Interobserver Agreement of Lung Ultrasound Findings of COVID-19

Andre Kumar, MD, MEd, Yingjie Weng, MS, Sally Graglia, MD, MPH, Sukyung Chung, PhD, Youyou Duanmu, MD, MPH, Farhan Lalani, MD, Kavita Gandhi, MD, PhD, Viveta Lobo, MD, Trevor Jensen, MD, Jeffrey Nahn, MD, John Kugler, MD

Received August 31, 2020, from the Department of Medicine, Stanford University School of Medicine, Stanford, California, USA (A.K., J.K.); Quantitative Sciences Unit, Stanford University, Stanford, California, USA (Y.W., S.C.); Department of Emergency Medicine, University of California San Francisco, San Francisco, California, USA (S.G., K.G., J.N.); Department of Emergency Medicine, Stanford University School of Medicine, Stanford, California, USA (Y.D., V.L.); and Department of Medicine, University of California San Francisco, San Francisco, California, USA (F.L., T.J.). Manuscript accepted for publication December 21, 2020.

Andre Kumar, MD, MEd is a paid consultant for Vave Health, which manufactures one of the ultrasound devices used in this study. His consultant duties include providing feedback on product development. The other authors do not have any items to disclose.

Address correspondence to Andre Kumar, MD, MEd, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 95401.

Email: akumar3@stanford.edu

Abbreviations

AAL, anterior axillary line; BMI, body mass index; CT, computerized tomography; ICU, intensive care unit; IRB, institutional review board; IRR, interrater reliability; ISM, inferior scapular margin; LUS, Lung ultrasound; PAL, posterior axillary line; POCUS, Point-of-care ultrasound; PSM, parasternal margin; SL, scapular line

doi:10.1002/jum.15620

Background—Lung ultrasound (LUS) has received considerable interest in the clinical evaluation of patients with COVID-19. Previously described LUS manifestations for COVID-19 include B-lines, consolidations, and pleural thickening. The interrater reliability (IRR) of these findings for COVID-19 is unknown.

Methods—This study was conducted between March and June 2020. Nine physicians (hospitalists: $n = 4$; emergency medicine: $n = 5$) from 3 medical centers independently evaluated $n = 20$ LUS scans ($n = 180$ independent observations) collected from patients with COVID-19, diagnosed via RT-PCR. These studies were randomly selected from an image database consisting of COVID-19 patients evaluated in the emergency department with portable ultrasound devices. Physicians were blinded to any patient information or previous LUS interpretation. Kappa values (κ) were used to calculate IRR.

Results—There was substantial IRR on the following items: normal LUS scan ($\kappa = 0.79$ [95% CI: 0.72–0.87]), presence of B-lines ($\kappa = 0.79$ [95% CI: 0.72–0.87]), ≥3 B-lines observed ($\kappa = 0.72$ [95% CI: 0.64–0.79]). Moderate IRR was observed for the presence of any consolidation ($\kappa = 0.57$ [95% CI: 0.50–0.64]), subpleural consolidation ($\kappa = 0.49$ [95% CI: 0.42–0.56]), and presence of effusion ($\kappa = 0.49$ [95% CI: 0.41–0.56]). Fair IRR was observed for pleural thickening ($\kappa = 0.23$ [95% CI: 0.15–0.30]).

Discussion—Many LUS manifestations for COVID-19 appear to have moderate to substantial IRR across providers from multiple specialties utilizing differing portable devices.

The most reliable LUS findings with COVID-19 may include the presence/count of B-lines or determining if a scan is normal. Clinical protocols for LUS with COVID-19 may require additional observers for the confirmation of less reliable findings such as consolidations.

Key Words—COVID-19; interobserver agreement; interrater; lung; POCUS; reliability; SARS-CoV-2; ultrasound

Introduction

Point-of-care ultrasound (POCUS) has the potential to transform healthcare delivery in the era of COVID-19 with its diagnostic expediency.[1] POCUS provides real-time interpretation of a patient's condition in an augmented clinical manner, which can immediately impact management decisions.[2,3] POCUS devices, particularly handheld devices, are often cheaper than traditional radiological equipment such as X-ray or

computerized tomography (CT) machines, which make POCUS ideal for COVID-19 surge scenarios where these resources may be limited.[4,5] Furthermore, POCUS may reduce the number of providers exposed to patients with COVID-19 by decreasing the need for radiological studies, which could result in decreases of personal protective equipment usage by radiological technicians and the resources needed to decontaminate larger radiological equipment.[4]

Previously reported pulmonary manifestations of COVID-19 with POCUS include B-lines, subpleural consolidations, pleural thickening, and absence of pleural effusions (Figure 1).[4–8] POCUS has been proposed to aid in the diagnosis of COVID-19,[9] as well as to predict intensive care unit (ICU) admission or death.[10] Given the potential for POCUS to predict outcomes among COVID-19 patients, there is a significant need to determine if these findings can be reliably interpreted among providers. Outside of COVID-19, previous investigations have found moderate to excellent interobserver agreement for many lung ultrasound (LUS) findings, including B-lines, consolidations, and effusions.[2,11,12]

In this study, we characterize the interobserver agreement of LUS findings that have been described for COVID-19. These images were collected using portable ultrasound devices, which may be more commonly utilized as a result of the COVID-19 pandemic.

Methods

Participants and Setting

This analysis was conducted as part of an ongoing study investigating the role of POCUS for COVID-19 that is being conducted at 4 medical centers in the United States (Clinicaltrials.gov Registration: NCT04384055). This investigation utilizes a prospectively collected database and includes adult patients who meet the following criteria: (1) presentation to the emergency department with symptoms[13] suspicious for COVID-19, (2) a positive nasopharyngeal RT-PCR for SARS-CoV-2, and (3) received an LUS during their emergency department course or subsequent hospitalization (up to 28 days from admission). This study has received institutional review board (IRB) approval at all participating sites. Waiver of consent was approved by our IRBs.

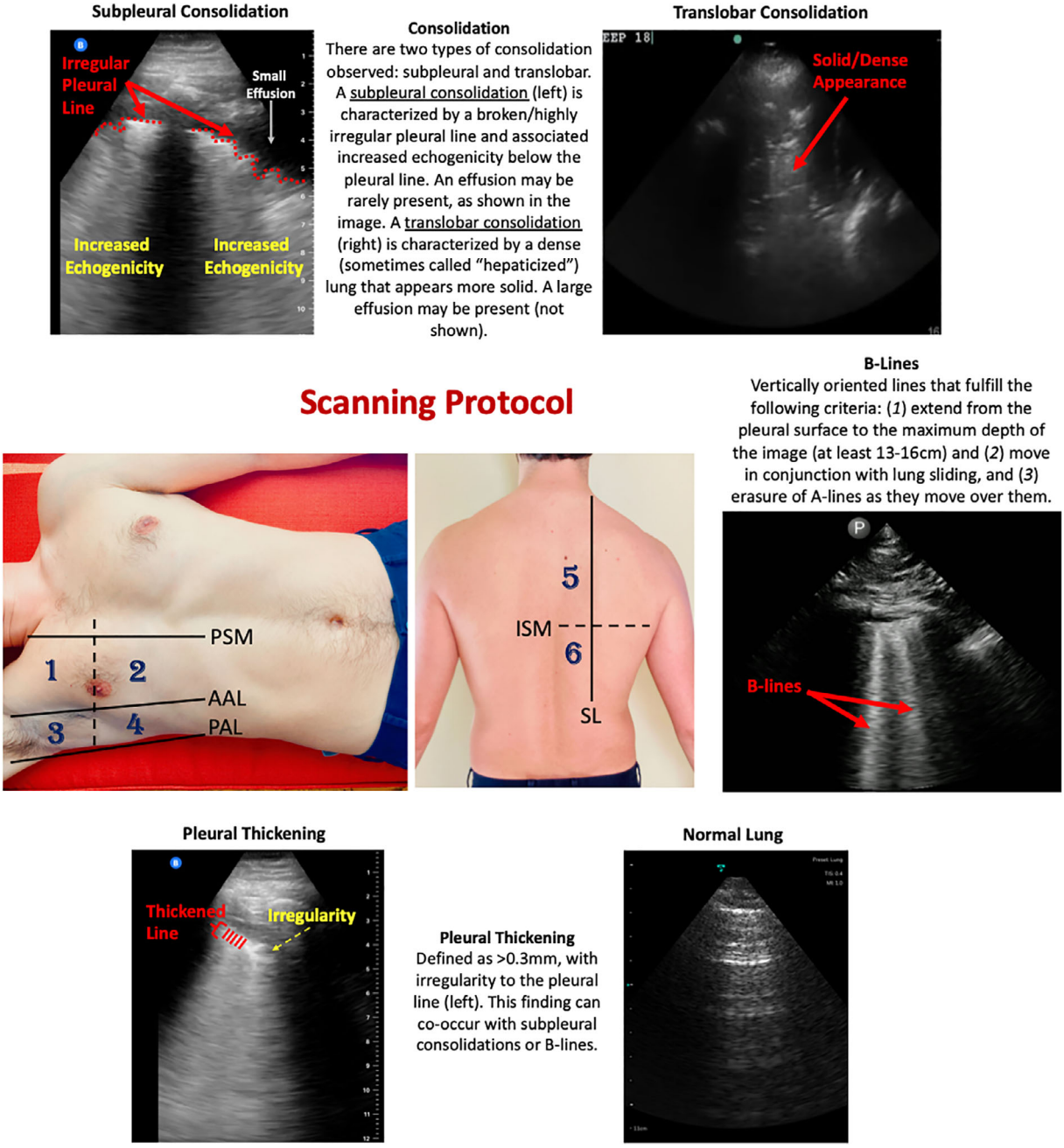
Study Procedure

In this study, a total of $n = 20$ LUS scans collected between March and June 2020 were randomly selected from our image database consisting of COVID-19 POCUS images. These scans originated from $n = 13$ patients. The difference between the number of scans and patients occurred because approximately half of the patients received multiple scans during their hospitalization (range 1–5 scans per patient). Whether a patient received multiple scans was based on patient factors (eg, duration of hospitalization) as well as provider discretion (please see Clinicaltrials.gov Registration: NCT04384055 for full protocol details). Any scans analyzed from the same patient were acquired on different days. All scans were collected using a 12-zone lung protocol, which we have previously described and is also demonstrated in Figure 1.[6] Each zone clip was 6 seconds in length.[6] The LUS studies were collected on the following devices: Butterfly iQ ($n = 9$), Vave Personal Ultrasound ($n = 6$), Philips Lumify ($n = 4$), and Sonosite M-turbo ($n = 1$), which represent the commercially available portable devices at our institutions. In our study protocol, providers can acquire these scans using any portable device available to them, and the types of devices used in this analysis were the result of random selection from our database.

This study included 9 physicians from 2 specialties (hospitalists $n = 4$; emergency medicine, $n = 5$). These 9 physicians independently evaluated $n = 20$ LUS scans ($n = 180$ independent observations). All POCUS scans were obtained by 3 of the study authors (YD, AK, and JK). These authors are credentialed in POCUS for patient care at their respective institutions. YD previously completed an Emergency Ultrasound Fellowship. JK is the director of POCUS for the Stanford Department of Medicine. AK is an instructor for the Society of Hospital Medicine POCUS Certification Program. Additional demographic details of the researchers involved in image interpretation can be found in the Appendix.

The physicians met for a 1-hour calibration session at the beginning of the study to review sample videos and discuss their real-time interpretations. These samples were not included in the interrater

Figure 1. Scanning protocol and lung ultrasound findings in COVID-19 patients. This study utilized a 12-zone protocol.[27] On each hemithorax, there are 6 zones. The exam begins on the patient's right side. Zones 1–2 (anterior zones) are between the parasternal margin (PSM) and the anterior axillary line (AAL) and are best obtained in the mid-clavicular line. Zones 3–4 (lateral zones) are between the AAL and posterior axillary line (PAL) and are best obtained in the mid-axillary line. The nipple line serves as a bisecting area between these zones. Zones 5–6 (posterior zones) are medial to the scapular line (SL) and are bisected by the inferior scapular margin (ISM). The zone areas are repeated on the contralateral hemithorax (starting with zone 7). This figure contains an overview of the observed ultrasound findings based on previously described terminology.[3,27,28]



analysis. The physicians were then instructed to independently assess the 20 de-identified studies and input their interpretation using a central electronic database (REDCap).[14] All physicians received a document that contained the scanning protocol and previously described definitions for each of the pathological findings (Figure 1).[3,12,15,16] No other instruction on image interpretation was provided to the group during the independent assessment period. The physicians who interpreted the images were blinded to any patient information or previous interpretation.

Statistical Analysis

The degree of agreement for kappa values was based on the research originally described by Cohen and later Landis & Koch.[17,18] In our study, Fleiss' kappa statistics and corresponding 95% confidence intervals were reported. In this scheme, we interpreted the Fleiss' kappa statistics[17] with the following criteria: κ <0 (no agreement), κ = 0–0.20 (none to slight agreement), κ = 0.21–0.40 (fair agreement), κ = 0.41–0.60 (moderate agreement), κ = 0.61–0.80 (substantial agreement), and κ = 0.81–1.0 (near perfect agreement). All analyses were performed using R statistical programming language, version 3.6.1 (Vienna, Austria).

Results

Study Population

Patient demographics and clinical characteristics are described in Table 1. The mean body mass index (BMI) for the patient population was 30.5 kg/m^2 (range 24.6–37.5). The mean age for the study cohort was 49.2 (SD 19.2). Approximately 57% of the patients were female (Table 1).

Normal Versus Abnormal Scan

Overall, there was substantial agreement on determining whether a scan contained no abnormalities, including absence of B-lines, consolidations, pleural thickening, or effusions (κ = 0.79 [95% CI: 0.72–0.87]; Table 2). There were 2 scans the majority of researchers ($n \geq 5$) identified as normal (Appendix).

Table 1. Patient Characteristics

Mean age (SD)	49.2 (19.2)
Mean BMI (kg/m^2; SD)	30.5 (3.9)
Female (%)	7 (54%)
Admitted to ICU (%)	6 (46%)
Medical history	
Hypertension	3 (23%)
Hyperlipidemia	3 (23%)
Diabetes	5 (38%)
Coronary artery disease	2 (15%)
Heart failure	0 (0%)
COPD	0 (0%)
Asthma	3 (23%)
Malignancy	0 (0%)

Note: $N = 20$ scans were randomly selected from our prospectively acquired database, which originated from $N = 13$ patients. The patient characteristics are displayed.
Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit.

B-lines

There was substantial interrater agreement on the presence of B-lines (κ = 0.79 [95% CI: 0.72–0.87]) or whether the scan contained 3 or more B-lines per lung field (κ = 0.72 [95% CI: 0.64–0.79]; Table 2). The presence of bilateral B-lines also demonstrated substantial agreement (κ = 0.70 [95% CI: 0.63–0.78]; Table 3). Similarly, the presence of B-lines had substantial agreement across the anterior (κ = 0.79 [95% CI: 0.72–0.87]), lateral (κ = 0.76 [95% CI: 0.69–0.83]), and posterior lung zones (κ = 0.70 [95% CI: 0.63–0.77]; Table 3). Notably, the total count of B-lines per scan had slight agreement (κ = 0.16 [95% CI: 0.14–0.17]). There were 18 scans the majority of researchers ($n \geq 5$) identified as containing B-lines (Appendix).

Consolidation

There was moderate interrater agreement on the presence of any consolidation (κ = 0.57 [95% CI: 0.50–0.64]; Table 2). When analyzed by consolidation subtype, the presence of subpleural consolidations had moderate agreement (κ = 0.49 [95% CI: 0.42–0.56]) and translobar consolidations had slight agreement (κ = 0.15 [95% CI: 0.08–0.23]). The presence of bilateral consolidation had fair agreement (κ = 0.28 [95% CI: 0.20–0.35]; Table 3). There was variable agreement of the presence of any type of consolidation by location: anterior (κ = 0.71 [95% CI: 0.63–0.78]), lateral (κ = 0.56 [95% CI: 0.48–

Table 2. Interobserver Agreement of Lung Ultrasound Findings Among COVID-19 Patients

Finding	Kappa Value	95% CI	Degree of Agreement
Normal examination	0.79	[0.72–0.87]	Substantial agreement
B-lines present	0.79	[0.72–0.87]	Substantial agreement
Three or more B-lines per lung field	0.72	[0.64–0.79]	Substantial agreement
Consolidation (translobar or subpleural)	0.57	[0.50–0.64]	Moderate agreement
Subpleural consolidation	0.49	[0.42–0.56]	Moderate agreement
Pleural effusion	0.49	[0.41–0.56]	Moderate agreement
Effusion size (small vs moderate)	0.47	[0.40–0.55]	Moderate agreement
Pleural thickening	0.23	[0.15–0.30]	Fair agreement
Total count of B-lines per scan	0.16	[0.14–0.17]	None to slight Agreement
Translobar consolidation	0.15	[0.08–0.23]	None to slight Agreement

Note: The degree of agreement was based on previously reported methodology. Seventeen effusion size was defined as <1 cm from the visceral and parietal pleura versus ≥1 cm.

Table 3. Interobserver Agreement of Lung Ultrasound by Lung Field Location

Finding	Kappa Value	95% CI	Degree of Agreement
B-lines			
Anterior B-lines	0.79	[0.72–0.87]	Substantial agreement
Lateral B-lines	0.76	[0.69–0.83]	Substantial agreement
Posterior B-lines	0.70	[0.63–0.77]	Substantial agreement
Bilateral B-lines	0.70	[0.63–0.78]	Substantial agreement
Consolidation			
Anterior consolidation	0.71	[0.63–0.78]	Substantial agreement
Lateral consolidation	0.56	[0.48–0.63]	Moderate agreement
Posterior zone consolidation	0.86	[0.78–0.93]	Near perfect Agreement
Presence of bilateral consolidation	0.28	[0.20–0.35]	Fair agreement
Pleural thickening			
Anterior pleural thickening	0.33	[0.26–0.40]	Fair agreement
Lateral pleural thickening	0.34	[0.26–0.41]	Fair agreement
Posterior pleural thickening	0.48	[0.41–0.55]	Moderate agreement
Presence of bilateral pleural thickening	0.33	[0.26–0.41]	Fair agreement
Effusions			
Presence of bilateral effusion	0.40	[0.32–0.47]	Fair agreement

Note: Lung zones (anterior, lateral, and posterior) are described in Figure 1.

0.63]), and posterior (κ = 0.86 [95% CI: 0.78–0.93]; Table 3). There were 14 scans the majority of researchers ($n \geq 5$) identified as containing consolidations (Appendix).

Pleural Thickening

There was fair agreement on the presence of pleural thickening (κ = 0.23 [95% CI: 0.15–0.30]; Table 2). Similarly, there was fair agreement on the presence of bilateral pleural thickening (κ = 0.33 [95% CI: 0.26–0.41]). When analyzed by location, pleural thickening demonstrated fair to moderate agreement across the anterior, lateral, and posterior zones (Table 3). There were 14 scans the majority of researchers ($n \geq 5$) identified as containing pleural thickening (Appendix).

Pleural Effusion

There was moderate agreement on the presence of pleural effusions (κ = 0.49 [95% CI: 0.41–0.56]) and size of effusion (κ = 0.47 [95% CI: 0.40–0.55]; Table 2). The presence of bilateral pleural effusions had fair agreement (κ = 0.40 [95% CI: 0.32–0.47]; Table 3). There were 3 scans the majority of researchers ($n \geq 5$) identified as containing effusions (Appendix).

Discussion

Previous investigations have demonstrated that COVID-19 has notable sonographic features, including B-lines, subpleural consolidations, pleural thickening, and minimal pleural effusions. These findings correlate well with findings observed with COVID-19 with computed tomography, including peripheral nodularity and ground-glass consolidations (Appendix).[8,19] Despite growing usage of LUS for respiratory disorders such as COVID-19, the interrater reliability of LUS for COVID-19 remains uncertain. In this study, several LUS findings demonstrated substantial agreement (eg, B-lines), while others demonstrated moderate to fair agreement (eg, consolidations, pleural thickening, or effusions). This study represents the first investigation of the interobserver agreement of LUS findings in COVID-19 and includes practitioners from multiple specialties who utilized several portable devices.

There are several implications of these findings. Previous authors have demonstrated the presence of B-lines can be used in the diagnostic evaluation for COVID-19.[6,9] Others have proposed that bilateral B-lines on LUS increase the risk of ICU admission or death with COVID-19.[10] Our results suggest that there is substantial interobserver agreement for the presence of B-lines across multiple provider specialties utilizing different handheld devices. Therefore, the presence of B-lines may represent a reliable diagnostic and prognostic clinical entity for COVID-19. Similarly, there was substantial agreement on whether an LUS scan was interpreted as abnormal versus normal. While the prognostic value of a normal LUS for COVID-19 remains uncertain, others have shown that chest radiograph abnormalities with COVID-19 are associated with ICU admission.[20] Furthermore, an abnormal LUS scan (outside of COVID-19) has prognostic implications across multiple diseases.[21–23] Future studies are needed to determine if LUS can reliably predict clinical outcomes with COVID-19.

How do these findings compare to the interrater reliability literature for LUS outside of COVID-19? Previous investigations have demonstrated moderate to substantial agreement for B-lines.[24–27] In contrast, there is only moderate to fair agreement for consolidation,[11,12,24] pleural irregularity,[24] and effusions.[11,24] Others have shown that there is substantial interrater reliability for LUS across differing probes (eg, linear versus phased array),[27,28] which is important given different portable devices were used in this investigation. Future investigations of LUS for COVID-19 should consider multiple observers to confirm less reliable findings and utilize a standardized interpretation protocol. This latter point may be important because there are various definitions of LUS findings in the literature, especially for consolidations or pleural thickening.[5,11,12,24] Although consolidations had moderate agreement in this study, the reliability of this finding might improve with more specific definitions and consensus-based guidelines. Similarly, this study defined pleural thickening as >0.3 mm increase with pleural irregularity based on previously described work (Figure 1),[29] but this definition may not be uniform.[3,15,16]

Even with an agreed upon definition for this study, there was only fair agreement for pleural thickening. Therefore, this finding might be best reserved for CT, where pleura thickness may be more easily ascertained.

There are several limitations to this study. Our study population was confined to patients who presented to the emergency department or were hospitalized, which limits the generalizability of these findings for providers in the outpatient setting. Although we randomly selected LUS studies from our database, we sampled fewer patients (n = 13) than the total number of scans analyzed (n = 20) because our database contained patients who had received multiple scans. The researchers in this study completed a 1-hour calibration session and had a definition sheet when interpreting images. Therefore, the IRR may be lower for certain findings among less-trained practitioners. Although this study utilized several portable devices, there was variability in the image quality of these devices (particularly when visualizing the pleural line), which may have affected the results regarding pleural thickening or subpleural consolidation. Certain pathological findings (B-lines, consolidations) may have been more represented in our database than others (pleural effusions). Despite these limitations, this study represents one of the first dedicated investigations into the interobserver agreement of LUS findings for COVID-19.

In conclusion, many LUS manifestations for COVID-19 appear to have substantial to moderate

IRR across providers from multiple specialties utilizing differing portable devices. More reliable findings included the determination of a normal scan, the presence and location of B-lines, and determining if ≥3 B-lines were present. Less reliable findings related to the presence or locations of consolidations, pleural thickening, or effusions. Since the presence of B-lines may have diagnostic and prognostic utility for COVID-19, this finding can likely be interpreted without additional oversight and can be incorporated into future clinical protocols. In contrast, other findings such as pleural thickening may be less reliable, and clinical protocols incorporating these findings may require quality assurance and precise definitions for accurate interpretation.

References

1. Moore CL, Copel JA. Point-of-care ultrasonography. *N Engl J Med* 2011; 364:749–757.
2. Kumar A, Liu G, Chi J, Kugler J. The role of technology in the bedside encounter. *Med Clin North Am* 2018; 102:443–451.
3. Lobo V, Weingrow D, Perera P, Williams SR, Gharahbaghian L. Thoracic ultrasonography. *Crit Care Clin* 2014; 30:93–117.v–vi.
4. Fox S, Dugar S. Point-of-care ultrasound and COVID-19. *Cleve Clin J Med* 2020. https://doi.org/10.3949/ccjm.87a.ccc019. [Epub ahead of print].
5. Mathews BK, Koenig S, Kurian L, et al. Clinical progress note: point-of-care ultrasound applications in COVID-19. *J Hosp Med* 2020; 15:353–355.
6. Kumar AD, Chung S, Duanmu Y, Graglia S, Lalani F, Gandhi K, et al. Lung ultrasound findings in patients hospitalized with Covid-19 [Internet]. Infectious Diseases (except HIV/AIDS). *medRxiv*; 2020. https://www.medrxiv.org/content/10.1101/2020.06.25.20140392v1.abstract
7. Volpicelli G, Gargani L. Sonographic signs and patterns of COVID-19 pneumonia. *Ultrasound J* 2020; 12:22.
8. Fiala MJ. Ultrasound in COVID-19: a timeline of ultrasound findings in relation to CT. *Clin Radiol* 2020; 75:553–554.
9. Peyrony O, Marbeuf-Gueye C, Truong V, et al. Accuracy of emergency department clinical findings for diagnosis of coronavirus disease 2019. *Ann Emerg Med* 2020; 76:405–412. https://doi.org/10.1016/j.annemergmed.2020.05.022.
10. Bonadia N, Carnicelli A, Piano A, et al. Lung ultrasound findings are associated with mortality and need of intensive care admission in COVID-19 patients evaluated in the Emergency Department. *Ultrasound Med Biol* 2020; 46:2927–2937. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7362856/.
11. Millington SJ, Arntfield RT, Guo RJ, et al. Expert agreement in the interpretation of lung ultrasound studies performed on mechanically ventilated patients. *J Ultrasound Med* 2018; 37:2659–2665.
12. Lichtenstein DA. Lung ultrasound in the critically ill. *Ann Intensive Care* 2014; 4:1.
13. CDC. *Coronavirus Disease 2019 (COVID-19)*. Atlanta, GA: Centers for Disease Control and Prevention; 2020 https://www.cdc.gov/coronavirus/2019-nCoV/hcp/clinical-criteria.html.
14. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; 42:377–381.
15. Reuss J. Sonography of the pleura. *Ultraschall Med* 2010; 31:8–22. quiz 23–25.
16. Doerschug KC, Schmidt GA. Intensive care ultrasound: III. Lung and pleural ultrasound for the intensivist. *Ann Am Thorac Soc* 2013; 10:708–712.
17. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971; 76:378–382.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174.
19. Ottaviani S, Franc M, Ebstein E, et al. Lung ultrasonography in patients with COVID-19: comparison with CT. *Clin Radiol* 2020; 75:877.e1–877.e6. https://doi.org/10.1016/j.crad.2020.07.024.
20. Liang W, Liang H, Ou L, et al. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Internal Medicine* 2020; 180:1081. http://dx.doi.org/10.1001/jamainternmed.2020.2033.
21. Gargani L, Bruni C, Romei C, et al. Prognostic value of lung ultrasound B-lines in systemic sclerosis. *Chest* 2020; 158:1515–1525.
22. Brusasco C, Santori G, Bruzzo E, et al. Quantitative lung ultrasonography: a putative new algorithm for automatic detection and quantification of B-lines. *Crit Care* 2019 Aug 28;23:288.
23. Saad MM, Kamal J, Moussaly E, et al. Relevance of B-lines on lung ultrasound in volume overload and pulmonary congestion: clinical correlations and outcomes in patients on hemodialysis. *Cardiorenal Med* 2018; 8:83–91.
24. Gravel CA, Monuteaux MC, Levy JA, Miller AF, Vieira RL, Bachur RG. Interrater reliability of pediatric point-of-care lung ultrasound findings. *Am J Emerg Med* 2020; 38:1–6.
25. Gullett J, Donnelly JP, Sinert R, et al. Interobserver agreement in the evaluation of B-lines using bedside ultrasound. *J Crit Care* 2015; 30:1395–1399.
26. Vieira JR, Castro MR d, Guimarães T de P, et al. Evaluation of pulmonary B lines by different intensive care physicians using

bedside ultrasonography: a reliability study. *Rev Bras Ter Intensiva* 2019; 31:354–360.

27. Chiem AT, Chan CH, Ander DS, Kobylivker AN, Manson WC. Comparison of expert and novice sonographers' performance in focused lung ultrasonography in dyspnea (FLUID) to diagnose patients with acute heart failure syndrome. *Acad Emerg Med* 2015; 22:564–573.
28. Gomond-Le Goff C, Vivalda L, Foligno S, Loi B, Yousef N, De Luca D. Effect of different probes and expertise on the interpretation reliability of point-of-care lung ultrasound. *Chest* 2020; 157: 924–931.
29. Soni NJ, Franco R, Velez MI, et al. Ultrasound in the diagnosis and management of pleural effusions. *J Hosp Med* 2015; 10: 811–816.