Lung ultrasound in diagnosing pneumonia in the emergency department: a systematic review and meta-analysis

Daniele Orso, Nicola Guglielmo and Roberto Copetti

Community-acquired pneumonia (CAP) is one of the most widespread and severe infectious diseases worldwide. In the emergency department (ED), there is still a need for a rapid and accurate tool that can diagnose CAP. Lung ultrasound (LUS) is a recent tool that is increasingly being for this purpose. So far, the LUS has been evaluated on a wide range of patients, but not yet on the specific population in the ED through a meta-analysis. Our aim was to assess the accuracy of the LUS in diagnosing CAP in this setting through a systematic review and a meta-analysis. A systematic research of literature was carried out for all published studies comparing the diagnostic accuracy of the LUS against chest radiography or computerized tomography scan in patients older than 18 years of age with clinical criteria for CAP assessed in the ED. We extracted the descriptive and quantitative data from eligible studies, and calculated the pooled sensitivity, specificity, and diagnostic odds ratio. We defined the summary receiver operating characteristic curve. Our initial search strategy yielded 10 377 studies, of which 17 (0.2%) were eligible. These studies provided a combined sample size of 5108 participants. The general risk of bias of the considered studies was quite low, but some concerns were highlighted. The diagnostic odds ratio was around 181 (I^2: 27%). The pooled area under the curve, sensitivity, and specificity were, respectively, 97, 92, and 93%. The LUS was found to be an accurate tool in diagnosing CAP in adult patients in the ED. More methodologically rigorous trials are needed.

***European Journal of Emergency Medicine* 25:312–321**

European Journal of Emergency Medicine 2018, **25**:312–321

Keywords: emergency department, meta-analysis, pneumonia, ultrasound

Department of Emergency Medicine, Ospedale Civile di Latisana (UD), Udine, Italy

Correspondence to Daniele Orso, MD, Department of Emergency Medicine, Ospedale Civile di Latisana (UD), A.A.S 2 'Bassa Friulana – Isontina', via Sabbionera 45, Latisana, Udine 33053, Italy
Tel: +39 329 016 2763; fax: +39 043 152 9461; e-mail: sd7782.do@gmail.com

Received 28 March 2017 **Accepted** 28 October 2017

Introduction

The annual prevalence of community-acquired pneumonia (CAP) in developed countries ranges from 1.6 to 16 cases per 1000 [1]. Approximately 20% of these will require hospitalization and the mortality rate can reach up to 48% [2]. The American College of Emergency Physicians defines CAP as 'the seventh leading cause of death in the USA, with 1.7 million hospital admissions per year' and the annual economic costs of CAP-related hospitalizations have been estimated at $9 billion [3]. Furthermore, the diagnosis of CAP is still a challenge for the emergency physician as it may present with a broad spectrum of symptoms. Therefore, there is an objective need for accurate imaging methods that enable the diagnosis of CAP in the emergency department (ED). In recent years, ultrasound techniques have advanced considerably. Several clinical trials and meta-analyses had been carried out to assess the accuracy of the lung ultrasound (LUS) in diagnosing pneumonia. The LUS has some advantages over the chest X-ray (CXR) and the chest computed tomography (CCT) scan. In fact, it allows the patient not to be exposed to ionizing radiation, is easily repeatable, and enables therapeutic decisions at bedside. Some studies have shown that the diagnostic accuracy of the CXR is suboptimal. The sensitivity of this method, according to the considered studies, varies from 73 to 88% [4–6]. The patient in the ED is often poorly mobilizable and the execution of the CXR is not perfectly adequate. In contrast, the CCT scan has a very high accuracy. However, it requires the patient to be moved from the ED to the radiology department, with a burden of resources and a potential reduction in patient safety [7,8]. Some statement papers have established some evidence-based recommendations in diagnosing pneumonia by the LUS. In particular, a subpleural echo-poor area (especially if there are fluid bronchograms) or a focal interstitial pattern (as focal B lines) are the most widely used patterns for diagnostic purpose in the literature.

The meta-analyses in the literature establishing the accuracy of the LUS have taken into account different clinical settings all together. We believe that it is not yet established whether the reliability of the LUS specifically in the ED is comparable to other settings such as the radiology department or the ICU through a meta-analysis. Our aim is to verify the diagnostic accuracy of the LUS in diagnosing CAP in adults in the ED through a systematic review and a meta-analysis that assembles several studies published in the literature.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, (*www.euro-emergencymed.com*).

DOI: 10.1097/MEJ.0000000000000517

Methods

Information sources

All prospective studies carried out in ED, which tested the LUS in the diagnosis of CAP in adults until 10 March 2017, were considered. A systematic literature search was carried out in Medline (1946–present). The search was also implemented in Embase (1974–present), the Cochrane Library (1898–present), and Google Scholar. A combination of a controlled vocabulary of keywords around 'pneumonia' and 'ultrasound' was used (more details in Supplementary Material, Supplemental digital content 1, *http://links.lww.com/EJEM/A188*). The search of studies was not limited on the basis of publication dates. Unpublished researches or conference communications were not considered. Only articles written in English were considered. All relevant titles and abstracts were searched for full text by two authors (D.O. and N.G.) independent of each other. References from selected studies and review were evaluated manually to identify any further relevant study for analysis [9]. The literature search and data analysis was carried out in March 2017.

Study selection and eligibility criteria

Inclusion criteria were as follows: enrollment of adult patients aged older than 18 years or more with clinical suspicion of pneumonia on the basis of respiratory signs and symptoms (in accordance with the guidelines of the main scientific societies) or acute respiratory failure, and evaluation of pneumonia on the basis of a combination of clinical data, laboratory results, and any of the following methods: chest imaging by CXR or CCT scan or both. We considered as the gold standard the CCT scan findings of pneumonia in a compatible clinical evaluation, but we also considered the studies that used CXR (see in the Study Selection section). We included studies that considered the ED as a clinical setting.

We excluded studies that enrolled children or pediatric patients. Two authors (D.O. and N.G.) carried out two independent searches for the eligible studies for the pooled analysis. Data retrieved from these studies by both researchers were compared. The decision on the eligibility of the studies was decided by agreement between the two researchers. We established *a priori* that all disputes were resolved through the consultation of the third investigator (R.C.).

Data items

The following data were extracted from each study: first author; year of publication; country where the study was carried out; data collection time (e.g. retrospective or prospective); sample size; mean age of the population; sex proportion; inclusion criteria; method of sampling (randomized, consecutive enrollment, or convenience sample); number of arms in the study (e.g. one or two arms as case–control studies); method of blinding; the involved centers (e.g. monocentric or multicentric study); sonographer qualification; expertise of the sonographer (by number of LUS procedures performed or by the time of LUS experience); LUS pattern definitions and especially the ultrasound diagnostic criteria considered in each study; evaluated areas, in particular, whether the posterior fields were evaluated; considered reference diagnostic standard; and number of proportion of true positives, true negatives, false positives, and false negatives.

Risk of bias in individual studies

Methodological quality was assessed through the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) standard [10]. Two reviewers (D.O. and N.G.) scored the QUADAS-2 checklist independent of each other. The qualitative evaluations of the two researchers were then compared. Disagreements were resolved by consulting the third reviewer (R.C.).

Summary measures and synthesis of results

We estimated the descriptive statistics, in particular, the sensitivity and specificity, the rate of false positives, the positive, and the negative likelihood ratios. We carried out the χ^2-test for heterogeneity in the sensitivities and specificities for the considered studies. We used the univariate approach to assess the cumulative diagnostic odds ratio (DOR), through a fixed-effect model following the Mantel–Haenszel approach, and through the random-effects model following the DerSimonian and Laird approach. We also evaluated the degree of heterogeneity using the Cochran Q-statistic and the inconsistency (I^2) test. We considered an I^2 less than 30% as low, moderate if between 30 and 60%, substantial if between 50 and 90%, and considerable if more than 75%. We also explored the proportional hazards model approach to evaluate the fitting under the homogeneity or heterogeneity hypothesis. Finally, we used the bivariate approach by Reitsma to construct the summary receiver operating characteristic (SROC) and calculate the corresponding summary area under the curve (AUC) [11].

To avoid any bias in the pooled analysis, we considered a subanalysis excluding any study with a high probability of bias in one of the QUADAS-2 items.

All statistical analyses were carried out using the R-CRAN project, ver. 3.3.1 [12]. It was implemented the R-package 'mada'.

Results

Studies' selection

We identified 10 377 studies that fitted our search strategy. Eighty (0.8%) studies were considered for further evaluation on the basis of the inclusion criteria. After excluding studies involved pediatric patients, case-reports, commentaries and narrative or systematic review, and meta-analysis, studies carried out in different settings from the ED, studies written in languages other than English, and

abstracts only, 22 studies were identified. Of these studies, one was excluded on the basis of assessment of the quality [13], two studies because they did not fulfill the inclusion criteria. [14,15], one was excluded because it was considered a very low-quality reference standard [16], and one because a subsample of a previous study was used in it [17]. Two studies were included after consulting the third reviewer [18,19]. Although there were some concerns in terms of the choice of reference standard in one of these two studies, we decided to include it in the pooled analysis as there is no certainty that these concerns would determine the complete inapplicability of the results [18]. For the second one, there were some concerns about the setting [19]. We decided to include it because at least a proportion of the sample was enrolled in the ED. One study evaluated every hemithorax as single independent observations [20]. We considered an effective sample size on the basis of the number of observations rather than the number of patients. A flow diagram of the trial selection process is shown in Fig. 1.

Studies characteristics

The main characteristics of eligible studies are shown in Table 1. The pooled sample size was 5108 patients. The mean age was 67 years and 2529 (48%) patients were men. Only one study was carried out in multiple centers [19]. Eight (47%) studies were carried out in Italy [18,20,21,23,25,26,32,33]; the remaining studies were carried out in USA, France, Iran, Turkey, Denmark, and China. Ten (59%) studies enrolled patients with suspected CAP and six (35%) enrolled patients with acute dyspnea. Seven (41%) studies enrolled a consecutive sample, two studies enrolled a convenience sample, and only one enrolled a randomized sample. In seven (41%) studies, the enrollment method was not specified. Only one study included two study arms (case–control design) [30]. We considered the blinding 'strict' if the radiologist who performed the CXR or the CCT scan was blinded to the ultrasound findings; if not, we considered the blinding just as 'present'. In case there was no blinding of the sonographer to clinical data (or, even more, to the radiological findings), we considered 'no blinding' at all. We judged seven (41%) studies to have a 'strict' blinding; two (12%) studies as blinding 'present', and in eight (47%) studies there was no blinding at all. In 15 (88%) studies, the sonographer was an emergency physician; in one of these, however, the sonographer could also be a radiologist [18]. One study refers only to the experience of the sonographer [19]. A study did not comment on the operator's qualification [20]. The minimum experience level considered necessary varies markedly between studies, ranging from 6 h to 10 years. Similarly, the minimum number of LUS procedures ranged from 50 to more than 400. Ten (59%) studies considered the LUS as positive if there was a subpleural consolidation or a focal B lines area. In one study, it was considered only the consolidation [23] and, in two, they were considered only the focal B lines areas [18,24]. In four studies the positive criteria were not specified. In 13 (76%) studies, the posterior areas of the lungs were explored and in two (12%) studies they were not. In two (12%) studies, this detail was not specified. Seven (41%) studies used as the reference standard the final diagnosis established by one or more independent adjudicators, considering all the observed instrumental and laboratory findings. Six (35%) studies used a CCT scan just in case of discrepancy between the LUS and the CXR. Three (18%) studies considered as the reference standard exclusively the CCT scan. In one study, the reference standard was only the CXR [18].

Risk of bias within studies

The overall quality of studies included in our meta-analysis was high enough (Fig. 2). In three studies [18,23,29], the patient selection criteria were not clear enough (risk of bias: uncertain). However, we judged that there was no significant concern in the applicability of the obtained results. In two studies [18,27], there were some concerns about the choice of the index test (e.g. focal B lines but not consolidation, or not specified at all); for this reason, we judged as unclear the applicability with respect to this item. In any case, we did not have a strong evidence of a high bias source. In three studies [22,24,34], we found some concerns in terms of the reference standard used. Although this finding makes questionable the applicability of the standard used, we did not find this to be a certain source of bias. Only one study was judged to be at high risk of bias in terms of the reference standard: Volpicelli *et al.* [18] chose as the reference standard the CXR instead of the CCT scan. We are not sure that this was the correct choice. In fact, the limits of the diagnostic accuracy of the CXR were highlighted after the publication of that work. For this reason, although we did not have sufficient information to be able to exclude the study of Volpicelli *et al.* [18] from the pooled analysis, we carried out a subanalysis excluding this study. In two studies [19,20], we detected minor concerns in terms of the timing at which the ultrasound examination was carried out. In particular, in both studies, the time limit within which the enrolled patients were subjected to the LUS is not well specified.

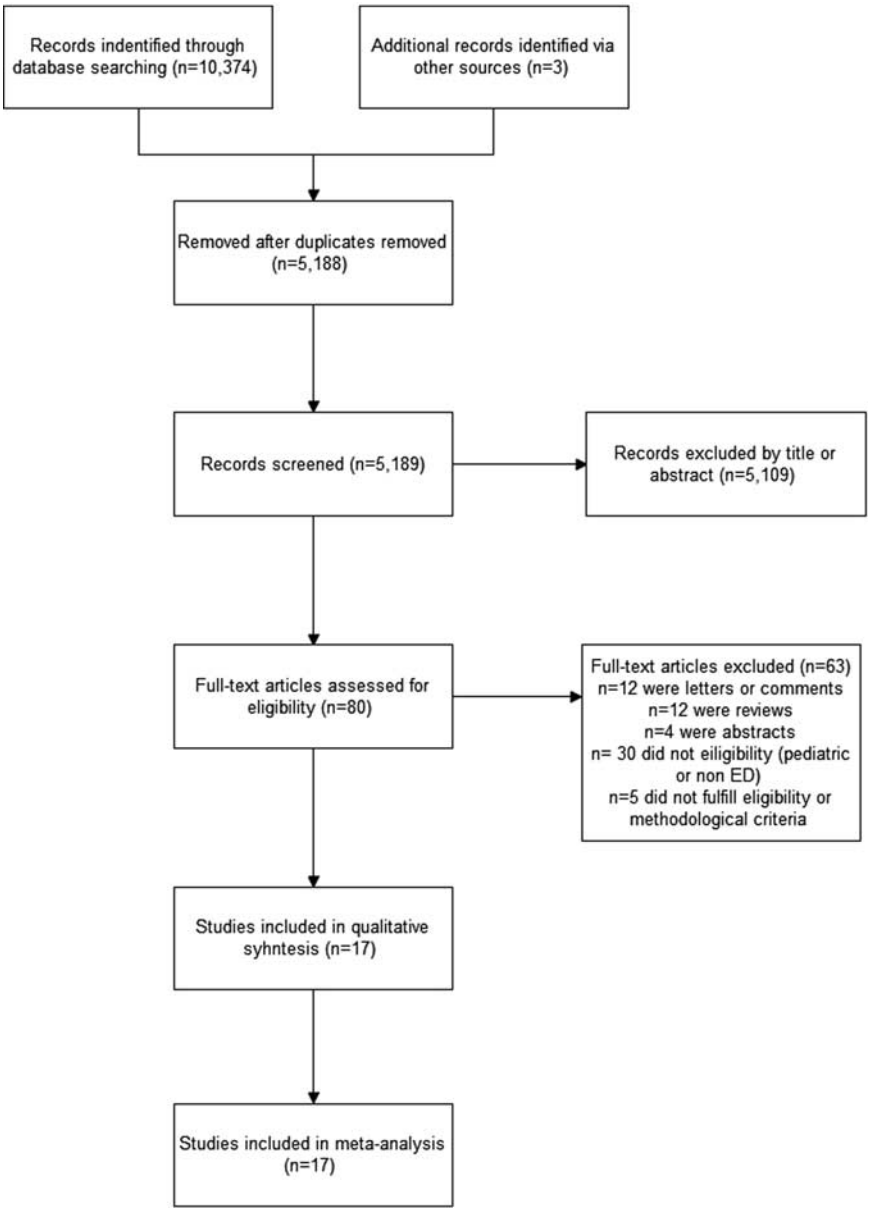
Results of individual studies

The sensitivity and the specificity of the considered studies are shown in Table 2. The sensitivity ranged from 0.68 [95% confidence interval (CI): 0.53–0.80] to 1.00 (95% CI: 1.00–1.00). The specificity ranged from 0.25 (95% CI: 0.03–0.80) to 1.00 (95% CI: 0.98–1.00). The variability among studies for both the sensitivity and the specificity was significant (χ^2 values, respectively, 264.73 and 463.10; $P < 0.001$ for both). The forest plots for the sensitivity and the specificity are presented in Fig. 3.

Synthesis of results

The positive likelihood ratio among the considered studies ranged from 512.97 (95% CI: 32.11–8196.15) to 1.31

Fig. 1



Flow diagram of the articles retrieved from the search of databases and the reasons for exclusions. ED, emergency department.

(95% CI: 0.59–2.92). The negative likelihood ratio among the studies ranged from 0.00 (95% CI: 0.00–0.01) to 0.38 (95% CI: 0.22–0.64). The DOR through the Mantel–Haenszel method for the fixed effects was 51.35 (Cochran Q-statistic: 39296.42) and the DOR through the DerSimonian and Laird method for random effects was 180.58 (95% CI: 65.84–495.31; logDOR: 5.20; 95% CI: 4.19–6.21; τ: 1.83; 95% CI: 0.00–3.28; Cochran Q-statistic: 21.98 with $P = 0.14$; Higgin's I^2: 27.21%). The forest plot of DOR through the random-effects model is shown in Fig. 4a.

Under the homogeneity hypothesis, we obtained a log-likelihood of − 97.18 [Akaike information criterion (AIC): 196.40; Bayesian information criterion (BIC): 197.20] and an AUC of 0.98 (0.98–0.99) with a χ^2 goodness-of-fit test of 302.77 ($P < 0.001$). Under heterogeneity, we achieved a log-likelihood of 44.78 (AIC: − 85.60; BIC: − 83.90) and an AUC of 0.97 (95% CI: 0.96–0.99) with a χ^2 goodness-of-fit test of 27.75 ($P = 0.023$).

The corresponding summary AUC through a bivariate approach was 0.97 (log-likelihood: 40.22; AIC: − 70.44; BIC: − 62.81), with a pooled sensitivity of 0.92 (95% CI: 0.86–0.95) and a specificity of 0.93 (95% CI: 0.86–0.97). The SROC is shown in Fig. 4b.

Table 1 Characteristics of the pooled studies considered

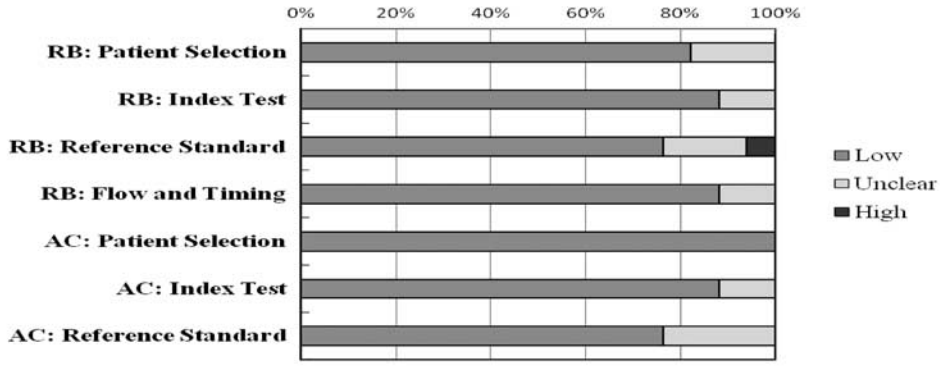
ID	References	Country	Setting	Type	Sample size	Mean age (years)	Male/female ratio
1	Zanobetti *et al.* [21]	Italy	ED	Prospective	2683	71	1367/1316
2	Dimitrios *et al.* [22]	USA	ED	Prospective	115	61	47/68
3	Interrigi *et al.* [23]	Italy	ED	Prospective	370	NS	NS
4	Mantuani *et al.* [24]	USA	ED	Prospective	57	58	36/21
5	Pagano *et al.* [25]	Italy	ED	Prospective	105	58	59/46
6	Corradi *et al.* [20]	Italy	ED	Prospective	32 (64)[a]	62	17/15
7	Nazerian *et al.* [26]	Italy	ED	Prospective	285	71	133/152
8	Taghizadieh *et al.* [27]	Iran	ED	Prospective	30	NS	NS
9	Gallard *et al.* [34]	France	ED	Prospective	130	80	70/60
10	Liu *et al.* [28]	China	ED	Prospective	179	72	100/79
11	Bourcier *et al.* [29]	France	ED	Prospective	144	78	72/72
12	Laursen *et al.* [30]	Denmark	ED	Prospective	158	73	61/97
13	Ünlüer *et al.* [31]	Turkey	ED	Prospective	72	66	35/37
14	Cortellaro *et al.* [32]	Italy	ED	Prospective	120	69	77/43
15	Reissig *et al.* [19]	Europe	NS	Prospective	362	64	228/134
16	Volpicelli *et al.* [18]	Italy	ED	Prospective	217	67	132/85
17	Parlamento *et al.* [33]	Italy	ED	Prospective	49	61	31/18

ID	Inclusion criteria	Sampling	Arms	Blinding	Centre	LUS operator (experience)	Diagnostic criteria (posterior areas: yes/no)	Control
1	Acute dyspnea	Consecutive	Single	Strict	Single	EP (>80 h/150 LUS)	Consolidation + focal B lines (yes)	Final diagnosis
2	Acute dyspnea	NS	Single	No	Single	EP (NS)	NS (yes)	Final diagnosis
3	Acute dyspnea	NS	Single	No	Single	EP (NS)	Consolidation (yes)	CXR/CCT
4	Acute dyspnea	Convenience	Single	Present	Single	EP (NS)	B lines (no)	Final diagnosis
5	Suspected CAP	NS	Single	Present	Single	EP (NS)	Consolidation + focal B lines (yes)	Final diagnosis
6	Suspected CAP	Consecutive	Single	No	Single	NS	Consolidation + focal B lines (yes)	CCT
7	Suspected CAP	Consecutive	Single	Strict	Single	EP (>1 year)	Consolidation + focal B lines (yes)	CCT
8	Suspected CAP	Consecutive	Single	No	Single	EP (NS)	NS (NS)	CXR/CCT
9	Acute dyspnea	NS	Single	Strict	Single	EP (>2 years)	Consolidation + focal B lines (yes)	Final diagnosis
10	Suspected CAP	Consecutive	Single	No	Single	EP (28 h/50 LUS)	Consolidation + focal B lines (yes)	CCT
11	Suspected CAP	NS	Single	No	Single	EP (24 h)	Consolidation + focal B lines (yes)	Final diagnosis
12	Acute dyspnea	Randomized	Two	Strict	Single	EP (>400 LUS)	NS (NS)	Final diagnosis
13	Suspected CAP	NS	Single	No	Single	EP (6 h)	Consolidation + focal B lines (yes)	CXR/CCT
14	Suspected CAP	Consecutive	Single	Strict	Single	EP (NS)	Consolidation + focal B lines (yes)	CXR/CCT
15	Suspected CAP	NS	Single	No	Multicentric	Experienced physicians (>100 LUS)	NS (yes)	CXR/CCT
16	NS	Consecutive	Single	Strict	Single	EP/radiologists (>200 LUS/year)	Focal B lines (no)	CXR
17	Suspected CAP	Convenience	Single	Strict	Single	EP (>10 years)	Consolidation + focal B lines (yes)	CXR/CCT

CCT, chest computed tomography; CXR, chest X-ray; ED, emergency department; EP, emergency physician; LUS, lung ultrasound; NS, not specified.
[a]The number of patients was 32, but the authors considered every hemithorax as independent; thus, the effective sample size was 64.

Fig. 2

	Risk of bias				Applicability Concerns		
Study	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
1	☺	☺	☺	☺	☺	☺	☺
2	☺	☺	😐	☺	☺	☺	😐
3	😐	☺	☺	☺	☺	☺	☺
4	☺	☺	😐	☺	☺	☺	😐
5	☺	☺	☺	☺	☺	☺	☺
6	☺	☺	☺	😐	☺	☺	☺
7	☺	☺	☺	☺	☺	☺	☺
8	☺	😐	☺	☺	☺	😐	☺
9	☺	☺	😐	☺	☺	☺	😐
10	☺	☺	☺	☺	☺	☺	☺
11	😐	☺	☺	☺	☺	☺	☺
12	☺	☺	☺	☺	☺	☺	☺
13	☺	☺	☺	☺	☺	☺	☺
14	☺	☺	☺	☺	☺	☺	☺
15	☺	☺	☺	😐	☺	☺	☺
16	😐	😐	☹	☺	☺	😐	😐
17	☺	☺	☺	☺	☺	☺	☺



Details of quality assessment by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.

Additional analysis

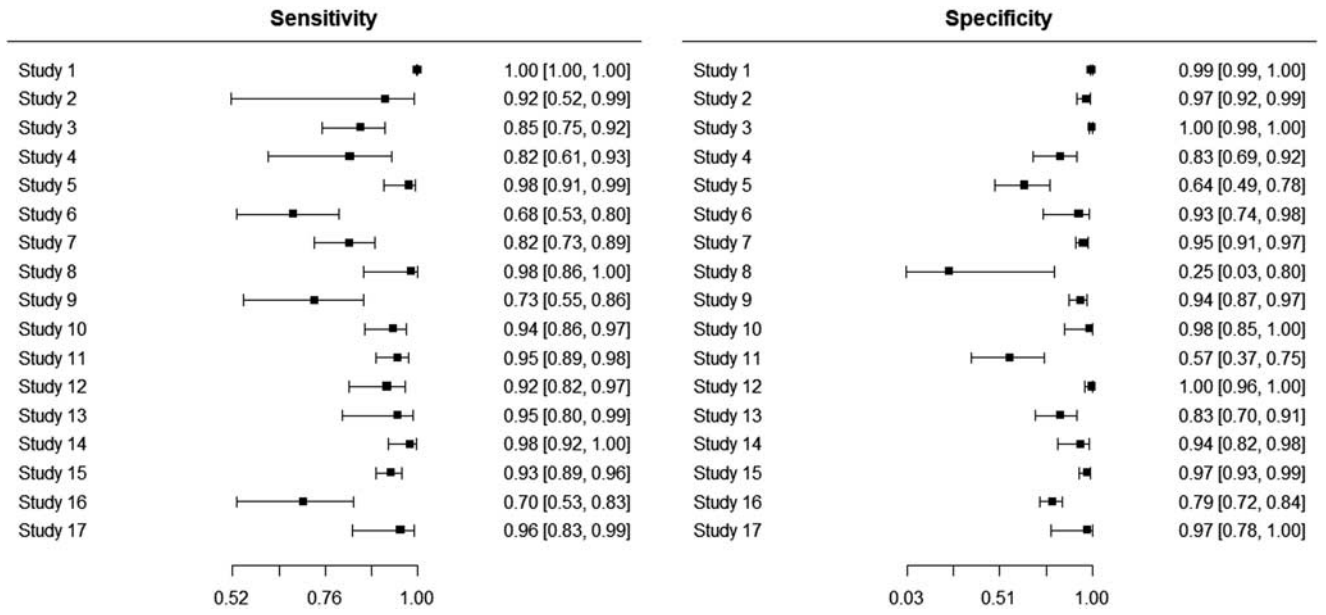
To avoid any source of bias, we carried out a subanalysis excluding any study with a high risk of bias in any QUADAS-2 item. For this reason, we recalculated the summary AUC by excluding the study by Volpicelli *et al.* [18]. The corresponding summary AUC was 0.97

Table 2 Descriptive statistics of the considered studies

ID	TP	FP	TN	FN	Sensitivity	Specificity	Positive LR	Negative LR	DOR
1	1086	10	1587	0	1.00 (1.00–1.00)	0.99 (0.99–1.00)	152.12 (83.25–277.98)	0.00 (0.00–0.01)	328536.90 (19231.78–5612402.42)
2	5	3	107	0	0.92 (0.52–0.99)	0.97 (0.92–0.99)	29.07 (10.08–83.81)	0.09 (0.01–1.22)	337.86 (15.46–7385.01)
3	60	0	300	10	0.85 (0.75–0.92)	1.00 (0.98–1.00)	512.97 (32.11–8196.15)	0.15 (0.08–0.26)	3462.90 (200.22–59892.72)
4	16	6	32	3	0.82 (0.61–0.93)	0.83 (0.69–0.92)	4.95 (2.38–10.27)	0.21 (0.08–0.55)	23.57 (5.65–98.31)
5	67	13	24	1	0.98 (0.91–0.99)	0.64 (0.49–0.78)	2.75 (1.79–4.23)	0.03 (0.01–0.17)	81.67 (14.20–469.52)
6	30	1	19	14	0.68 (0.53–0.80)	0.93 (0.74–0.98)	9.49 (2.00–44.94)	0.35 (0.22–0.54)	27.34 (4.64–161.26)
7	72	9	189	15	0.82 (0.73–0.89)	0.95 (0.91–0.97)	17.26 (9.21–32.34)	0.18 (0.12–0.29)	93.30 (39.81–218.67)
8	29	1	0	0	0.98 (0.86–1.00)	0.25 (0.03–0.80)	1.31 (0.59–2.92)	0.07 (0.00–2.56)	19.67 (0.28–1377.85)
9	20	6	97	7	0.73 (0.55–0.86)	0.94 (0.87–0.97)	11.71 (5.38–25.49)	0.29 (0.15–0.53)	41.00 (12.94–129.91)
10	80	0	27	5	0.94 (0.86–0.97)	0.98 (0.85–1.00)	52.42 (3.36–817.92)	0.07 (0.03–0.15)	805.00 (43.10–15034.29)
11	117	9	12	6	0.95 (0.89–0.98)	0.57 (0.37–0.75)	2.19 (1.36–3.55)	0.09 (0.04–0.21)	23.79 (7.49–75.54)
12	52	0	102	4	0.92 (0.82–0.97)	1.00 (0.96–1.00)	189.74 (11.94–3016.20)	0.08 (0.03–0.19)	2391.67 (126.36–45268.28)
13	27	7	37	1	0.95 (0.80–0.99)	0.83 (0.70–0.91)	5.69 (2.94–11.00)	0.06 (0.01–0.30)	91.67 (14.84–566.20)
14	80	2	37	1	0.98 (0.92–1.00)	0.94 (0.82–0.98)	15.71 (4.73–52.18)	0.02 (0.00–0.10)	805.00 (102.50–6322.32)
15	211	3	127	15	0.93 (0.89–0.96)	0.97 (0.93–0.99)	34.87 (12.40–98.09)	0.07 (0.04–0.11)	497.07 (152.66–1618.55)
16	22	39	147	9	0.70 (0.53–0.83)	0.79 (0.72–0.84)	3.33 (2.33–4.76)	0.38 (0.22–0.64)	8.84 (3.83–20.40)
17	32	0	17	1	0.96 (0.83–0.99)	0.97 (0.78–1.00)	34.41 (2.24–529.74)	0.05 (0.01–0.22)	758.33 (29.32–19615.20)

DOR, diagnostic odds ratio; FN, false negative; FP, false positive; LR, likelihood ratio; TN, true negative; TP, true positive.

Fig. 3



Forest plots of pooled sensitivity and specificity.

(log-likelihood: 40.61; AIC: – 71.22; BIC: – 63.89), with a pooled sensitivity of 0.92 (95% CI: 0.87–0.96) and a specificity of 0.94 (95% CI: 0.87–0.97).

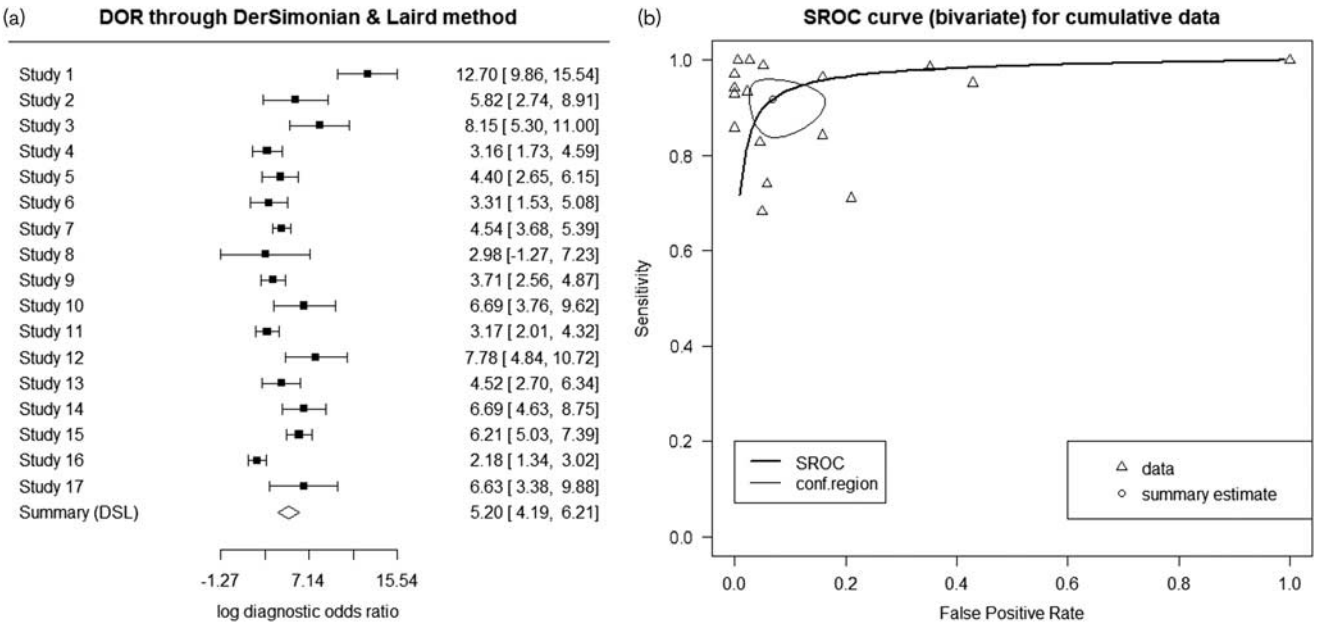
Discussion

We found that the LUS had a quite high sensitivity (92%) and specificity (93%) in diagnosing CAP in adults in the ED setting. Excluding the study by Volpicelli *et al.* [18], which had used a questionable reference standard as the CXR, we did not find an increase in accuracy of the LUS, but the overall diagnostic accuracy was already high enough (97%). The large pooled sample yields a stable result even without considering studies with a small sample size such as that by Volpicelli *et al.* [18].

We found that there is still some uncertainty about the best reference standard. In fact, if the almost excellent diagnostic capabilities of the CCT scan seem obvious, in contrast, it does not seem ethically justifiable to indiscriminately expose every patient to ionizing radiation.

It appears that studies in the literature are gradually increasing in number, and then as the pooled sample increases, the degree of heterogeneity is reducing. The study of Zanobetti *et al.* [21], which enrolled more than

Fig. 4



(a) Forest plot of pooled diagnostic odds ratio (DOR); (b) the summary receiver operating characteristic (SROC) of the lung ultrasound in diagnosing community-acquired pneumonia in adult patients in emergency department (ED). DSL, DerSimonian and Laird method.

two thousand patients, obtained some very substantial results in this respect. However, reaching a sufficiently large sample to achieve a good level of power does not seem to be sufficiently taken into account in current clinical trials.

In terms of the strictness of blinding, the studies in the literature had different strategies. We believe that more attention should be paid to this aspect in further studies, although we did not find substantial sources of bias from the considered studies.

To our knowledge, this is the first meta-analysis that takes into account specifically the ED setting. In fact, we believe that it is methodologically questionable to include studies carried out in the ICU or in the medical ward to establish the diagnostic accuracy of the LUS in diagnosing CAP. A proportion of pneumonia that occurs in hospitalized patients could be attributed to hospital-acquired pneumonia or induced ventilator-associated pneumonia in patients intubated for more than 24 h. Although, theoretically at least, it is conceivable that there are no considerable differences in the diagnostic accuracy of the LUS in these specific target populations compared to CAP patients, there are currently no primary studies demonstrating it.

In terms of the sonographic findings for pneumonia, some evidence-based recommendations produced by a panel of experts are present in the literature [35]. These recommendations are sufficiently robust, in our opinion, to be able to be considered in the use of definitions related to the LUS.

Compared with the meta-analysis already reported in the literature, our population sample is more focused on our clinical question [36–42]. Despite this, we obtained a very similar value of AUC in comparison with most of the other meta-analyses (98% in Alzahrani *et al.* [36]; 96% in Long *et al.* [37]; 96% in Xia *et al.* [38]; 97% in Ye *et al.* [39]; 98% in Chavez *et al.* [40]). These data testify to the usefulness of the LUS in diagnosing CAP, despite heterogeneous populations. Hu *et al.* [41] obtained a higher AUC (of about 99%) than ours. However, we noticed that they considered numerous studies involving a pediatric (and a newborn) population. We decided, instead, to exclude studies that enrolled pediatric patients. This is the reason for the difference in accuracy between their work and ours. Llamas-Álvarez *et al.* [42] obtained a lower AUC value (93%) than ours. We assume that this difference is mainly because of the different studies considered. Llamas-Álvarez and colleagues did not consider the study by Zanobetti and colleagues, which included more than 2000 patients. Furthermore, the considered setting by Llamas-Álvarez and colleagues was quite different from ours; in fact, they considered studies carried out in the ED, in the ICU, and also in the medical ward. We cannot report the degree of heterogeneity as Llamas-Álvarez and colleagues did not report it. However, they recognized a certain degree of heterogeneity that we believe can be attributed to the difference in the populations considered (e.g. ventilator-associated

pneumonias and CAPs together). The heterogeneity of the studies considered by us was low ($I^2 = 27.21\%$). Although was much smaller than all the previous meta-analyses, it was still not optimal. These data reflect the methodological diversity of the analyzed studies, as described previously (target population, enrollment, blinding, sonographer's expertise, and reference standard). We believe that some more methodologically rigorous studies would be useful to standardize as much as possible the different populations considered and the diagnostic references.

Conclusion

The LUS proved to be a sufficiently useful and accurate tool to diagnose CAP in an adult population in the ED.

Some concerns have been raised on the robustness of the obtained results, because of the lack of a well-standardized methodology in the studies included (in particular on the choice of reference standards, the experience of the sonographer, and the ultrasound patterns considered significant for CAP). More methodologically rigorous studies are needed.

Acknowledgements

Conflicts of interest

There are no conflicts of interest.

References

1 Llop CJ, Tuttle E, Tillotson GS, LaPlante K, File TM. Antibiotic treatment patterns, costs, and resource utilization among patients with community acquired pneumonia: a US cohort study. *Hosp Pract* 2017; **45**:1–8.

2 Welte T, Torres A, Nathwani D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. *Thorax* 2012; **67**:71–79.

3 Nazarian DJ, Eddy OL, Lukens TW, Weingart SD, Decker WW. Clinical policy: critical issues in the management of adult patients presenting to the emergency department with community-acquired pneumonia. *Ann Emerg Med* 2009; **54**:704–731.

4 Hagaman JT, Rouan GW, Shipley RT, Panos RJ. Admission chest radiograph lacks sensitivity in the diagnosis of community-acquired pneumonia. *Am J Med Sci* 2009; **337**:236–240.

5 Maughan BC, Asselin N, Carey JL, Sucov A, Valente JH. False-negative chest radiographs in emergency department diagnosis of pneumonia. *R I Med J* 2014; **97**:20–23.

6 Hayden GE, Wrenn KW. Chest radiograph vs. computed tomography scan in the evaluation of pneumonia. *J Emerg Med* 2009; **36**:266–270.

7 Brenner DJ, Hall EJ. Computed tomography – an increasing source of radiation exposure. *N Engl J Med* 2007; **357**:2277–2284.

8 Trovato FM, Catalano D, Trovato GM. Thoracic ultrasound: an adjunctive and valuable imaging tool in emergency, resource-limited settings and for a sustainable monitoring of patients. *World J Radiol* 2016; **8**:775–784.

9 Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, *et al*. Meta-analysis of observational studies in epidemiology (MOOSE) group. Meta-analysis of observational studies in epidemiology. *JAMA* 2000; **283**:2008–2012.

10 Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al*. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**:529–536.

11 Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; **149**:889–897.

12 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available at: *https://cran.r-project.org/index.html*. [Accessed 4 March 2017].

13 Sperandeo M, Carnevale V, Muscarella S, Sperandeo G, Varriale A, Filabozzi P, *et al*. Clinical application of transthoracic ultrasonography in inpatients with pneumonia. *Eur J Clin Invest* 2011; **41**:1–7.

14 Volpicelli G, Caramello V, Cardinale L, Cravino M. Diagnosis of radio-occult pulmonary conditions by real-time chest ultrasonography in patients with pleuritic pain. *Ultrasound Med Biol* 2008; **34**:1717–1723.

15 Dexheimer Neto FL, Andrade JM, Raupp ACT, Townsend DS, Beltrami FG, Brisson H, *et al*. Diagnostic accuracy of the bedside lung ultrasound in emergency protocol for the diagnosis of acute respiratory failure in spontaneously breathing patients. *J Bras Pneumol* 2015; **41**:58–64.

16 Özkan B, Ünlüer EE, Akyol EE, Karagöz A, Bayata MS, Akoğlu H, *et al*. Stethoscope versus point-of-care ultrasound in the differential diagnosis of dyspnea: a randomized trial. *Eur J Emerg Med* 2015; **22**:440–443.

17 Nazerian P, Cerini G, Vanni S, Gigli C, Zanobetti M, Bartolucci M, *et al*. Diagnostic accuracy of lung ultrasonography combined with procalcitonin for the diagnosis of pneumonia: a pilot study. *Crit Ultrasound J* 2016; **8**:17.

18 Volpicelli G, Caramello V, Cardinale L, Mussa A, Bar F, Frascisco MF. Detection of sonographic B-lines in patients with normal lung or radiographic alveolar consolidation. *Med Sci Monit* 2008; **14**:122–128.

19 Reissig A, Copetti R, Mathis G, Mempel C, Schuler A, Zechner P, *et al*. Lung ultrasound in the diagnosis and follow-up of community-acquired pneumonia. *Chest* 2012; **142**:965–972.

20 Corradi F, Brusasco C, Garlaschi A, Paparo F, Ball L, Santori G, *et al*. Quantitative analysis of lung ultrasonography for the detection of community-acquired pneumonia: a pilot study. *Biomed Res Int* 2015; **2015**:868707.

21 Zanobetti M, Scorpiniti M, Gigli C, Nazerian P, Vanni S, Innocenti F, *et al*. Point-of-care ultrasonography for evaluation of acute dyspnea in the emergency department. *Chest* 2017; **151**:1295–1301.

22 Dimitrios P, Secko M, Gullett J, Stone M, Zehtabchi S. Clinician-performed bedside ultrasound in improving diagnostic accuracy in patients presenting to the emergency department with acute dyspnea. *West J Emerg Med* 2017; **18**:382–389.

23 Interrigi MC, Trovato FM, Catalano D, Trovato GM. Emergency thoracic ultrasound and clinical risk management. *Ther Clin Risk Manag* 2017; **13**:151–160.

24 Mantuani D, Frazee BW, Fahimi J, Nagdev A. Point-of-care multi-organ ultrasound improves diagnostic accuracy in adults presenting to the emergency department with acute dyspnea. *West J Emerg Med* 2016; **17**:46–53.

25 Pagano A, Numis FG, Visone G, Pirozzi C, Masarone M, Olibet M, *et al*. Lung ultrasound for diagnosis of pneumonia in emergency department. *Intern Emerg Med* 2015; **10**:851–854.

26 Nazerian P, Volpicelli G, Vanni S, Gigli C, Betti L, Bartolucci M, *et al*. Accuracy of lung ultrasound for the diagnosis of consolidations when compared to chest computed tomography. *Am J Emerg Med* 2015; **33**:620–625.

27 Taghizadieh A, Ala A, Rahmani F, Nadi A. Diagnostic accuracy of chest x-ray and ultrasonography in detection of community acquired pneumonia: a brief report. *Emerg (Tehran)* 2015; **3**:114–116.

28 Liu X, Lian R, Tao Y, Gu C, Zhang G. Lung ultrasonography: an effective way to diagnose community-acquired pneumonia. *Emerg Med J* 2015; **32**:433–438.

29 Bourcier JE, Paquet J, Seinger M, Gallard E, Redonnet JP, Cheddadi F, *et al*. Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. *Am J Emerg Med* 2014; **32**:115–118.

30 Laursen CB, Sloth E, Lassen AT, Christensen RD, Lambrechtsen J, Madsen PH, *et al*. Point-of-care ultrasonography in patients admitted with respiratory symptoms: a single-blind, randomised controlled trial. *Lancet Respir Med* 2014; **2**:638–646.

31 Ünlüer EE, Karagoz A, Senturk GO, Bayata S. Bedside lung ultrasonography for diagnosis of pneumonia. *Hong Kong J Emerg Med* 2013; **20**:98–104.

32 Cortellaro F, Colombo S, Coen D, Duca PG. Lung ultrasound is an accurate diagnostic tool for the diagnosis of pneumonia in the emergency department. *Emerg Med J* 2012; **29**:19–23.

33 Parlamento S, Copetti R, di Bartolomeo S. Evaluation of lung ultrasound for the diagnosis of pneumonia in the ED. *Am J Emerg Med* 2009; **27**:379–384.

34 Gallard E, Redonnet JP, Bourcier JE, Deshaies D, Largeteau N, Amalric JM, *et al*. Diagnostic performance of cardiopulmonary ultrasound performer by the emergency physician in the management of acute dyspnea. *Am J Emerg Med* 2015; **33**:352–358.

35 Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, *et al*. International liaison committee on lung ultrasound (ILC-CUS) for the international consensus conference on lung ultrasound (ICC-LUS). International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med* 2012; **38**:577–591.

36 Alzahrani SA, Al-Salamah MA, Al-Madani WH, Elbarbary MA. Systematic review and meta-analysis for the use of ultrasound versus radiology in diagnosing of pneumonia. *Crit Ultrasound J* 2017; **9**:6.

37 Long L, Zhao HT, Zhang ZY, Wang GY, Zhao HL. Lung ultrasound for the diagnosis of pneumonia in adults: a meta-analysis. *Medicine (Baltimore)* 2017; **96**:e5713.
38 Xia Y, Ying Y, Wang S, Li W, Shen H. Effectiveness of lung ultrasonography for diagnosis of pneumonia in adults: a systematic review and meta-analysis. *J Thorac Dis* 2016; **8**:2822–2831.
39 Ye X, Xiao H, Chen B, Zhang SY. Accuracy of lung ultrasonography versus chest radiography for the diagnosis of adult community-acquired pneumonia: review of the literature and meta-analysis. *PLoS One* 2015; **10**:e0130066.
40 Chavez MA, Shams N, Ellington LE, Naithani N, Gilman RH, Steinhoff MC, *et al*. Lung ultrasound for the diagnosis of pneumonia in adults: a systematic review and meta-analysis. *Respir Res* 2014; **15**:50.
41 Hu QJ, Shen YC, Jia LQ, Guo SJ, Long HY, Pang CS, *et al*. Diagnostic performance of lung ultrasound in the diagnosis of pneumonia: a bivariate meta-analysis. *Int J Clin Exp Med* 2014; **7**:115–121.
42 Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J. Accuracy of lung ultrasound in the diagnosis of pneumonia in adults: systematic review and meta-analysis. *Chest* 2017; **151**:374–382.