CrossMark

ORIGINAL RESEARCH

Lung real time three-dimensional imaging in critically ill ventilated patients: a global diagnosis concordance study

Jean-Marc Le Goff[1] · Claire Roger[1] · Benjamin Louart[1] · Pierre Géraud Claret[1] ·
Aurélien Daurat[1] · Stéphanie Bulyez[1] · Jean-Yves Lefrant[1] · Xavier Bobbia[1] ·
Laurent Muller[1]

Abstract Lung ultrasound (LUS) increases clinical diagnosis performance in intensive care unit (ICU). Real-time three-dimensional (3-D) imaging was compared with two-dimensional (2-D) LUS by assessing the global diagnosis concordance. In this single center, prospective, observational, pilot study, one trained operator performed a 3-D LUS immediately after a 2-D LUS in eight areas of interest on the same areas in 16 ventilated critically ill patients. All cine loops were recorded on a computer without visible link between 2-D and 3-D exams. Two experts blindly reviewed cine loops. Four main diagnoses were proposed: normal lung, consolidation, pleural effusion and interstitial syndrome. Fleiss κ and Cohen's κ values were calculated. In 252 LUS cine loops, the concordance between 2-D and 3-D exams was 83.3% (105/126), 77.6% (99/126) and 80.2% (101/126) for the trained operator and the experts respectively. The Cohen's κ coefficient value was 0.69 [95% Confidence Interval (CI) 0.58–0.80] for expert 1 meaning a substantial agreement. The inter-rater reliability was very good (Fleiss' κ value=0.94 [95% CI 0.87–1.0]) for 3-D exams. The Cohen's κ was excellent for pleural effusion (κ=0.93 [95% CI 0.76–1.0]), substantial for normal lung diagnosis (κ=0.68 [95% CI 0.51–0.86]) and interstitial syndrome (κ=0.62 [95% CI 0.45–0.80]) and fair for consolidation diagnoses (κ=0.47 [95% CI 0.30–0.64]). In ICU ventilated patients, there was a substantial concordance between 2-D and 3-D LUS with a good inter-rater reliability. However, the diagnosis concordance for lung consolidation is poor.

Keywords Lung ultrasound · LUS · 3-Dimensional ultrasonography · ICU · Pilot study · Concordance

1 Introduction

Lung ultrasonography (LUS) has been increasingly used as a surrogate of chest X-ray in the intensive care unit (ICU) since more than 15 years. Lung ultrasound is the second type of ultrasound exam performed in critically ill patients [1]. The diagnosis accuracy of LUS is superior to conventional chest X-ray in ICU or emergency department, especially for pleural effusion, atelectasis and pneumothorax [2–4]. Finally, the daily use of LUS was shown to be associated with a reduction of the number of computerized tomography (CT) scanner or conventional chest X-ray without increasing the morbi–mortality [5–7].

Lichtenstein and coworkers emphasized the idea that a very simple and low-cost device with a microconvex 5 MHz probe was sufficient for LUS in ICU patients. The feasibility of LUS with such device was largely demonstrated [3].

However, the rapid development of new ultrasound techniques such as 3-D imaging raises the question of their potential usefulness for LUS. The 3-D LUS semiology has never been described. Therefore, the primary objective of the present study was to determine the global diagnosis concordance between 2-D and 3-D LUS in ventilated critically ill patients. The secondary endpoint aimed at assessing the inter-observer reliability. Finally, the concordance between 2-D and 3-D exam was assessed considering four

✉ Jean-Yves Lefrant
jean-yves.lefrant@wanadoo.fr

1 Division of Anaesthesia Intensive Care, Pain and Emergency Medicine, University Hospital of Nîmes, Place du Professeur Robert Debré, 30029 Nîmes, France

Springer

specific lung diagnoses: normal lung, interstitial syndrome, consolidation and pleural effusion.

2 Materials and methods

2.1 Study design

The present trial was a prospective, single center, observational, pilot study. As it was strictly observational, the Institutional Review Board (IRB) approved the study protocol (no. 15/07.06, IRB of Centre Hospitalier Universitaire de Nîmes) before the enrollment of the first patient. According to the French law, the IRB waived patient informed consent. However, an information letter was given to each patient when conscious or relatives who could refuse to participate. Anonymous data were declared to the French Data protection authority (CNIL Commission Nationale de l'Informatique et des Libertés) (declaration number: 1878494 30th July 2015).

Patients were enrolled in the study from July the 20th to August the 10th 2015.

2.2 Population description

Inclusion criteria were: patient ≥18-year-old admitted to ICU, invasive ventilation and need for LUS examination.

Non-inclusion criteria were: age <18, patients without social insurance, refusal for non-sedated patient, chest drainage and extensive dressings.

Moreover, patients in whom LUS examination was not available were excluded.

2.3 Lung ultrasound protocol

The imaging was performed with EPIQ 7 with a probe X5-1 (5–1 MHz Matrix transducer, Philips®, Eindhoven, Netherlands) allowing 2-D and 3-D real time examinations.

In a first time, a non expert trained operator (defined as a daily LUS practice ≥5 years), performed a 2-D LUS in eight lung areas (four on the right side and four on the left side) in each patient as described by international guidelines published in 2012 [8] (Fig. 1). Immediately after, the same operator performed a 3-D LUS with the same ultrasonography machine. A simple switch of the machine from 2-D to 3-D permitted to perform the two exams in 2-D and 3-D LUS, respectively.

For 2-D and 3-D analysis, one cine loop over five cardiac cycles was stored on a computer hard drive.

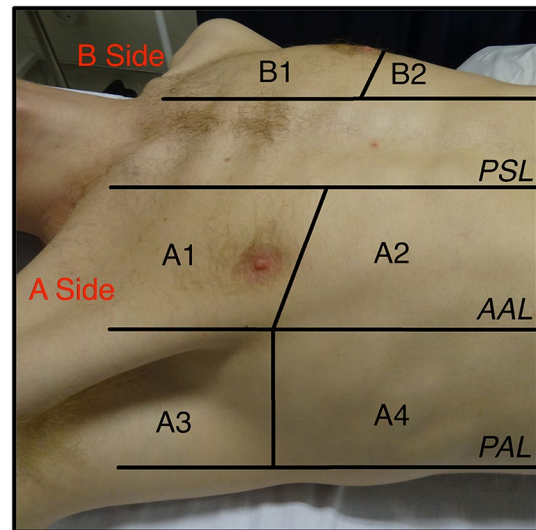According to literature, four main LUS diagnoses were possible [8]:

Fig. 1 Areas of interest for LUS according to literature [8]. *A* side denotes the *right* side of patient; *B* is for the *left* side. *PSL* parasternal line, *AAL* anterior axillary line, *PAL* posterior axillary line
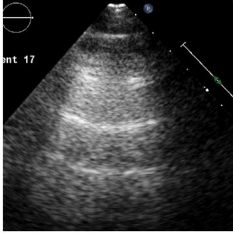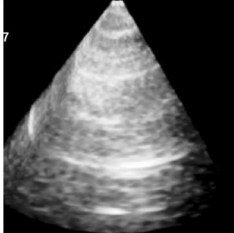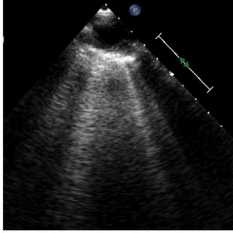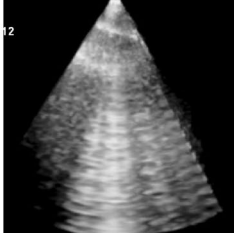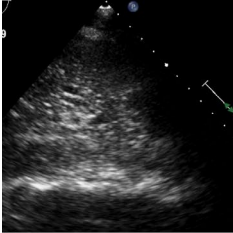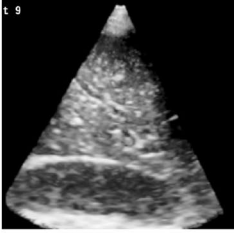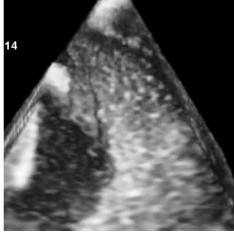
- Normal lung: defined as sliding pleural line with A-lines without B-lines patterns.
- Interstitial syndrome: defined as presence of B-lines, discrete laser-like vertical hyperechoic reverberation artifacts arising from pleural line to the bottom of the ultrasonography screen without fading, moving synchronously with lung sliding.
- Pleural effusion: defined as a generally hypoechoic space between parietal and visceral pleura with or without consolidation.
- Consolidation without pleural effusion: defined as a subpleural echo-poor region or one tissue-like echo-texture.
- Because the study period was short (3 weeks), there were not enough patients with pneumothorax for analyzing this diagnosis.

In a second time, two experts (1 and 2) performed a review of these loops. Expert was defined as an intensivist with a national diploma of Critical Care ultrasonography with a daily practice of LUS since more than 5 years and regular teaching of ultrasonography in medical university.

All exams were blinded and stored without visible connection between 2-D and 3-D exams. The experts had to make a diagnosis between those four possibilities and to quote by a numerical scale the quality of the exam from 0 (poorest quality) to 10 (best quality).

The experts blindly and randomly analyzed the exams on a computer (Fig. 2).

Fig. 2 Comparative exams 2-D and 3-D for main diagnoses. *2-D* two-dimensional, *3-D* three-dimensional

	2-D	3-D
Normal Lung		
Interstitial Syndrome		
Consolidation		
Pleural effusion		

2-D is for two-dimensional and 3-D for three-dimensional.

2.4 Studied parameters

In each patient, sex, age, main diagnosis and Simplified Acute Physiology Score II (SAPSII) score [9] were recorded.

2.5 Statistical analysis

Assumptions for the sample size were based on our daily practice in our ICU. The number of patients needed was calculated in order to be able to estimate a kappa (κ) coefficient over 0.6 [10, 11] with a one-sided alpha-risk of 5%, a power of 90% and expecting the following proportions: normal lung (40%), interstitial syndrome (30%), lung consolidation (10%) and pleural effusion (20%). Therefore, 114 LUS examinations were mandatory.

We tested the concordance by unweighted κ test of Cohen [11] between 2-D and 3-D LUS diagnosis for expert 1 to the primary end point.

For secondary objectives, the inter-rater reliability was tested for 3-D LUS by a Fleiss' Kappa [12], unweighted κ of Cohen for main diagnoses were calculated and quality evaluation scale value were compared by Wilcoxon signed-rank test for paired samples. A value of κ ranges from −1 (total disagreement) to 1 (perfect agreement), a value of 0 indicates no agreement than chance [11]. We used the Landis and Koch table to characterize κ values: <0 indicating no agreement, from 0 to 0.20 a slight agreement, from 0.21 to 0.40 a fair agreement, from 0.41 to 0.60 a moderate agreement, from 0.61 to 0.80 a substantial agreement and from 0.81 to 1.0 an almost perfect agreement [13].

Data are expressed in mean value with standard deviation (SD). Qualitative variables were expressed in absolute number with percentage (%).

Statistical significance was set at 5% for every test used.

3 Results

The study enrolled 16 patients. Demographic data and the mains of ICU admission are shown in Table 1. One hundred and twenty-eight cine loops of LUS examinations were performed in 2-D and 126 cine loops were performed in 3-D. Two cine loops were excluded because of poor quality in 2-D exams leading to analyze 126 cine loops in 2-D and 126 in 3-D (Fig. 2).

3.1 Primary endpoint

The ultrasound semiology described for 2-D LUS was the same in 3-D LUS (Fig. 3). The observed concordance between 2-D and 3-D exam was 83.3% (105/126) for the operator, 80.2% (101/126) for expert 1 and 77.6% (99/126) for expert 2.

The Cohen's κ value was 0.69 [95% CI 0.58–0.80] for expert 1 meaning a substantial agreement between 2-D and 3-D diagnosis.

3.2 Secondary endpoints

The inter-rater reliability was very good according the Fleiss' κ value, which was 0.94 [95% CI 0.87–1.0] for 3-D exams (Table 2).

Comparing with main diagnosis, for expert 1, the Cohen's κ was excellent for pleural effusion (κ=0.93 [95% CI 0.76–1.0]), substantial for normal lung diagnosis (κ=0.68 [95% CI 0.51–0.86]) or interstitial syndrome (κ=0.62 [95% CI 0.45–0.80]) and fair for consolidation diagnoses (κ=0.47 [95% CI 0.30–0.64]) (Table 3).

Table 1 Demographic data

Age, mean (SD) (years)	64.5 (16.5)
Male sex, no. (%)	13 (81.3)
SAPS II, mean (SD)	48.9 (25.5)
Body mass index, mean (SD)	25.0 (5.88)
Diagnoses	
Septic shock, no. (%)	3 (18.8)
Trauma, no. (%)	2 (12.5)
Cardiogenic pulmonary oedema, no. (%)	1 (6.3)
Cardiac arrest, no. (%)	3 (18.8)
Acute pancreatis, no. (%)	1 (6.3)
Stroke, no. (%)	2 (12.5)
Acute respiratory distress syndrome, no. (%)	3 (18.8)
Rupture of abdominal aortic aneurysm, no. (%)	1 (6.3)

SD standard deviation, *SAPS II* Simplified Acute Physiology Score II calculated from 12 physiological variables and three disease-related variables during 24 h period; scores range from 0 to 163 with higher scores indicating more severe disease
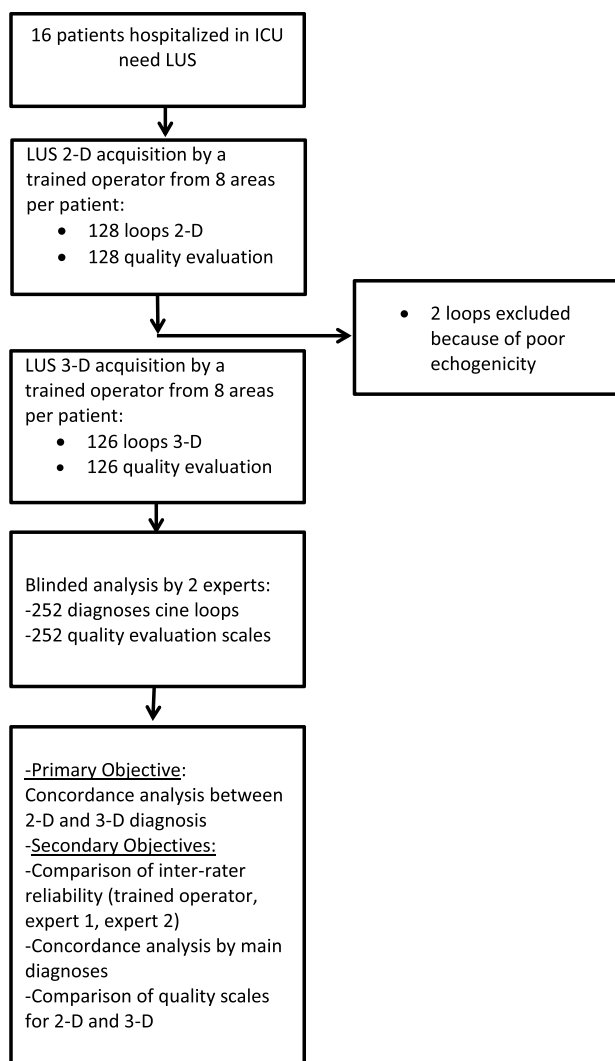

Fig. 3 Flow chart. *ICU* intensive care unit, *LUS* lung ultrasonography, *2-D* two-dimensional, *3-D* three-dimensional

The quality numerical scales were 6.8 ± 1.6 in 2-D versus 6.7 ± 1.4 in 3-D for the operator (p=0.50), 7.3 ± 1.7 in 2-D versus 7.2 ± 1.9 in 3-D for expert 1 (p=0.76)

Table 2 Concordance tests for 2-D and 3-D imaging by using the Landis and Koch scale [12]

	Kappa	95% CI	Agreement
Unweighted Kappa for first expert	0.69	[0.58–0.80]	Substantial
Diagnosis-wise Kappa			
Normal lung	0.68	[0.51–0.86]	Substantial
Interstitial syndrome	0.62	[0.45–0.80]	Substantial
Consolidation	0.47	[0.30–0.64]	Moderate
Pleural effusion	0.93	[0.76–1.0]	Almost perfect

CI confidence interval, *2-D* two-dimensional, *3-D* three-dimensional

Table 3 Inter-rater reliability for 2-D and 3-D exams [11, 12]

	2-D Fleiss' Kappa [95% CI]	3-D Fleiss' Kappa [95% CI]
All diagnoses	0.95 [0.88–1.0]	0.94 [0.87–1.0]
Normal lung	0.96 [0.87–1.0]	0.93 [0.83–1.0]
Interstitial syndrome	0.95 [0.85–1.0]	0.92 [0.82–1.0]
Consolidation	0.93 [0.83–1.0]	0.87 [0.77–0.97]
Pleural effusion	0.95 [0.85–1.0]	1.0 [0.9–1.0]

CI confidence interval, *2-D* two-dimensional, *3-D* three-dimensional

and 7.5 ± 1.4 in 2-D versus 7.1 ± 1.7 in 3-D for expert 2 ($p = 0.09$) (Fig. 1).

4 Discussion

The present study showed a substantial agreement ($\kappa = 0.69$) between 2-D and 3-D real-time LUS. The interoperator reliability was excellent ($\kappa = 0.95$). Nevertheless, the concordance was poor for lung consolidation ($\kappa = 0.47$) whereas it was substantial for normal lung ($\kappa = 0.68$), interstitial syndrome ($\kappa = 0.62$) and almost perfect for pleural effusion ($\kappa = 0.92$).

Modern ultrasound machines allow real-time 3-D ultrasound imaging (4-D). The usefulness of the technique has been described in fetal echocardiography [14]. It has also been used in adult echocardiography for cardiac chamber quantification or valvular disease [15, 16]. To the best of our knowledge, this is the first study showing the feasibility of LUS by using a 3-D ultrasound probe.

Even if daily practice of LUS was recommended in critically ill patient, it is rarely used in ICU as it represents only 20% of the ultrasound exams [1]. This could reflect that this technique is not easy to implement and to teach, especially because basic images are more based on artifacts than on anatomical structures. This is particularly true for lung sliding and interstitial syndrome (B lines artifacts). In order to improve quality of images and diagnosis accuracy, it could be hypothesized that 3-D ultrasound could be useful.

In clinical practice, the present study shows that the global quality of imaging with 2-D is different from 3-D LUS without determining which is the better. This was particularly true for lung consolidation. The 2-D semiology for consolidation syndrome is probably not strictly concordant to 3-D semiology. As CT scanner imaging was not performed in the present study, this raises the need for describing a specific semiology for 3-D LUS. Moreover, considering the price of ultrasonography machine with 3-D probes (comparing to 2-D conventional devices), it is at present hard to see how this could be developed outside the context of clinical research for describing the imaging of LUS in 3-D and reporting its clinical implications.

The second difficulty for 3-D LUS use in clinical practice is the volume data acquisition of images. When positioning the 3-D LUS probe on the thorax, the physician obtains a cone-shaped image. By using machine trackball, this cone can be moved in every position of space. It is hard to determine the optimal position of such cone in order to get the best diagnosis performance.

Some limitations should be advanced.

First, the number of patients is low. This is due to the fact that the ultrasonography machine was lent to our unit for only 3 weeks. We increased the statistical power of this study by performing an important number of exams of each region of interest. This leaded to analyze 252 LUS exams. Second, the post hoc analysis of videos loops probably reduces the diagnosis accuracy because the experts cannot control the transducer position on thorax. The numerical quality scale was 7.3 ± 1.8 in 2-D and 7.2 ± 1.9 in 3-D for expert 1. These two values are moderate meaning that post hoc computer analysis is more difficult than performing a real-time ultrasound examination. Third, we choose to present 3-D loops in same orientation to 2-D loops. That implies a loss of a great part of interest of 3-D imaging. Forth, we used the cardiac phased array probe whereas the LUS has been extensively described by using a 5 MHz microconvex probe. However, International evidence-based recommendations for point-of-care lung untrasound proposed that the choice of the transducer could be based on physician preference [17]. Finally, we did not compare 2-D and 3-D LUS exams to a gold standard as CT scanner.

5 Conclusion

This first pilot study shows a substantial concordance between 2-D and 3-D LUS diagnoses in ventilated critically ill patients. The inter-observer reliability was good. Nevertheless, the diagnosis concordance for lung consolidation was poor. As no comparison with CT scan was performed, larger studies are needed to precise the exact role for 3-D LUS in clinical practice.

Compliance with ethical standards

Conflict of interest The authors declare no competing interest.

References

1. Zieleskiewicz L, Muller L, Lakhal K, Meresse Z, Arbelot C, Bertrand P-M, et al. Point-of-care ultrasound in intensive care units: assessment of 1073 procedures in a multicentric, prospective, observational study. Intensive Care Med. 2015;41(9):1638–47.

2. Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure: the BLUE protocol. Chest. 2008;134(1):117–25.

3. Lichtenstein D, Goldstein I, Mourgeon E, Cluzel P, Grenier P, Rouby J-J. Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome. Anesthesiology. 2004;100(1):9–15.

4. Vignon P, Chastagner C, Berkane V, Chardac E, François B, Normand S, et al. Quantitative assessment of pleural effusion in critically ill patients by means of ultrasonography. Crit Care Med. 2005;33(8):1757–63.

5. Xirouchaki N, Magkanas E, Vaporidi K, Kondili E, Plataki M, Patrianakos A, et al. Lung ultrasound in critically ill patients: comparison with bedside chest radiography. Intensive Care Med. 2011;37(9):1488–93.

6. Peris A, Tutino L, Zagli G, Batacchi S, Cianchi G, Spina R, et al. The use of point-of-care bedside lung ultrasound significantly reduces the number of radiographs and computed tomography scans in critically ill patients. Anesth Analg. 2010;111(3):687–92.

7. Zieleskiewicz L, Cornesse A, Hammad E, Haddam M, Brun C, Vigne C, et al. Implementation of lung ultrasound in polyvalent intensive care unit: impact on irradiation and medical cost. Anaesth Crit Care Pain Med. 2015;34(1):41–4.

8. Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, et al. International evidence-based recommendations for point-of-care lung ultrasound. Intensive Care Med. 2012;38(4):577–91.

9. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA. 1993;270(24):2957–63.

10. Flack VF, Afifi AA, Lachenbruch PA, Schouten HJA. Sample size determinations for the two rater kappa statistic. Psychometrika. 1988;53(3):321–5.

11. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20(1):37–46.

12. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378–82.

13. Landis JR, Koch GG. The Measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.

14. Adriaanse BME, van Vugt JMG, Haak MC. Three- and four-dimensional ultrasound in fetal echocardiography: an up-to-date overview. J Perinatol. 2016;36:685–93.

15. Lang RM, Tsang W, Weinert L, Mor-Avi V, Chandra S. Valvular heart disease. The value of 3-dimensional echocardiography. J Am Coll Cardiol. 2011;58(19):1933–44.

16. Maréchaux S, Le Goffic C, Ennezat P-V, Semichon M, Castel A-L, Delelis F, et al. Quantitative assessment of primary mitral regurgitation using left ventricular volumes: a three-dimensional transthoracic echocardiographic pilot study. Eur Heart J Cardiovasc Imaging. 2014;15(10):1133–9.

17. Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, et al. International evidence based recommendations for point-of-care lung ultrasound. Intensive Care Med. 2012;38:577–91.