ORIGINAL RESEARCH CONTRIBUTION

Comparison of Expert and Novice Sonographers' Performance in Focused Lung Ultrasonography in Dyspnea (FLUID) to Diagnose Patients With Acute Heart Failure Syndrome

Alan T. Chiem, MD, MPH, Connie H. Chan, MD, Douglas S. Ander, MD, Andrew N. Kobylivker, MD, and William C. Manson, MD

Abstract

Objectives: The goal of this study was to examine the ability of emergency physicians who are not experts in emergency ultrasound (US) to perform lung ultrasonography and to identify B-lines. The hypothesis was that novice sonographers are able to perform lung US and identify B-lines after a brief intervention. In addition, the authors examined the diagnostic accuracy of B-lines in undifferentiated dyspneic patients for the diagnosis of acute heart failure syndrome (AHFS), using an eight-lung-zone technique as well as an abbreviated two-lung-zone technique.

Methods: This was a prospective, cross-sectional study of patients who presented to the emergency department (ED) with acute dyspnea from May 2009 to June 2010. Emergency medicine (EM) resident physicians, who received a 30-minute training course in thoracic US examinations, performed lung ultrasonography on patients presenting to the ED with undifferentiated dyspnea. They attempted to identify the presence or absence of sonographic B-lines in eight lung fields based on their bedside US examinations. An emergency US expert blinded to the diagnosis and patient presentation, as well as to the residents' interpretations of presence of B-lines, served as the criterion standard. A secondary outcome determined the accuracy of B-lines, using both an eight-lung-zone and a two-lung-zone technique, for predicting pulmonary edema from AHFS in patients presenting with undifferentiated dyspnea. Two expert reviewers who were blinded to the US results determined the clinical diagnosis of AHFS.

Results: A cohort of 66 EM resident physicians performed lung US on 380 patients with a range of 1 to 28 examinations, a mean of 5.8 examinations, and a median of three examinations performed per resident. Compared to expert interpretation, lung US to detect B-lines by inexperienced sonographers achieved the following test characteristics: sensitivity 85%, specificity 84%, positive likelihood ratio (+LR) 5.2, negative likelihood ratio (−LR) 0.2, positive predictive value (PPV) 64%, and negative predictive value (NPV) 94%. Regarding the secondary outcome, the final diagnosis was AHFS in 35% of patients (134 of 380). For novice sonographers, one positive lung zone (i.e., anything positive) had a sensitivity of 87%, a specificity of 49%, a +LR of 1.7, a −LR of 0.3, a PPV of 50%, and an NPV of 88% for predicting AHFS. When all eight lung zones were determined positive (i.e., totally positive) by novice sonographers, the sensitivity was 19%, specificity was 97%, +LR was 5.7, −LR was 0.8, PPV was 76%, and NPV was 68%

for predicting AHFS. The areas under the curve for novice and expert sonographers were 0.77 (95% CI = 0.72 to 0.82) and 0.76 (95% CI = 0.71 to 0.82), respectively.

Conclusions: Novice sonographers can identify sonographic B-lines with similar accuracy compared to an expert sonographer. Lung US has fair predictive value for pulmonary edema from acute heart failure in the hands of both novice and expert sonographers.

Despite advances in care over several decades, acute heart failure syndrome (AHFS) continues to be the most common discharge diagnosis of patients over age 65 years, with a 1-month rehospitalization rate of 20% to 30%.[1,2] To address this problem, the National Institutes of Health has highlighted some major research priorities in AHFS, including enrollment of patients at presentation to the emergency department (ED), and the need to develop adjuncts to physical examination such as point-of-care (POC) ultrasound (US) to improve accuracy in diagnosis of AHFS phenotype and severity, as well as to monitor treatment.[3–5] This is because the physical examination findings of acute heart failure are insensitive and/or nonspecific. For example, the findings of rales, jugular venous distension, and lower-extremity edema have been shown to be unreliable in identifying patients with AHFS. An S3 gallop is specific, but presents in less than 25% of AHFS patients.[6,7] Treatment may be delayed or initiated inappropriately while waiting for laboratory and chest radiograph results. Clinicians would be able to direct treatment sooner if there were a quick bedside method of determining the underlying cause of acute dyspnea.

Point-of-care US is standard curriculum in emergency medicine (EM) residency programs.[8] In a few small studies, POC cardiac and pulmonary US have been demonstrated to reliably identify diastolic dysfunction, systolic dysfunction, and the presence of systemic and pulmonary congestion.[9–13] ED-based enrollment in conjunction with POC US allows for monitoring of response to treatment, along with the use of traditional criteria for improvement, such as symptomatic improvement, amount of diuresis, and B-type natriuretic peptide (BNP) decrease.

Sonographic B-lines can be used to identify interstitial syndrome, of which pulmonary edema from AHFS is the most common etiology. B-lines occur as a result of fluid accumulation in the pulmonary interstitium allowing sound waves to conduct and resonate through the normally aerated lung parenchyma, creating a vertical, linear artifact (Figure 1). With greater amounts of interstitial fluid, more B-lines become visible. B-lines are diagnostic of alveolar-interstitial syndrome, which can be due to acute respiratory distress syndrome, cardiogenic and noncardiogenic pulmonary edema, and pulmonary fibrosis. While there are key differences in the sonographic findings, that discussion is beyond the scope of this article. However, prior studies have shown that bilateral lung zones that each contain three or more evenly spaced B-lines are highly predictive of pulmonary edema.[14–17] More recently, there has been evidence to show that presence of thoracic B-lines in dyspneic patients correlates with an elevated BNP and worsened prognosis in AHFS.[18–22] As a result of these observational studies, lung US in the evaluation of pulmonary edema received strong Level B recommendations in an international consensus panel.[23]

However, most of the studies use a small number of sonographers who are typically fellowship-trained in emergency ultrasonography or with similar backgrounds, limiting generalizability to the general emergency physician (EP).[11–13] Other studies include inferior vena cava (IVC) assessment and focused cardiac US. These theoretically increase the specificity for interstitial syndrome due to acute heart failure, but require more

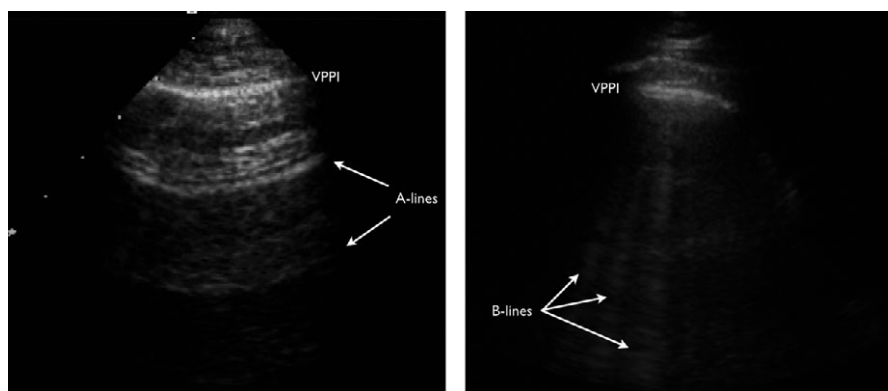**Figure 1.** Thoracic ultrasonography for the evaluation of sonographic B-lines. The image on the left shows "A-lines" which are reverberation artifact from the visceral-parietal pleural interface. This is a normal finding. The image on the right shows multiple vertically oriented B-lines, which are typically due to interstitial fluid that allows for conduction of sound waves deep into the lung parenchyma.

training and procedure time to perform IVC and focused cardiac US. This may deter some physicians from adopting US in assessing patients with dyspnea. Using these additional measures may also lead to a higher false-negative rate, given that half of AHFS presentations have preserved or normal ejection fractions.[1–5] For example, Anderson et al.[12] showed that a combination of B-line ultrasonography with IVC or ejection fraction assessment decreased sensitivity from 70% to under 40%.

Accurate and rapid diagnosis of AHFS may lead to earlier treatment and improved patient outcomes.[5] Previous research indicates a moderately high sensitivity and specificity when experienced investigators perform lung US.[11–17] This study investigated the diagnostic accuracy of lung US performed by a large group of novice sonographers in patients presenting to the ED with a chief complaint of dyspnea. If the test characteristics are similar between expert and novice sonographers, then the identification of sonographic B-lines would be more applicable to the general EP and serve as a useful POC test. This would allow for earlier targeted treatment initiation and perhaps serve as a useful reassessment tool to monitor response to treatment as it is noninvasive, potentially easy to perform, and does not expose patients to additional radiation.

The primary objective was to assess the ability of EM resident physicians with minimal training to perform lung US and identify sonographic B-lines in real time. We compared their performance with expert sonographer review and calculated sensitivity, specificity, likelihood ratios (LRs), and positive and negative predictive values (PPV, NPV). We predicted that novice interpretation would exceed 80% in sensitivity and specificity compared to expert sonographer review. A secondary outcome was to characterize the diagnostic accuracy of resident-performed US in predicting pulmonary edema due to AHFS, as determined by blinded expert reviewers. We assessed the role of increasing lung zone positivity on the diagnosis of AHFS. We calculated receiver operating characteristic (ROC) curves and the area under the curve (AUC) for novice and expert interpretation, and predicted that both would have at least fair predictive ability, with AUC > 0.7.[24]

METHODS

Study Design

This was a prospective, cross-sectional study of novice sonographers who performed US on a convenience sample of patients presenting to the ED with a chief complaint of dyspnea. The hospital's institutional review board approved the study. This clinical trial was preregistered with ClinicalTrials.gov (NCT00833144). Patients provided written informed consent prior to enrollment. Residents were not required to participate, and no consent was obtained, but the procedures were part of their usual practice.

Study Setting and Population

This study was conducted from May 2009 to June 2010 at an inner-city ED with over 100,000 patient visits per year. It is home to an EM training program featuring 56 residents and an emergency US fellowship program. The study period overlapped two academic years, so a total of 72 residents were eligible to participate. Study sonographers were EM resident physicians on clinical duty who were recruited by undergraduate research associates (RAs).

Undergraduate RAs screened patients presenting to the ED with chief complaints of dyspnea or shortness of breath 7 days a week from 8 a.m. to midnight. Exclusion criteria included patients younger than 18 years old, incarcerated patients, pregnant patients, dialysis patients, non–English-speaking patients, and patients with dyspnea clearly caused by other diagnoses (e.g., stab wound to the chest). Attending physicians on duty were asked to make the determination if potential subjects' dyspnea were from a clear cause not related to acute heart failure (e.g., pneumothorax), in which case they were excluded. Patients were not enrolled if chest radiographic findings were already available and known to the sonographer. Patients were also excluded if they were on positive pressure ventilation or receiving nebulizer treatment or too ill to provide written consent.

Study Protocol

After informed consent was obtained, each patient received an eight-zone thoracic US study performed by a resident physician, using a curvilinear transducer (5–2 MHz, HD-11XE, Philips Healthcare, Andover, MA) on thoracic examination preset. RAs recorded the physician's bedside interpretation of whether each lung field was positive for three or more B-lines. Video clips were also recorded for each thoracic US study for later review. RAs also collected baseline treatment interventions and demographic data by reviewing the triage form and order sheet during the patient screening process.

All EM residents had received brief training in lung US and were eligible for participation in the study. Sonographers were not necessarily the treating physicians for the patients, nor explicitly blinded to the clinical status of the patient. However, they typically performed the US prior to chart review or patient examination, as patients were screened within 10 minutes of ED rooming. Resident experience in US consisted of a 2-week rotation during intern year, as well as several lectures and workshops during residency conference. None of the US training prior to the study featured thoracic US for the assessment of B-lines.

Lung US training consisted of a brief 30-minute lecture on technique and recognition of sonographic B-lines by the emergency US director. Specifically, they were taught the Volpicelli method of assessing eight lung zones, with a positive zone defined as ≥3 B-lines seen (Figure 2).[17] These eight lung zones have been used in previous lung US research.[11–13] No posttraining evaluation was given to assess for competency.

Data Collection and Processing

There were four components of data collection: the POC US data, the POC clinical data, the expert sonographer interpretation of US data, and an abstracted chart review of subjects to determine their final primary diagnoses. RAs collected the POC data during the patients'
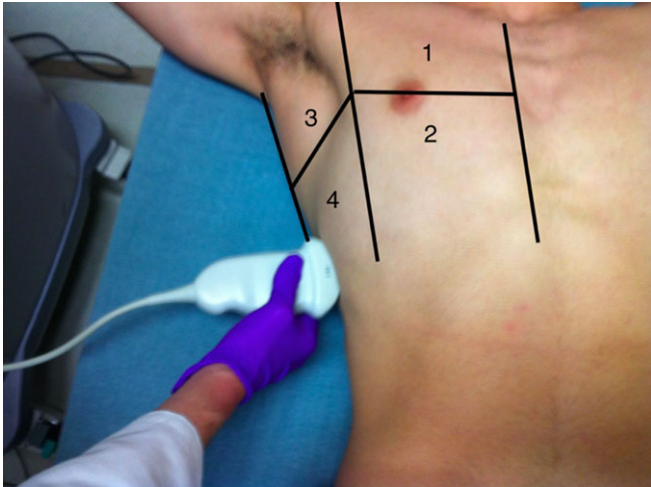
Figure 2. The right chest divided into four lung zones according to the Volpicelli method. Zones 1 and 2 correspond to the anterior superior and inferior lung zones, respectively. On the left chest, zones 5 and 6 correspond to these zones. Zones 3 and 4 correspond to the axillary superior and inferior zones, with corresponding zones 7 and 8 on the left chest. The vertical borders from lateral to medial are posterior axillary line, anterior axillary line, and sternal border. The horizontal border is the nipple line or fourth intercostal space.

visits, the expert interpretation of US data during weekly image review, and abstracted data for each chart review at least one month after the ED visit.

Both novice and expert interpretation of each lung zone per patient was recorded dichotomously as positive or negative. A lung zone was positive if three or more B-lines were identified. Positivity was also defined according to prior studies based on the pattern and number of lung zones positive.[11,17] For example, "anything positive" is defined as one or more lung zones being interpreted as having ≥3 B-lines. "Minimally positive" is defined as one or more lung zones per side being positive, while "positive" is defined as two or more lung zones per side being positive. "Very positive" refers to three or more lung zones per side being positive, while "all positive" is defined as all lung zones exhibiting ≥3 B-lines. In addition, we assessed whether an abbreviated two-lung zone examination may perform similarly by examining bilateral positivity in superior anterior lung zones (zones 1 and 5) or superior axillary lung zones (zones 3 and 7). The chief reason for focusing on superior lung zones in the two-lung zone method is the interference of the heart in visualization of the left anterior and axillary inferior lung zones, which is common given the prevalence of cardiomegaly in the study population.

Study investigators trained RAs to perform chart abstraction for chart review according to current recommendations.[25,26] RAs used a chart abstraction form that obtained data from the ED resident physician note, initial chest radiograph reading, echocardiography results, initial BNP level, and discharge summary when available. The bedside US results were not included in the chart abstraction. Also, the ED note featured a template system (T-System, Dallas, TX) that had prepopulated history,

symptoms, and signs relevant to undifferentiated dyspnea, allowing for clear chart abstraction. The discharge summaries were dictated by the inpatient resident physicians. Independent verification and error checking by the study investigators were performed on 10% of the charts. All data were entered into Microsoft Excel spreadsheets prior to analysis.

Outcome Measures

The primary outcome for this study was the ability of novice sonographers to perform lung US and to identify sonographic B-lines compared to an expert sonographer. The secondary outcome was the accuracy of novice and expert sonographer interpretation of B-lines to diagnose AHFS, using two independent expert reviewers as the criterion standard. We examined the utility of both a full eight-lung-zone examination and the two-lung-zone examination in diagnosing AHFS.

The expert sonographer conducted weekly reviews on the recorded thoracic US studies. This sonographer had completed 3 years of emergency residency training and a 1-year fellowship in emergency US, in addition to being certified as a registered diagnostic medical sonographer and registered diagnostic cardiac sonographer. The expert sonographer (WCM), who was blinded to the residents' interpretations as well as all clinical patient information, provided independent interpretations of the saved images. Interpretations by the expert sonographer served as the criterion standard for the first study objective.

Two independent physicians, one EP with a research background in heart failure and one cardiologist, reviewed each patient chart using the abstracted chart data. Based on this information, the two expert reviewers independently determined if the principal cause of the patient's dyspnea was AHFS. This method of review is commonly used in clinical trials involving heart failure, as acute heart failure is considered a syndrome with myriad causes and presentations.[11,17]

In 30 cases the expert sonographer, also blinded to US results and discharge diagnosis, served as the tiebreaker when the two expert physician reviewers disagreed on the cause of the patient's dyspnea. In each of these select cases, if the expert sonographer had already reviewed the patient's lung US images, these results were not provided to the expert sonographer during this determination.

Sample Size Calculation

Prior to patient enrollment, a power analysis was performed to determine the total number of patients needed to evaluate the study question. Based on previous studies, we estimated our prevalence of AHFS at 25%, and the sensitivity and specificity would each be 0.80. Assuming an alpha of 0.05, a beta of 0.2, and a desired precision of 0.10, we calculated a sample size of 249 patients. For our secondary outcomes involving ROC analysis, we assigned a null AUC of 0.60, as AUCs under 0.60 indicate poor predictive value. We set our hypothesized AUC at 0.70, as AUCs greater than 0.70 have at least fair predictive value.[24] Using these criteria, our sample size calculation for ROC analysis was 348 patients. We assumed that all resident physicians with

varying levels of US experience could participate and therefore did not account for clustering.

Data Analysis

Using the definitions of positivity, sensitivity, specificity, LRs, and predictive values, along with 95% confidence intervals (CI), were calculated using SPSS, version 22.0. In the assessment of novice sonographer performance in identifying B-lines, the criterion standard was the expert interpretation. In the assessment of B-line lung US in predicting AHFS, the same calculations were performed using both novice and expert sonographer interpretation, with the criterion standard being the expert reviewers' assessment of whether or not AHFS was the primary cause of the patient's dyspnea. To calculate overall diagnostic accuracy, ROC curves were calculated using the same thresholds for positivity (e.g., anything positive, minimally positive, positive, very positive, and totally positive) for expert and novice sonographer interpretation. AUC values with 95% CIs were calculated for both novice and expert sonographer interpretations.

RESULTS

A total of 392 patients were enrolled during the study period. US images were deemed inadequate or incomplete for 12 patients (3%). These patients were excluded, leaving 380 subjects who were included for analysis (Figure 3).

Patients ranged in age from 24 to 87 years old, with a mean (±SD) age of 55 (±12) years. Subjects were primarily African American (93%) and male (55%), with 46% reporting histories of congestive heart failure. Prior to the US examination, 36% of patients had AHFS treatment initiated, either with nitroglycerin or with furosemide. The typical AHFS findings on examination

were seen in less than 70% of patients with AHFS (Table 1). In terms of study sonographers, 66 of a possible 72 EM residents performed the study US, with a range of 1 to 28 examinations (mean ± SD = 5.8 ± 5.1 examinations) and median of three examinations performed per sonographer (Figure 4).

Novice Sonographers' Ability to Identify B-lines

There were 242 patients who each had at least one lung zone with ≥3 B-lines, as interpreted by novice sonographers. The average prevalence of positivity for each lung zone was 22% (range of 16% to 27%). Novice sonographers also interpreted 165 patients as having one or more lung zones positive bilaterally ("minimally positive"), 116 patients with two or more lung zones positive bilaterally ("positive"), 72 patients with three or more lung zones positive bilaterally ("very positive"), and 33 patients with all eight lung zones positive ("totally positive"). In the two-lung zone method, novice sonographers interpreted 84 patients as having bilateral superior anterior positive lung zones and 95 having bilateral superior axillary positive lung zones. In comparison, the expert sonographer interpreted a smaller number of patients as being positive for each of these categories (Figure 5). Table 2 shows the comparison of expert sonographer interpretation of lung US to detect B-lines to inexperienced sonographers.

Novice Versus Expert Interpretation and Primary Diagnosis of AHFS

Regarding the secondary outcome, the primary reason for dyspnea was AHFS in 35% of patients (134 of 380). Overall agreement between experts was 92%, with 30 cases that required the expert sonographer to serve as the arbitrator. Table 3 shows the results for novice sonographers, and Table 4 shows the results for expert sonographers. The AUCs for novice and expert sonographer were 0.77 (95% CI = 0.72 to 0.82) and 0.76 (95% CI = 0.71 to 0.82), respectively (Figure 6).
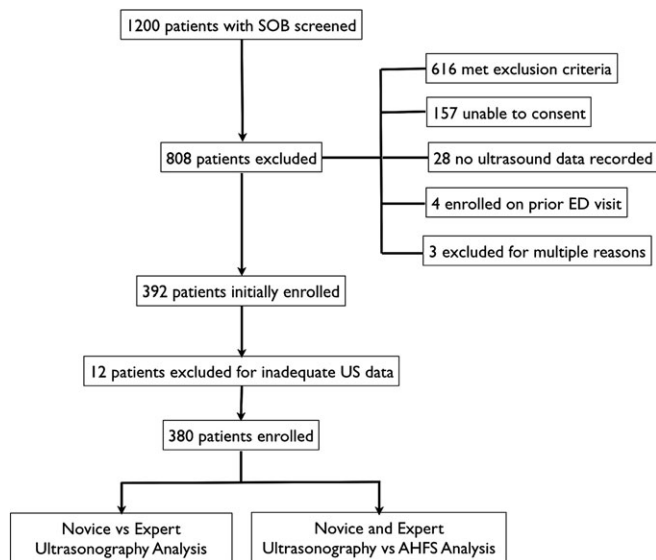

Figure 3. Patients presenting with dyspnea were screened, with 808 patients excluded. An additional 12 subjects were excluded due to inadequate ultrasound data, leaving 380 patients eligible for analysis. SOB = shortness of breath.

Table 1
Patient Demographics and Presenting Symptoms

Characteristic	No AHFS	AHFS
Age > 65 yr (n = 365)	40 (17)	32 (24)
Male (n = 380)	125 (51)	84 (63)
White (n = 363)	16 (7)	4 (3)
African American (n = 363)	213 (92)	126 (95)
History of CHF (n = 312)	48 (25)	95 (81)
BNP > 500 pg/mL (n = 249)	15 (12)	97 (78)
Received treatment prior to US (n = 380)	64 (25)	74 (55)
JVD (n = 171)	6 (7)	52 (62)
Rales or crackles (n = 171)	12 (14)	46 (55)
Bilateral lower extremity edema (n = 174)	21 (24)	52 (61)
S3 gallop (n = 168)	0 (0)	22 (27)

Data are reported as n (%).
AHFS = acute heart failure syndrome; BNP = B-type natriuretic peptide; CHF = congestive heart failure; JVD = jugular venous distension.
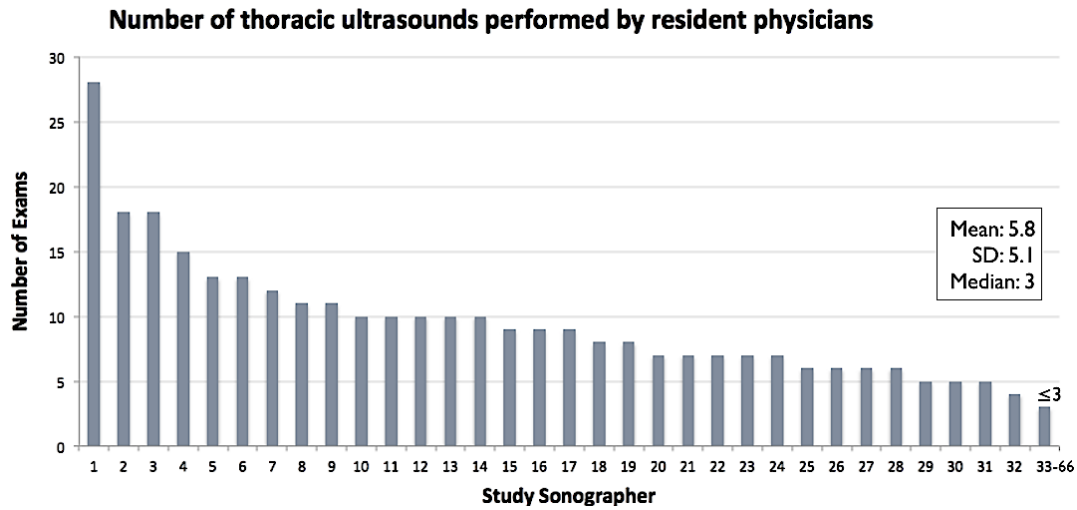
Number of thoracic ultrasounds performed by resident physicians

Mean: 5.8
SD: 5.1
Median: 3

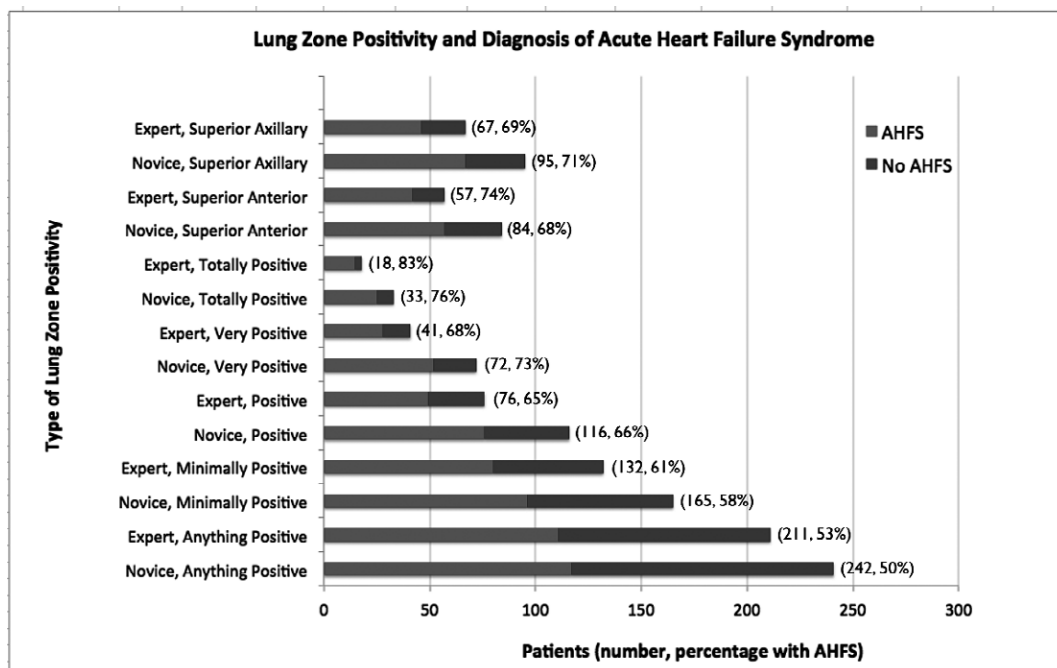**Figure 4.** List of study sonographers and number of lung ultrasound studies performed.

Figure 5. Lung zone positivity and diagnosis of acute heart failure syndrome (AHFS). Note that the percentage with AHFS is synonymous with the positive predictive value.

DISCUSSION

This is one of the first studies to examine a large group of novice sonographers in determining the accuracy of B-line US in the diagnosis of AHFS. Prior studies show sensitivities and specificities of 80% to 90% in the hands of a small group of expert sonographers.[11,12,19] Our study shows that clinicians with brief 30-minute training can obtain adequate images in patients over 95% of the time and interpret B-lines with similar sensitivity and specificity compared to our expert sonographer.

The resident physicians who enrolled patients had typical training in emergency US for EM residency. While some senior residents had performed several hundred US in other areas, no residents had prior training in lung US to assess for B-lines. Therefore, the residents who participated in this study had similar experience in lung US as general EPs without prior training.

Our results suggest that, despite some disagreement between novice and expert sonographer interpretation, the accuracy of B-line US in diagnosing AHFS is similar. While the study sonographers performed the US prior to history taking and physical examination, they were not explicitly blinded to the clinical status of the patient. Therefore, they may have relied on the patient's appearance to increase or decrease the pretest probability of AHFS. Therefore, blinding of the patient's clinical status to the expert sonographer may have lowered

Table 2
Novice Sonographer Test Characteristics for Assessing Sonographic B-lines per Each Lung Zone, Using Expert Sonographer Interpretation as the Criterion Standard

Lung Zone	% with ≥3 B lines	Sensitivity (95% CI)	Specificity (95% CI)	+ LR (95% CI)	−LR (95% CI)	PPV (95% CI)	NPV (95% CI)
3	23	8 (71–87)	82 (76–86)	4.4 (3.3–5.7)	0.2 (0.2–0.4)	64 (55–72)	91 (86–94)
1	23	8 (80–94)	84 (79–88)	5.6 (4.2–7.4)	0.1 (0.1–0.2)	66 (57–74)	95 (92–98)
5	20	8 (80–92)	86 (82–90)	6.2 (4.6–8.4)	0.2 (0.1–0.3)	65 (55–74)	95 (92–97)
7	22	8 (76–91)	83 (78–87)	5.1 (3.8–6.7)	0.2 (0.1–0.3)	64 (55–73)	94 (90–97)
4	27	8 (80–93)	85 (80–89)	5.8 (4.3–7.4)	0.1 (0.1–0.2)	72 (63–79)	94 (90–97)
2	21	8 (74–90)	80 (75–85)	4.3 (3.3–5.3)	0.2 (0.1–0.3)	59 (50–67)	94 (89–96)
6	16	8 (79–95)	85 (80–88)	5.8 (4.4–7.7)	0.1 (0.1–0.3)	57 (47–66)	97 (94–99)
8	22	8 (74–90)	85 (80–89)	5.5 (4.1–7.3)	0.2 (0.1–0.3)	66 (57–74)	93 (89–96)
Overall	22	85 (83–88)	84 (82–85)	5.2 (4.7–5.8)	0.2 (0.1–0.2)	64 (61–67)	94 (93–95)

Lung zones are listed in descending order from right to left, superior then inferior. For example, lung zone 3 is right superior axillary, while lung zone 4 is right inferior axillary. All except LRs are percentages.
+LR = positive likelihood ratio; −LR = negative likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.

Table 3
Novice Sonographer Test Characteristics for Assessing Pulmonary Edema From Acute Heart Failure Syndrome, Using A Two-expert Chart Review Method as the Criterion Standard

Lung Zones Positive	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	PPV (95% CI)	NPV (95% CI)
Anything positive (≥1 of 8)	87 (81–92)	49 (42–55)	1.7 (1.5–2.0)	0.3 (0.2–0.4)	50 (42–55)	88 (81–92)
Minimally positive (1 ≥ per side)	72 (64–79)	72 (66–77)	2.5 (2.0–3.2)	0.4 (0.3–0.5)	58 (50–66)	82 (76–87)
Positive (≥2 per side)	57 (48–65)	84 (79–88)	3.5 (2.5–4.8)	0.5 (0.4–0.6)	66 (56–74)	78 (72–83)
Very positive (≥3 per side)	40 (32–49)	92 (88–95)	4.9 (3.1–7.9)	0.7 (0.6–0.8)	73 (61–82)	74 (68–78)
Totally positive (4 per side)	19 (13–26)	97 (94–98)	5.7 (2.6–12.3)	0.8 (0.8–0.9)	76 (57–88)	68 (63–73)
Superior anterior (zones 1 and 5)	43 (35–51)	89 (85–92)	3.9 (2.6–5.8)	0.7 (0.6–0.8)	68 (57–77)	74 (69–79)
Superior axillary (zones 3 and 7)	51 (42–59)	89 (84–92)	4.5 (3.0–6.6)	0.6 (0.5–0.7)	71 (60–79)	77 (72–82)

Increasing level of lung zones positive for B-lines is listed in descending order. All except LRs are percentages.
+LR = positive likelihood ratio; −LR = negative likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.

Table 4
Expert Sonographer Test Characteristics for Assessing Pulmonary Edema From Acute Heart Failure Syndrome, Using a Two-expert Chart Review Method as the Criterion Standard

Lung Zones Positive	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	−LR (95% CI)	PPV (95% CI)	NPV (95% CI)
Anything positive (≥1 of 8)	85 (78–90)	58 (52–64)	2.1 (1.7–2.4)	0.3 (0.2–0.4)	53 (46–59)	88 (82–92)
Minimally positive (1 ≥ per side)	62 (53–70)	78 (73–83)	2.8 (2.2–3.7)	0.5 (0.4–0.6)	61 (52–69)	79 (73–84)
Positive (≥2 per side)	38 (30–46)	89 (84–92)	3.4 (2.2–5.1)	0.7 (0.6–0.8)	65 (53–75)	72 (67–77)
Very positive (≥3 per side)	22 (15–29)	95 (91–97)	4.0 (2.1–7.4)	0.8 (0.7–0.9)	68 (52–81)	69 (64–74)
Totally positive (4 per side)	12 (7–18)	99 (96–100)	9.2 (2.7–31.3)	0.9 (0.8–1.0)	83 (58–96)	67 (62–72)
Superior anterior (zones 1 and 5)	31 (24–40)	94 (90–96)	5.1 (3.0–8.9)	0.7 (0.6–0.8)	74 (60–84)	72 (66–76)
Superior axillary (zones 3 and 7)	34 (27–43)	92 (87–94)	4.0 (2.5–6.4)	0.7 (0.6–0.8)	69 (56–79)	72 (66–77)

Increasing level of lung zones positive for B-lines is listed in descending order. All except LRs are percentages.
+LR = positive likelihood ratio; −LR = negative likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.

test characteristics, just as exposure to clinical status may have helped raise resident sonographers' test characteristics. This may lead to a bias toward the null; i.e., there is no difference in test characteristics between novice and expert sonographers and the diagnosis of acute heart failure. However, while the pretest probability should change based on exposure to clinical status, the LRs usually should not change. Therefore, the LRs for both expert and novice sonographers were compa-

rable to that of prior studies by Liteplo et al.[11] Of note, the lower end of the 99% CI range for +LRs reported in this previous study was under 2.0 for all degrees of positivity.[11] As expected from our large sample size, our CI ranges were narrower for every level of positivity reported, with most above 2.0. In addition, Anderson et al.[12] reported sensitivity and specificity of two or more positive lung zones bilaterally of 34 and 91%, respectively. This compares favorably to our expert
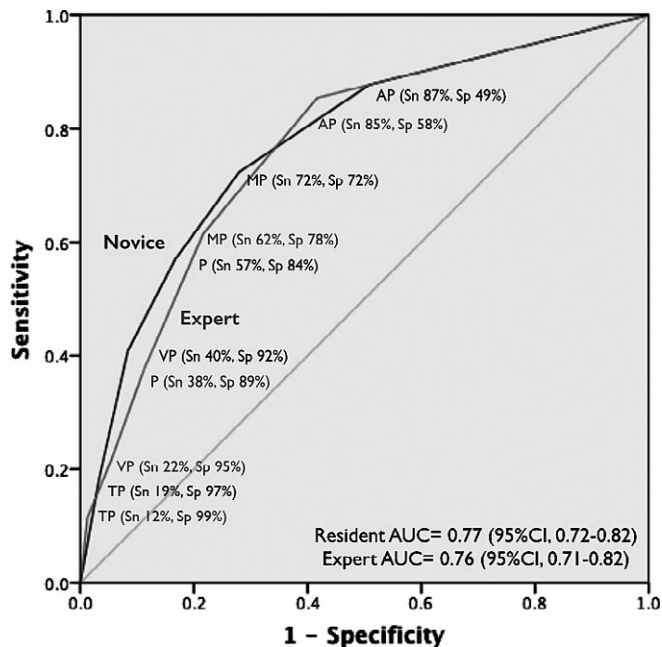
Figure 6. Receiver operating characteristic curves for novice sonographer and expert sonographer thoracic B-line ultrasonography in diagnosing pulmonary edema from AHFS. AUC = area under the curve; AP = anything positive; MP = minimally positive; P = positive; Sn = sensitivity; Sp = specificity; TP = totally positive. VP = very positive.

performance of 38 and 89%, as well as our novice performance of 57 and 84%, for two or more positive bilateral lung zones. Therefore, while we acknowledge a potential bias toward the null for expert sonographers who were not provided clinical data, we also believe that our data are consistent with those of prior studies.

Similarly, the expert reviewers were blinded to the US results, but not to each patient's history and physical information. In a case where a resident physician's B-line interpretation agreed with the expert reviewer diagnosis of AHFS, but not with the expert sonographer's B-line interpretation, this may have been due to exposure to the patient's clinical status. The expert sonographer, however, was blinded to clinical status and his accuracy relied solely on the definition of ≥3 B-lines per zone for positivity. Therefore, in false-positive cases by novice sonographers due to clinical exposure, the expert sonographer interpretation would have accurately interpreted the examination as negative for B-lines.

In our comparison of a two-lung-zone approach versus assessment of all eight lung zones, the former approach had high specificity but low sensitivity. This is clinically useful, as this simplified examination can identify pulmonary edema when the lung zones are positive. Therefore, treatment may be initiated solely on bilaterally positive lung zones. However, if this approach yields negative lung zones, then clinicians may elect to evaluate the other six lung zones to increase test sensitivity.

LIMITATIONS

Image acquisition for B-line US was technically easy, as our expert sonographer deemed only 3% of our 392

studies inadequate. No expert sonographers were present during image acquisition to assess novice technique, so we cannot assess whether inadequate scans were due to improper technique or patient factors such as morbid obesity. While this is a limitation, the authors believe that these methods more closely represent the real-world setting where most EPs are not experts at bedside US. More in-depth training may improve the rate of adequate scans as well as interpretation accuracy, but there is no agreed-upon optimal number of supervised scans to achieve proficiency. In addition, more experience with the procedure should lead to improved accuracy. However, this was not assessed by our study.

As pulmonary edema has many potential causes, we considered comparison of US to chest radiography, but there are two limitations to this approach. First, a chest x-ray is not the criterion standard for assessing pulmonary edema. EP interpretation of pulmonary edema pattern on chest x-ray agrees with a radiologist in less than 50% of cases.[27,28] Between resident and attending radiologists, the agreement in diagnosis for pulmonary edema in heart failure is fair, with a kappa of 0.29.[29] This is important, because the radiologist interpretation is often not contemporaneous with initial patient management, and EPs are likely to miss a significant proportion of pulmonary edema based on their interpretation of chest radiographs. Second, the vast majority of pulmonary edema is due to acute left-sided heart failure, regardless of primarily systolic or diastolic dysfunction. We therefore decided to use a method of strict chart review by expert physicians to establish the most likely diagnosis of AHFS, a method that has been used in prior AHFS studies, since the diagnosis for AHFS relies on a constellation of historical, clinical, laboratory, and imaging data. There is no single diagnostic test to evaluate whether a patient's presenting symptom of dyspnea is primarily due to AHFS. A multiphysician chart review provides the best surrogate for this test.[11,12,19]

There are pertinent false-positives and false-negatives to mention. Subjects may have pulmonary edema, but due to noncardiogenic causes. Expert reviewers would assign another cause of dyspnea in these cases other than AHFS, leading to false-positives. Likewise, patients may have acute right heart failure or primarily systemic congestion with paucity of pulmonary congestion, which would lead to false-negatives. In addition, in cases of flash pulmonary edema from acute diastolic dysfunction, the BNP elevation may be delayed by several hours.[30] Subjects who had BNP levels assessed therefore may have pulmonary edema on US but not elevated BNPs, unless repeat BNPs were performed. Therefore, the expert reviewers may have been influenced to interpret a subject with pulmonary edema and a normal BNP as due to a noncardiogenic cause. This was rare, as fewer than five subjects had pulmonary edema from noncardiogenic causes as their primary diagnoses.

Our study design also likely excluded two extreme spectra of dyspneic patients: those who, by attending physician determination, had other clear sources for dyspnea than acute heart failure, and those who either

were on positive pressure ventilation or too ill to consent. The former group had an extremely low pretest probability of acute heart failure, while the latter group had a high pretest probability of acute heart failure. By eliminating these groups from our study, we are able to increase the utility of B-line US, since it is in patients with intermediate pretest probability of acute heart failure that this procedure would benefit most in diagnosis and management.

In addition, given the large number of calculations performed, we considered the increased probability of false-positives generated due to multiple comparisons. However, these comparisons are highly correlated, and the test characteristics agree well with prior similar studies. So while we decided not to account for increased Type I error, we believe that our results are valid based on these points.

Last, this was a single-site, prospective, cross-sectional study, so findings for novice sonographers may not apply to other institutions. No patient-centered outcomes were assessed, nor did we examined the added utility in diagnosis of AHFS. Future research plans include assessing the effect of B-line US on clinical decision-making, as well as on its ability to affect patient outcomes such as intensive care unit length of stay, ventilator days, the rate of worsening renal function, and inpatient mortality. We also did not study the effect of number of US performed on interpretative accuracy, nor the minimum number of scans needed to achieve competency.

CONCLUSIONS

Inexperienced sonographers can identify sonographic B-lines with greater than 80% sensitivity and specificity compared to an expert sonographer after a brief tutorial. Lung ultrasonography has fair predictive value for pulmonary edema from acute heart failure in the hands of both novice and expert sonographers.

References

1. Peacock WF, Braunwald E, Abraham W, et al. National Heart, Lung, and Blood Institute Working Group on emergency department management of acute heart failure. J Am Coll Cardiol 2010;56:343–51.
2. Gheorghiade M, Zannad F, Sopko G, et al. Acute heart failure syndromes: current state and framework for future research. Circulation 2005;112:3958–68.
3. Collins SP, Storrow AB, Kirk JD, Pang PS, Diercks DB, Gheorghiade M. Beyond pulmonary edema: diagnostic, risk stratification, and treatment challenges of acute heart failure management in the emergency department. Ann Emerg Med 2008;51:45–57.
4. Silvers SM, Howell JM, Kosowsky JM, Rokos IC, Jagoda AS. Clinical policy: critical issues in the evaluation and management of adult patients presenting to the emergency department with acute heart failure syndromes. Ann Emerg Med 2007;49:627–68.
5. Weintraub NL, Collins SP, Pang PS, et al. Acute heart failure syndromes: emergency department presentation, treatment, and dispositions: current approaches and future aims. A scientific statement from the American Heart Association. Circulation 2010;122:1975–96.
6. Vader JM, Drazner MH. Clinical assessment of heart failure: utility of symptoms, signs, and daily weights. Heart Fail Clin 2009;5:149–60.
7. Wang CS, FitzGerald JM, Schulzer M, et al. Does this dyspneic patient in the emergency department have congestive heart failure? JAMA 2005;294:1944–56.
8. American College of Emergency Physicians Board of Directors. Emergency Ultrasound Guidelines. ACEP Policy Statement. Available at: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCAQFjAA&url=http%3A%2F%2Fwww.acep.org%2Fworkarea%2Fdownloadasset.aspx%3Fid%3D32878&ei=tHrbVOygBoTwUJjCgugE&usg=AFQjCNH9DDA5uttBic8rMaCgbbr7lrHBnA&sig2=i0ZivqbtTPoyTtn01Sdfuw&bvm=bv.85761416,d.d24&cad=rja. Accessed Feb 11, 2015.
9. Nazerian P, Vanni S, Zanobetti M, et al. Diagnostic accuracy of emergency Doppler echocardiography for identification of acute left ventricular heart failure in patients with acute dyspnea: comparison with Boston Criteria and N-terminal Pro BNP. Acad Emerg Med 2010;17:18–26.
10. Sharma GV, Woods PA, Lambrew CT, et al. Evaluation of a noninvasive system for determining left ventricular filling pressure. Arch Intern Med 2002;162:2084–8.
11. Liteplo AS, Marill KA, Villen T, et al. Emergency thoracic ultrasound in the differentiation of the etiology of shortness of breath (ETUDES): sonographic B-lines and N-terminal pro-brain-type natriuretic peptide in diagnosing congestive heart failure. Acad Emerg Med 2009;16:201–10.
12. Anderson KL, Jeng KY, Fields JM, Panebianco NL, Dean AJ. Diagnosing heart failure among acutely dyspneic patients with cardiac, inferior vena cava, and lung ultrasonography. Am J Emerg Med 2013;31:1208–14.
13. Cibinel GA, Casoli G, Elia F, et al. Diagnostic accuracy and reproducibility of pleural and lung ultrasound in discriminating cardiogenic causes of acute dyspnea in the emergency department. Intern Emerg Med 2012;7:65–70.
14. Lichtenstein D, Meziere G, Biderman P, Gepner A, Barre O. The comet-tail artifact. An ultrasound sign of alveolar-interstitial syndrome. Am J Respir Crit Care Med 1997;156:1640–6.
15. Lichtenstein D, Meziere G. A lung ultrasound sign allowing bedside distinction between pulmonary edema and COPD: the comet-tail artifact. Intensive Care Med 1998;24:1331–4.
16. Lichtenstein DA. Ultrasound in the management of thoracic disease. Crit Care Med 2007;35(5 Suppl):S250–61.

17. Volpicelli G, Mussa A, Garofalo G, et al. Bedside lung ultrasound in the assessment of alveolar-interstitial syndrome. Am J Emerg Med 2006;24: 689–96.
18. Frassi F, Gargani L, Tesorio P, Raciti M, Mottola G, Picano E. Prognostic value of extravascular lung water assessed with ultrasound lung comets by chest sonography in patients with dyspnea and/or chest pain. J Card Fail 2007;13:830–5.
19. Gargani L, Frassi F, Soldati G, Tesorio P, Gheorghiade M, Picano E. Ultrasound lung comets for the differential diagnosis of acute cardiogenic dyspnoea: a comparison with natriuretic peptides. Eur J Heart Fail 2008;10:70–7.
20. Manson WC, Bonz JW, Carmody K, Osborne M, Moore CL. Identification of sonographic B-lines with linear transducer predicts elevated B-type natriuretic peptide level. West J Emerg Med 2011;12:102–6.
21. Prosen G, Klemen P, Strnad M, Grmec S. Combination of lung ultrasound (a comet-tail sign) and N-terminal pro-brain natriuretic peptide in differentiating acute heart failure from chronic obstructive pulmonary disease and asthma as cause of acute dyspnea in prehospital emergency setting. Crit Care 2011;15:R114.
22. Goonewardena SN, Gemignani A, Ronan A, et al. Comparison of hand-carried ultrasound assessment of the IVC and N-terminal pro-BNP for predicting readmission after hospitalization for acute decompensated heart failure. JACC Cardiovasc Imaging 2008;1:595–601.
23. Volpicelli G, Elbarbary M, Blaivas M, et al. International evidence-based recommendations for point-of-care lung ultrasound. Intensive Care Med 2012;38:577–91.
24. Metz CE. Basic principles of ROC analysis. Semin Nucl Med 1978;8:283–98.
25. Gilbert EH, Lowenstein SR, Koziol-McLain J, Barta DC, Steiner J. Chart reviews in emergency medicine research: where are the methods? Ann Emerg Med 1996;27:305–8.
26. Kaji AH, Schriger D, Green S. Looking through the retrospectoscope: reducing bias in emergency medicine chart review studies. Ann Emerg Med 2014;64:292–8.
27. Al aseri Z. Accuracy of chest radiography interpretation by emergency physicians. Emerg Radiol 2009;16:111–4.
28. Gatt ME, Spectre G, Paltiel O, Hiller N, Stainikowicz R. Chest radiographs in the emergency department: is the radiologist really necessary? Postgrad Med J 2003;79:214–7.
29. Feldmann EJ, Jain VR, Rakoff S, Haramati LB. Radiology residents' on-call interpretation of chest radiographs for congestive heart failure. Acad Radiol 2007;14:1264–70.
30. Maisel A, Mueller C, Adams K Jr, et al. State of the art: using natriuretic peptide levels in clinical practice. Eur J Heart Fail 2008;10:824–39.