfile3.docx
- Additionalfile4.docx
- Additionalfile5.docx
- Additionalfile6.docx