REVIEW

Diagnostic accuracy of lung ultrasound in diagnosis of ARDS and identification of focal or non-focal ARDS subphenotypes: a systematic review and meta-analysis

Maud M. A. Boumans[1], William Aerts[2], Luigi Pisani[4,8], Lieuwe D. J. Bos[3,9], Marry R. Smit[2,7] and Pieter R. Tuinman[2,5,6,7*]

Abstract

Background Acute respiratory distress syndrome (ARDS) is a life-threatening respiratory condition with high mortality rates, accounting for 10% of all intensive care unit admissions. Lung ultrasound (LUS) as diagnostic tool for acute respiratory failure has garnered widespread recognition and was recently incorporated into the updated definitions of ARDS. This raised the hypothesis that LUS is a reliable method for diagnosing ARDS.

Objectives We aimed to establish the accuracy of LUS for ARDS diagnosis and classification of focal versus non-focal ARDS subphenotypes.

Methods This systematic review and meta-analysis used a systematic search strategy, which was applied to Pub-Med, EMBASE and cochrane databases. Studies investigating the diagnostic accuracy of LUS compared to thoracic CT or chest radiography (CXR) in ARDS diagnosis or focal versus non-focal subphenotypes in adult patients were included. Quality of studies was evaluated using the QUADAS-2 tool. Statistical analyses were performed using "*Mada*" in Rstudio, version 4.0.3. Sensitivity and specificity with 95% confidence interval of each separate study were summarized in a Forest plot.

Results The search resulted in 2648 unique records. After selection, 11 reports were included, involving 2075 patients and 598 ARDS cases (29%). Nine studies reported on ARDS diagnosis and two reported on focal versus non-focal ARDS subphenotypes classification. Meta-analysis showed a pooled sensitivity of 0.631 (95% CI 0.450–0.782) and pooled specificity of 0.942 (95% CI 0.856–0.978) of LUS for ARDS diagnosis. In two studies, LUS could accurately differentiate between focal versus non-focal ARDS subphenotypes. Insufficient data was available to perform a meta-analysis.

Conclusion This review confirms the hypothesis that LUS is a reliable method for diagnosing ARDS in adult patients. For the classification of focal or non-focal subphenotypes, LUS showed promising results, but more research is needed.

Keywords ARDS, Lung ultrasound, Respiratory medicine, Diagnostic test accuracy

*Correspondence:
Pieter R. Tuinman
p.tuinman@amsterdamumc.nl
Full list of author information is available at the end of the article

Boumans *et al. Critical Care*   (2024) 28:224

Page 2 of 15

Introduction

Acute respiratory distress syndrome (ARDS) is a life-threatening respiratory condition, which manifests as acute hypoxemic respiratory failure often requiring mechanical ventilation [1, 2].

The prevalence of ARDS is high, estimated around 10% of all intensive care unit (ICU) admissions and mortality remains high [3]. Diagnosis is based on a broad set of clinical and radiological criteria lacking specificity. This results in physiological, biological, and radiological heterogeneity. The imaging criterion used in the diagnosis of ARDS is subject to considerable variability between observers and techniques. Accurate diagnosis of ARDS is of importance for the adequate management as well as clear definition of patient populations included in clinical research [3, 4].

Radiological heterogeneity in ARDS has been studied extensively and has revealed that morphological subphenotypes exist, which respond differently to ventilator treatment strategies [5]. Non-focal ARDS is defined by diffuse or patchy loss of aeration, which generally responds better to recruitment maneuvers. Focal ARDS shows predominant dorso-inferior consolidations and generally responds better to prone positioning [4]. In a prospective randomized clinical trial, a personalized treatment strategy was found to be superior only when classification was accurate and treatment was aligned with lung morphology [6], highlighting the importance of correct classification.

Since 2012, ARDS has been diagnosed by the Berlin definition [7]. In the last decade, several developments have prompted expansion of this definition. The Kigali modification is proposed for resource-constrained settings without access to CT, variable access to CXR or invasive pulse-oximetric methods for evaluating oxygenation [8]. The use of High-Flow Nasal Oxygen (HFNO) to manage ARDS patients with severe hypoxemic patients has been investigated and the use of it exacerbated during the COVID-19 pandemic. Finally, the use of LUS is increasing rapidly in patients with acute respiratory failure. To address these changes, a global consensus was reached to update the ARDS definition [9].

This global definition acknowledges the limitations of CT or CXR for ARDS diagnosis and includes LUS as a diagnostic tool. Multiple studies have investigated LUS as a tool for identifying ARDS and differentiation between focal and non-focal subphenotypes. However, the accuracy has not yet been well defined and the approaches of LUS among studies vary. In this systematic review and meta-analysis, we aim to assess the diagnostic accuracy of LUS for diagnosing ARDS and for classification of focal versus non-focal subphenotypes.

Objectives

Primary objective

To evaluate the diagnostic accuracy of LUS for diagnosing ARDS in adult patients.

Secondary objective

To evaluate the accuracy of LUS in classification of focal versus non-focal ARDS subphenotypes in adult patients.

Method

Study design

This is a systematic review and meta-analysis. The protocol was written according to the PRISMA-P guidelines (preferred reporting items for systematic reviews and meta-analysis protocols) and pre-registered at PROSPERO, registration number CRD42023413462.

Inclusion criteria

We aimed to include all studies that investigated the accuracy of LUS for ARDS diagnosis compared to the Berlin definition (requiring CT or CXR) and all studies that investigated the accuracy of LUS for classification of focal versus non-focal subphenotypes compared to CT or CXR. Cohort, case–control, cross-sectional and observational studies were included. Patient population included adult patients presenting with acute respiratory failure and/or need for mechanical ventilation. The index test was LUS, all LUS protocols were included. The reference standard for ARDS diagnosis was the Berlin definition, which used CT or CXR data. The reference standard for focal versus non-focal subphenotypes was CT and/or CXR.

Exclusion criteria

We excluded studies in pediatric patients, case reports and case series, articles published in any other language than English and studies without a reference CT or CXR. The authors made the decision to exclude non-English studies as evidence suggests that the impact on the results of systematic reviews is negligble and translation of non-English scientific papers may lead to errors in interpretation [10–13].

Literature search

A medical librarian experienced in organizing systematic reviews was consulted to construct a robust search strategy. The search was conducted up until the 6th of February 2023, in MEDLINE via PubMed, EMBASE and Cochrane Database of Systematic Reviews. The search

Boumans *et al. Critical Care*   (2024) 28:224

Page 3 of 15

methods and terms used are presented in the Supplementary material.

Selection of studies

Results were managed in an online database, Rayyan. Two independent reviewers (MB and WA) first evaluated title and abstracts. Afterwards full texts were assessed for eligibility. References of included studies were screened, and potentially eligible studies were evaluated for inclusion. Disagreement between the first two reviewers was resolved by consensus meetings with a third reviewer (PRT).

Data extraction

Data was extracted by two independent reviewers (MB and WA). From each included study, characteristics were extracted and summarized. These included author, study design, study period, year of publication, patient characteristics, number of ARDS and non-ARDS patients, chosen LUS protocol and reference standard, primary and secondary outcomes and study limitations.

Outcomes

The primary outcome was the accuracy of LUS in diagnosing ARDS. The secondary objective was to establish the accuracy of LUS in classification of focal versus non-focal ARDS subphenotypes. This included accuracy, sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, accuracy, AUROC and Youden's index. Whenever a specific measure was not described in the study, raw data was used for calculation by the reviewers. Specificity, sensitivity, negative predictive value and positive predictive value were calculated by extracting raw data and implementing it in a 2×2 contingency table.

Quality assessment

For quality assessment of the included studies, the QUADAS-2 tool (Quality Assessment of Diagnostic Accuracy Studies) was used by two independent reviewers (MB and WA). Disagreement was resolved by consensus meetings with a third reviewer (PRT).

Statistical analysis and data synthesis

Pooled sensitivity and specificity estimates were obtained by applying a bivariate analysis on the raw study data, using the package "*Mada*" in Rstudio, version 4.0.3. This resulted in a summary receiver operating characteristic (sROC) curve. The sensitivity and specificity with 95% confidence interval of each separate study were summarized in a Forest plot.

For the studies that evaluated different LUS protocols or scores within the same patient cohort, the protocol that was recommended by the authors was analyzed. If no recommendation was made, the protocol with the highest accuracy was included. In the diagnostic accuracy analysis of the study of Smit et al., two different cutoff points for the LUS-ARDS score were used: a high cutoff point (LUS-ARDS score > 27) for optimal specificity and low cutoff point (LUS-ARDS score > 8) for optimal sensitivity in two different cohorts. For this meta-analysis, the sensitivity and specificity of the recommended protocol or the one with the highest accuracy was used, together with the high cutoff point for both cohorts and the low cutoff point of both cohorts from the study of Smit et al.

Furthermore, we calculated a pooled specificity and sensitivity in which we excluded studies with an ARDS incidence greater than 30%.

Results

Selection of studies

Details regarding search strategy and study selection process are summarized in Fig. 1. After searching three databases, 2648 unique records were identified. After screening, retrieval, and assessment for eligibility, 11 studies were included in the review and underwent qualitative assessment.

Study characteristics

Characteristics of included studies are summarized in Table 1. Most of the included studies were conducted prospectively. Two were post-hoc studies. One study retrospectively analysed results of two previously reported studies. Settings included medical, surgical, and mixed ICUs or a combination thereof. One study was conducted on a medical ward and one in the emergency department. For the index test, a variety of LUS protocols were used as shown in Table 1. None of the included studies made note of the ultrasound settings used in their LUS protocols for detecting B-lines.

Of the nine studies investigating ARDS diagnosis, two studies used a quantitative analysis and awarded points for "A", "B" or "C" profiles. Seven of these studies had investigators diagnose ARDS based on descriptive parameters. These parameters were multiple B-lines, pleural line abnormalities, reduced or absent lung sliding, consolidations.

The accuracy of LUS to differentiate focal from non-focal ARDS subphenotypes was the primary outcome in two studies. These two studies both applied a quantitative analysis, where points were awarded for "A", "B" or "C" profiles and the total amount of points was added up and analysed.
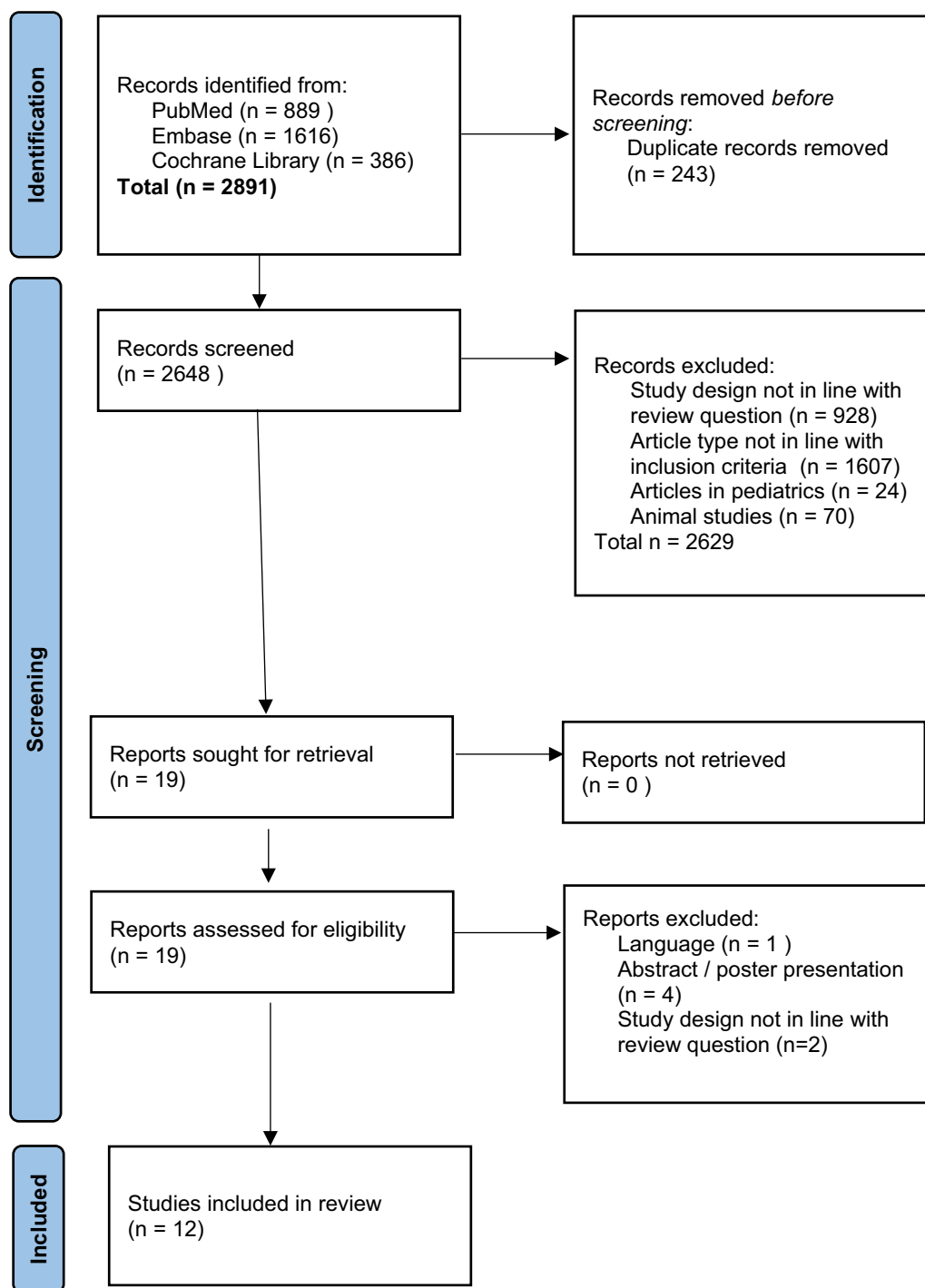
```
┌─────────────────────────────────────────────────────────────────────────────────────────┐
│ Identification                                                                            │
│  ┌────────────────────────────────────┐      ┌──────────────────────────────────────┐     │
│  │ Records identified from:           │ ───► │ Records removed before               │     │
│  │    PubMed (n = 889 )               │      │ screening:                           │     │
│  │    Embase (n = 1616)               │      │    Duplicate records removed         │     │
│  │    Cochrane Library (n = 386)      │      │    (n = 243)                         │     │
│  │    Total (n = 2891)                │      │                                      │     │
│  └────────────────────────────────────┘      └──────────────────────────────────────┘     │
└─────────────────────────────────────────────────────────────────────────────────────────┘
```

Records identified from:
   PubMed (n = 889)
   Embase (n = 1616)
   Cochrane Library (n = 386)
   Total (n = 2891)

Records removed *before screening*:
   Duplicate records removed (n = 243)

Records screened (n = 2648)

Records excluded:
   Study design not in line with review question (n = 928)
   Article type not in line with inclusion criteria (n = 1607)
   Articles in pediatrics (n = 24)
   Animal studies (n = 70)
Total n = 2629

Reports sought for retrieval (n = 19)

Reports not retrieved (n = 0)

Reports assessed for eligibility (n = 19)

Reports excluded:
   Language (n = 1)
   Abstract / poster presentation (n = 4)
   Study design not in line with review question (n=2)

Studies included in review (n = 12)

Fig. 1 PRISMA flow diagram presenting the study search strategy and study selection process

Outcomes

Eleven studies were included in total, involving a total of 2075 patients, of whom 598 (29%) had ARDS. Of these eleven studies, nine focused on LUS for ARDS diagnosis, involving 1977 patients, of whom 500 (25%) had ARDS.

The two studies focusing on differentiation between focal and non-focal subphenotypes involved 98 patients, all of whom had ARDS.

1138 (55%) of the included patients were mechanically ventilated, 383 (34%) of whom had ARDS. 467 patients

Boumans *et al. Critical Care*   (2024) 28:224

Page 5 of 15

Table 1 Study characteristics

References	Study design	Study period	Setting	Patient characteristics	N of ARDS/N patients included	Index test	Index test protocol	Reference standard	Primary outcome[a]	Secondary outcome(s)	N of operators	Study limitations
Daabis et al. [14]	Prospective	NS	Respiratory and general ICU	Acute respiratory failure	10/100	LUS	ARDS when either B-profile + PLAPS[b] Or B-profile present	History	ARDS	–	1	Patients with several final diagnoses were excluded Not specified which patients received chest X-Ray and/or chest CT Criteria for ARDS diagnosis were not specified
Bass et al. [15]	Prospective	2013	General ICU	Mechanically ventilated	35/77	SpO$_2$/FiO$_2$ and bilateral LUS	ARDS when ≥ 3 B-lines in single frozen frame in ≥ 1 lung field	PaO$_2$/FiO$_2$ ≤ 300 and chest radiograph consistent with ARDS	ARDS	–	1	Repeat LUS assessments performed unblinded for medical records Sole inclusion of intubated patients Protocol using only B-lines, disregarding other LUS signs Patient repositioning avoided, reducing visibility of posterior lung fields
Huang et al. [16]	Prospective	2014–2016	General ICU	Elderly, suspected of having ARDS	33/51	12 zone LUS protocol	ARDS when bilateral lung fields with either: Interstitial syndrome Consolidation Normal or abnormal pleural line or effusion	Berlin definition	ARDS	LUS + NT-proBNP and P/F ratio for ARDS diagnosis	NS	Finite study arms Unclear interval between LUS and thoracic CT

Table 1 (continued)

References	Study design	Study period	Setting	Patient characteristics	N of ARDS/N patients included	Index test	Index test protocol	Reference standard	Primary outcome[a]	Secondary outcome(s)	N of operators	Study limitations
See et al. [17]	Retrospective	2014–2017	Medical ICU	Noninvasive or invasive mechanically ventilated	216/456	Berlin LUS: Berlin definition using LUS as imaging criterium	ARDS when ≥ 1 region per hemithorax with > 2 B-lines or consolidation	Berlin definition	ARDS	Index test as predictor of ventilator free days, ICU/hospital LOS and ICU/hospital mortality	NS	Likely interobserver variability within each imaging modality Only medical ICU patients were included LUS were performed by respiratory therapists, limiting applicability
Pisani et al. [18]	Post-hoc	2016–2017	Medical and surgical ICU	Mechanically ventilated	35/152	Global and regional LUS Scores using 12 zone protocol	Lung aeration score[c] Optimal cutoff: global lung aeration score 15	Berlin definition	ARDS	–	1	No collection of LUS features such as lung sliding, pleural line abnormalities and subpleural Consolidations, which might add to diagnostic accuracy
Pierrakos et al. [4]	Post hoc	2021/2018[c]	General ICU	Amsterdam: mechanically ventilated Lombardy: mechanically ventilated with ARDS	51/51	Global and regional LUS Scores	Lung aeration scores[c] using Amsterdam, Lombardy and Piedmont method Anterior LUS ≥ 2 was strongly associated with nonfocal ARDS	CT scan	Focal vs nonfocal lung morphology	–	NS	Validity of difference between lateral and posterior LUS scores in Amsterdam not fully assessed No COVID-19 related ARDS cases were included

Table 1 (continued)

References	Study design	Study period	Setting	Patient characteristics	N of ARDS/N patients included	Index test	Index test protocol	Reference standard	Primary outcome[a]	Secondary outcome(s)	N of operators	Study limitations
Costamagna et al. [19]	Prospective	NS	General ICU	Mechanically ventilated and ARDS	47/47	Total, ventral, intermediate and dorsal LUS Scores	Lung aeration score[c] Ventral LUS ≥3 strongly associated with non-focal ARDS	CT scan	Focal vs nonfocal lung morphology	–	NS	Different spatial resolution of CT and LUS may affect the evaluation of lung aeration Cohort did not include patients with Covid-19 associated ARDS
Baid et al. [20]	Prospective	2019–2021	ED	Acute respiratory failure	35/237	Anterolateral and posterior thoracic LUS using BLUE-protocol	ARDS when Subpleural anterior consolidations Reduced/absent lung sliding Thickened/fragmented pleura Inhomogeneous distributed B-lines	Berlin definition	ARDS and other chest pathologies	Time to formulate POCUS diagnosis compared to composite reference	1	Uneven distribution of differentials of acute onset dyspnoea Possible clinical bias for reference diagnosis, as ED consultants were involved in active patient management
Chaitra and Hattholi [21]	Prospective	2017	General ICU	Acute respiratory failure	17/130	6 zone LUS using BLUE-protocol	2 ARDS criteria: (1) bilateral lung sliding + B3-lines (2) bilateral lung sliding, B3-lines + consolidation	NS	ARDS	–	1	Pleural effusion not included as distinct category ARDS and severe pneumonia not differentiated clinically Unclear diagnostic criteria for ARDS Unclear if LUS operators were blinded for clinical data

Boumans et al. Critical Care   (2024) 28:224

Page 8 of 15

Table 1 (continued)

References	Study design	Study period	Setting	Patient characteristics	N of ARDS/N patients included	Index test	Index test protocol	Reference standard	Primary outcome[a]	Secondary outcome(s)	N of operators	Study limitations
Arthur et al. [22]	Prospective	2016–2017	Medical Ward	Respiratory or cardiac reports and thoracic X-Ray performed	22/321	7 point LUS on each hemithorax (4 in sicker patients)	ARDS when Multiple B-lines, sub-pleural consolidations Air bronchograms	Final diagnosis at discharge	ARDS and other chest pathologies	Diagnostic accuracy of LUS in chest pathologies when CT scan was included in composite reference standard	1	Large frame between LUS and CXR Chance for incorporation bias, CXR included in composite reference standard Central chest and airway pathologies not seen on LUS
Smit et al. [23]	Prospective	2019–2021	General ICU	Mechanically ventilated	97/453	12-region LUS score	LUS-ARDS score 1 additional point in abnormal pleural line, consolidations, dynamic air bronchogram, pleural effusions LUS-ARDS score cutoffs: Low >8, high >27	Berlin definition	ARDS	LUS compared with clinical data and CXR Accuracy of LUS in scenarios where expert panel was not certain about ARDS diagnosis	3	Not all LUS signs associated with ARDS included in model Not all patients received index test because incomplete LUS regions Pragmatic sample size of validation cohort Low N nonpulmonary ARDS

Studies are presented according to publication date

For the index test, a variety of LUS protocols were used. Six studies analysed 12 lung regions. One study evaluated 14 points, 7 per hemithorax. In one study, the LUS protocol was not specified. A-profile is defined as lung sliding with A-lines and < 2 B-lines per view. B-profile is defined as > 3 B-lines per view. C-profile is a consolidation

NS not specified, *LUS* lung ultrasound, *LOS* length of stay, *CXR* chest X-ray, *ED* emergency department, *NCIS* non-cardiogenic interstitial syndrome, *ABG* arterial blood gas

[a] Primary outcome measured diagnostic accuracy in each study included

[b] PLAPS: Posterolateral Alveolar and/or Pleural Syndrome. Presence of consolidation or effusion at the most posterior part of the lung immediately superior to the diaphragm (PLAPS point)

[c] Lung aeration score: A total of 12 lung regions is scored according to the lung profile observed per zone. Zones are given a score of 0 in A-profile; 1 in > 2 B-lines; 2 in B-lines 50% pleural line; 3 in C pattern. Global aeration score is calculated by summing the scores of all 12 lung regions, while anterior, lateral or posterior aeration scores are calculated by summing their respective lung zones

Boumans *et al. Critical Care*   (2024) 28:224

Page 9 of 15

presented to the emergency department with acute respiratory failure, of whom 62 (13%) had ARDS.

Outcomes and diagnostic parameters of the included studies are presented in Online Appendix 1.

A meta-analysis comparing all studies with the diagnostic accuracy parameters using a low cutoff point from the study of Smit et al. resulted in a pooled sensitivity of 0.70 (95% CI 0.504–0.837) and pooled specificity of 0.93 (95% CI 0.789–0.977) and an area under the curve (AUC) of 0.88. Comparing the results of included studies with the high cutoff point with the study of Smit et al. resulted in a pooled sensitivity of 0.640. (95% CI 0.459–0.782) and pooled specificity of 0.94 (95% CI 0.861–0.979) and an AUC of 0.88. The Forest plots and sROC of these analyses can be found in Figs. 2 and 3, respectively.

After excluding studies with an ARDS incidence greater than 30%, a pooled sensitivity of 0.66 (95% CI 0.431–0.831) and pooled specificity of 0.96 (95% CI 0.825–0.993) and an AUC of 0.88 was found in the low cutoff group. A pooled sensitivity of 0.57 (95% CI 0.392–0.735) and pooled specificity of 0.97 (95% CI 0.930–0.990) and an AUC of 0.91 was found in the high cutoff group.

Two studies on LUS for focal versus non-focal ARDS subphenotypes also reported high specificity (71% and 100%) and sensitivity (94% and 100%).

Risk of *bias*

Risk of bias and applicability of studies included are summarized in Table 2. In studies assessing accuracy of LUS in ARDS diagnosis, one study was scored as high risk of bias in the patient selection domain, as this study excluded patients with non-respiratory or rare causes of acute respiratory failure or patients with multiple diagnoses at the end of hospitalization. For the same reason, this study was scored as high on applicability concerns in the patient selection domain.

One study was scored as unclear risk of bias as it was not specified whether researchers performing the index test were blinded for clinical data.

Four studies were scored as unclear risk of bias in the reference standard domain, as none of these studies specified the criteria used for ARDS diagnosis. For the same reason, they were scored as unclear on applicability concerns in the reference standard domain.

Four studies were scored as unclear risk of bias in the flow and timing domain, as none of these studies specified the time interval between the index test (LUS) and the diagnosis of ARDS by the reference standard.

A funnel plot was conducted to study if publication bias is present (Fig. 4).

The two studies assessing accuracy of LUS for focal versus non-focal ARDS subphenotypes diagnosis scored low on all domains of risk of bias and applicability.

Discussion

The major findings of this systematic review and meta-analysis in adult patients on the diagnostic accuracy of LUS for ARDS compared to reference standards which included the Berlin Definition are: (1) LUS has a high pooled specificity and a moderate sensitivity in ARDS diagnosis; (2) LUS can accurately diagnose focal versus non-focal ARDS subphenotypes with a high specificity and sensitivity; (3) The prevalence of ARDS found in the included studies matches previously reported prevalence and incidence rates of ARDS in patients with respiratory failure and ICU admissions [3, 16]. This indicates generalizability of the results.

The high pooled specificity found in the current study indicates that LUS is an adequate tool for diagnosing ARDS, but less adequate in ruling out ARDS. These findings are surprising as before LUS was found to be more sensitive compared to the Berlin definition, largely due to lack of specificity of LUS criteria. This contrasting finding could be largely attributed to the specific cutoff point chosen for the meta-analysis based on cutoff points used in the included studies. To evaluate the effect of a high or low chosen cutoff point on the results, we included both the high (LUS-ARDS score > 27) and low (LUS-ARDS score > 8) cutoff points of both the cohorts in the study of Smit et al. in the meta-analysis. As shown in the results, the pooled sensitivity and specificity are almost the same, underlining the finding that LUS is a specific tool for diagnosing ARDS.

Sensitivity varied considerably between studies. An explanation for lower sensitivity could be an imperfect reference standard. Some studies only used CXR as reference standard and CXR has been shown to be an unreliable diagnostic method alone for diagnosing ARDS [24]. The difference in sensitivity between the different studies might also be attributed to the absence of a standardized LUS ARDS definition. Unclear definitions on ARDS criteria might lead to over- or under classification and impact sensitivity. Also, ARDS diagnosis might be missed by LUS if presenting with A-profile in focal ARDS subphenotype. It was hypothesized that lung ultrasound in these cases may lack sensitivity because of the posterior-predominant aspect of ARDS and these regions are more difficult to scan using LUS in critically ill patients [17], especially in less experienced operators. Furthermore, ARDS might be more difficult to detect on LUS in initial stages of the disease. In summary, LUS is an accurate tool for diagnosing ARDS and possibly less for ruling out ARDS.

To assess whether a high incidence of ARDS impacted results, we also evaluated the pooled sensitivity and specificity while in- and excluding studies with an incidence
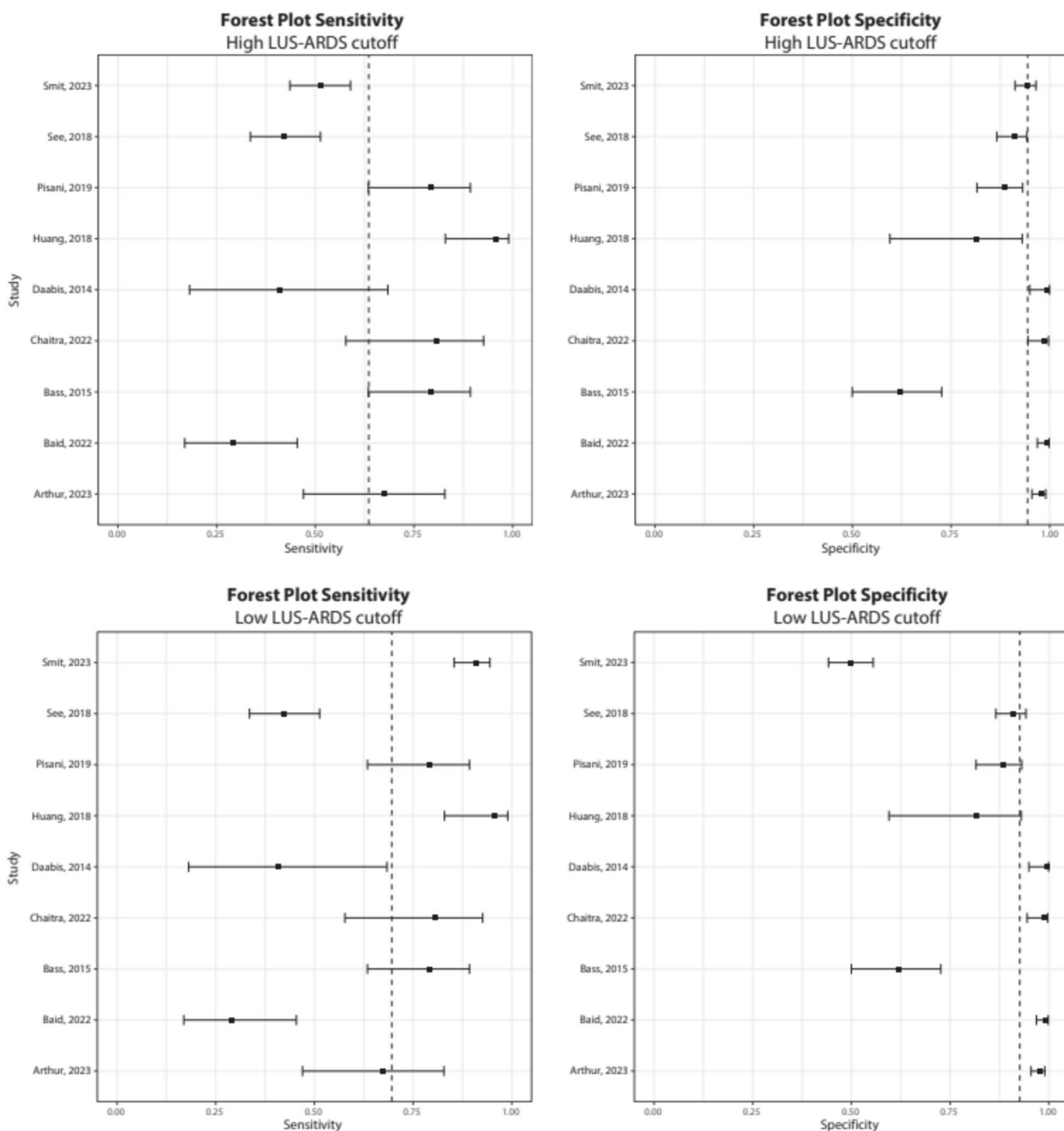
Boumans *et al. Critical Care*   (2024) 28:224

Page 10 of 15

Fig. 2 Forest plots comparing diagnostic accuracy parameters from all studies with parameters from the study of Smit et al., after using a low and high cutoff point as suggested by the authors of this study. The dashed line represents median values of sensitivity and specificity. Sensitivity and confidence intervals: Smit (high cutoff): 0.51 [0.44–0.59] Smit (low cutoff): 0.91 [0.85–0.94], Daabis: 0.41 [0.18–0.68], Bass: 0.79 [0.63–0.89], Huang: 0.96 [0.83–0.99], See: 0.42 [0.34–0.51], Pisani: 0.79 [0.63–0.89], Baid: 0.29 [0.17–0.45], Chaitra: 0.81 [0.58–0.93], Arthur: 0.67 [0.47–0.83]. Specificity and confidence intervals: Smit (high cutoff): 0.94 [0.91–0.97] Smit (low cutoff): 0.50 [0.44–0.55], Daabis: 0.99 [0.95–1.00], Bass: 0.62 [0.50–0.73], Huang: 0.82 [0.60–0.93], See: 0.91 [0.87–0.94], Pisani: 0.89 [0.82–0.93], Baid: 0.99 [0.97–1.00], Chaitra: 0.99 [0.94–1.00], Arthur: 0.98 [0.95–0.99]

of ARDS over 30%. As can be seen in the results, this does not have a significant effect on the pooled results.

Different ARDS subphenotypes might require different treatment strategies, e.g. prone positioning for focal ARDS and lung-recruitment maneuvers for non-focal ARDS. Correct identification of ARDS subphenotype will impact treatment and outcomes in the future. This underlines the importance of accurate distinction
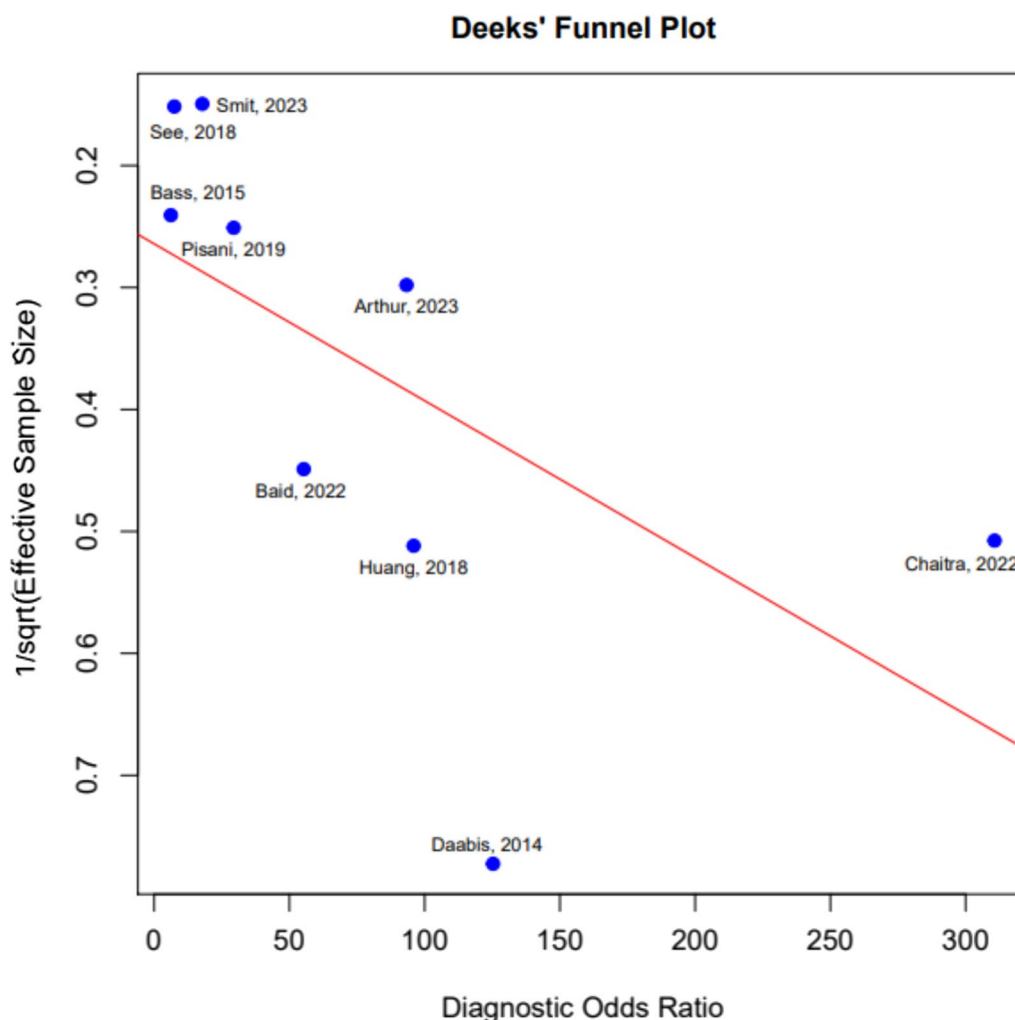
Fig. 3 sROC curve after meta-analysis of nine studies assessing diagnostic accuracy of LUS in ARDS diagnosis

Table 2 Risk of bias and applicability assessment of included studies, following the QUADAS-2 scoring tool

Study	Risk of bias				Applicability		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Studies comparing LUS to Berlin definition in ARDS diagnosis							
Pisani et al. [18]	Low	Low	Low	*Unclear*	Low	Low	Low
Arthur et al. [22]	Low	*Low*	*Unclear*	Low	Low	Low	*Unclear*
Baid et al. [20]	Low	*Low*	*Unclear*	Low	Low	Low	*Unclear*
Bass et al. [15]	Low	Low	Low	Low	Low	Low	Low
Chaitra and Hattiholi [21]	Low	*Unclear*	*Unclear*	*Unclear*	Low	Low	*Unclear*
Daabis et al. [14]	*High*	Low	*Unclear*	*Unclear*	*High*	Low	*Unclear*
Huang et al. [16]	Low	Low	Low	*Unclear*	Low	Low	Low
See et al. [17]	Low	Low	Low	Low	Low	Low	Low
Smit et al. [23]	Low	Low	Low	Low	Low	Low	Low
Studies comparing LUS to Chest CT in ARDS phenotype diagnosis							
Pierrakos et al. [4]	Low	Low	Low	Low	Low	Low	Low
Costamagna et al. [19]	Low	Low	Low	Low	Low	Low	Low

Boumans *et al. Critical Care*   (2024) 28:224

Page 12 of 15

Deeks' Funnel Plot

Fig. 4 Deek's Funnel plot for assessment of publication bias

between ARDS subphenotypes. Two studies investigated the accuracy of LUS for classification of focal versus non-focal ARDS subphenotypes compared to CT. The reported sensitivity and specificity were good [4, 18], indicating LUS as a reliable diagnostic tool for this distinction. However, since only few studies addressing this subject have been published, there is a need for further validation. In addition, a study is being conducted to see if tailoring mechanical ventilation to lung morphology assessed by LUS reduces mortality in ARDS patients compared to standard ventilation (NCT05492344).

There are many advantages of using LUS as a replacement for CT in ARDS diagnostic work-up. It can lead to an earlier start of treatment and reduce workload on ICU personnel. Furthermore, patient transport is a high-risk procedure in ICU patients and can lead to severe adverse events. LUS reduces exposure to irradiation, lowers health care costs and provides real-time examination of lung parenchyma, which also makes it ideal for follow-up of disease severity and facilitates bedside titration of ventilatory parameters. Correct subtype identification of ARDS using LUS will aid in implementing the correct treatment strategy and adequate timing of prone positioning.

There are more advantages of LUS expanding beyond ARDS diagnosis. Especially in light of the recent COVID-19 pandemic, but also in case of other respiratory emergencies, LUS can contribute to earlier identification of diseases in the emergency department (ED), allowing for earlier start of treatment and earlier admission to the ward or ICU and thus reducing the workload on ED personnel.

Of course, performing and interpreting LUS requires training. However, studies show that after 8 h of training, clinicians show proficiency in the interpretation of LUS images [25].

Boumans *et al. Critical Care*   (2024) 28:224

Page 13 of 15

Currently, no consistent and/or optimal LUS protocol for diagnosing ARDS exists based on the included studies. However, among the included studies there are overlapping criteria used to establish the diagnosis. Multiple (> 2 per view) bilateral, non-homogenous B-lines, present in at least one area per hemithorax and the presence of subpleural consolidations were seen as indicative of ARDS in all studies. These criteria were confirmed by studies differentiating cardiogenic edema from non-cardiogenic interstitial syndrome, which includes ARDS but also interstitial lung disease [26–28]. Pleural line abnormalities such as irregular thickening or fragmentation were also observed, as well as areas of preserved lung parenchyma and pleural effusion. Consolidation accompanied by pleural effusion was a marker of compression atelectasis, and therefore seen as not indicative of ARDS. It seems imperative that in the future these 'typical LUS findings' and/or a LUS-ARDS score should be included in the new ARDS diagnostic criteria, as it was recently shown that including LUS per se increased the occurrence rate of ARDS [29].

Amongst the included studies, 6, 12 and 14 point lung ultrasound protocols were described. The optimal amount of lung regions to be included in a LUS protocol remains a subject of discussion, as 6 point protocols are more practically applicable but might lack the amount of diagnostic information compared to 12 or 18 point LUS protocols. Recent studies have shown that 6 point LUS protocols can yield similar diagnostic results compared to 12 point protocols. However, these studies have not been validated in ARDS patients [18, 30].

ARDS remains unrecognized in 20–50% of all cases, while mortality rates remain high [3]. This underlines the importance of an improved definition. Since 2012, the Berlin definition has been used as the golden standard in diagnosing patients with ARDS [7]. In 2016, the Kigali modification was proposed and validated for low-income countries, in which the bilateral opacities had to be present on either LUS or CXR [31]. Several studies already used the Kigali modification in settings with scarce access to more advanced imaging modalities [32, 33].

Also, in the global ARDS definition, lung ultrasound was proposed as an alternative for the imaging criterium for ARDS diagnosis. This review shows a good specificity but moderate sensitivity for LUS in ARDS diagnosis. LUS as a diagnostic tool for ARDS seems promising, and we support implementation of LUS in the global ARDS definition. However, we would like to highlight the importance of further prospective research and standardization of ARDS LUS definitions. In addition, an interesting recent development is the pairing of LUS with a deep learning model, which was able to distinguish COVID ARDS, non-COVID ARDS and hydrostatic pulmonary edema [34].

Strengths and limitations

This systematic review has several strengths. To the best of our knowledge, this is the first systematic review on this subject, despite LUS already being added to the global ARDS definition. Secondly, the rigorous search strategy and method used are important strengths of this review. Another strength is the number of included studies and relatively large sample size. Of the eleven included studies, seven studies had a study population of 100 or more participants.

This review also has several limitations. Our meta-analyses included a number of small studies. Research analyzing the effect of small studies on treatment effect found that these trials often report a more positive treatment effect [35]. This might also hold true for diagnostic accuracy studies. Therefore, we advise to be careful with the interpretation of the smaller trials and limited number of cases, where a false-negative or positive results can affect accuracy more strongly than larger studies. There was sufficient data between the nine studies investigating the accuracy of LUS in ARDS diagnosis, which allowed for a meta-analysis to be performed. The studies in this review showed high sensitivity and specificity for distinguishing between focal and non-focal ARDS subphenotypes, but only two studies focused on this research question. A limited number of studies can influence the beneficial effect of the outcome researched [36]. Furthermore, there is a moderate heterogeneity between the included studies. Because there is no clear consensus yet on how to diagnose ARDS based on LUS alone, each study applied a different LUS ARDS definition, which could influence the sensitivity and specificity between different studies. A recent study showed a higher occurrence rate of ARDS when adding bilateral abnormalities as found by LUS. The authors concluded that incorporating well-defined LUS ARDS criteria in the new ARDS definitions (Kigali and Global Definition) will improve sensitivity and specificity [29]. In none of the included studies, the ultrasound preset was defined. The choice of preset can significantly influence the detection of B-lines, thereby impacting both sensitivity and specificity.

This meta-analysis assessed the risk of publication bias by funnel plot analysis. However, this should be interpreted with caution as this is primarily designed for interventional studies and it is unclear if publication bias exists for diagnostic accuracy test studies [37].

Boumans *et al. Critical Care*   (2024) 28:224

Page 14 of 15

Conclusion

This systematic review and meta-analysis demonstrate that LUS is a reliable diagnostic tool for ARDS in adult patients. The high specificity indicates that it is especially good for diagnosing ARDS, with a moderate sensitivity making it moderately reliable for ruling out ARDS. These results support the implementation of LUS in the global ARDS definition. However, given the significant heterogeneity amongst included studies, this review warrants the need for further clinical research. We would recommend a standardized LUS ARDS definition, including a preferred LUS protocol including presets for B-line recognition. This will improve homogeneity between studies and improve clinical applicability in real-life settings.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13054-024-04985-1.

> Supplementary Material 1.

Author contributions
MB and WA wrote the main manuscript text and prepared the tables and figures. MS provided statistical analysis and Figs. 2 and 3. PRT was consulted in case of disagreement between MB and WA. All authors reviewed the manuscript.

Availability of data and materials
All data supporting the findings of this systematic review can be found in the manuscript or supplementary material.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Intensive Care Medicine, Noordwest Ziekenhuisgroep, Wilhelminalaan 12, Alkmaar, The Netherlands. [2]Department of Intensive Care Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, Amsterdam, The Netherlands. [3]Department of Intensive Care Medicine, Amsterdam UMC, Amsterdam Medisch Centrum, Meibergdreef 9, Amsterdam, The Netherlands. [4]Mahidol-Oxford Tropical Medicine Research Unit (MORU), Mahidol University, Bangkok 10400, Thailand. [5]Amsterdam Leiden IC Focused Echography (ALIFE), Amsterdam, The Netherlands. [6]Amsterdam Cardiovascular Sciences, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, Amsterdam, The Netherlands. [7]Amsterdam Institute for Immunity and Infectious Diseases, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, Amsterdam, The Netherlands. [8]Department of Precision-Regenerative Medicine and Jonic Area (DiMePRe-J), Section of Anesthesiology and Intensive Care Medicine, University of Bari "Aldo Moro", Bari, Italy. [9]Laboratory of Experimental Intensive Care and Anesthesiology (LEICA), University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands.

References

1. Meyer NJ, Gattinoni L, Calfee CS. Acute respiratory distress syndrome. Lancet. 2021;398(10300):622–37. https://doi.org/10.1016/s0140-6736(21)00439-6.
2. Fanelli V, Vlachou A, Ghannadian S, Simonetti U, Slutsky AS, Zhang H. Acute respiratory distress syndrome: new definition, current and future therapeutic options. J Thorac Dis. 2013;5(3):326–34. https://doi.org/10.3978/j.issn.2072-1439.2013.04.05.
3. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, Gattinoni L, van Haren F, Larsson A, McAuley DF, Ranieri M, Rubenfeld G, Thompson BT, Wrigge H, Slutsky AS, Pesenti A. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA. 2016;315(8):788. https://doi.org/10.1001/jama.2016.0291.
4. Pierrakos C, Smit MR, Pisani L, Paulus F, Schultz MJ, Constantin J-M, Chiumello D, Mojoli F, Mongodi S, Bos LD. Lung ultrasound assessment of focal and non-focal lung morphology in patients with acute respiratory distress syndrome. Front Physiol. 2021;12: 730857. https://doi.org/10.3389/fphys.2021.730857.
5. Khan YA, Fan E, Ferguson ND. Precision medicine and heterogeneity of treatment effect in therapies for ARDS. Chest. 2021;160(5):1729–38. https://doi.org/10.1016/j.chest.2021.07.009.
6. Constantin J-M, Jabaudon M, Lefrant J-Y, Jaber S, Quenot J-P, Langeron O, Ferrandière M, Grelon F, Seguin P, Ichai C, Veber B, Souweine B, Uberti T, Lasocki S, Legay F, Leone M, Eisenmann N, Dahyot-Fizelier C, Dupont H, Nanadougmar H. Personalised mechanical ventilation tailored to lung morphology versus low positive end-expiratory pressure for patients with acute respiratory distress syndrome in France (the live study): a multicentre, single-blind, randomised controlled trial. Lancet Respir Med. 2019;7(10):870–80. https://doi.org/10.1016/s2213-2600(19)30138-9.
7. Ranieri VM, Rubenfeld GD, Taylor Thompson B, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS. Acute respiratory distress syndrome. JAMA. 2012;307(23): 2526. https://doi.org/10.1001/jama.2012.5669.
8. Riviello ED, Kiviri W, Twagirumugabe T, Mueller A, Banner-Goodspeed VM, Officer L, Novack V, Mutumwinka M, Talmor DS, Fowler RA. Hospital incidence and outcomes of the acute respiratory distress syndrome using the Kigali modification of the Berlin definition. Am J Respir Crit Care Med. 2016;193(1):52–9. https://doi.org/10.1164/rccm.201503-0584oc.
9. Matthay MA, Arabi Y, Arroliga AC, Bernard G, Bersten AD, Brochard LJ, Calfee CS, Combes A, Daniel BM, Ferguson ND, Gong MN, Gotts JE, Herridge MS, Laffey JG, Liu KD, Machado FR, Martin TR, McAuley DF, Mercat A, Wick KD. A new global definition of acute respiratory distress syndrome. Am J Respir Crit Care Med. 2024;209(1):37–47. https://doi.org/10.1164/rccm.202303-0558ws.
10. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, Mierzwinski-Urban M, Clifford T, Hutton B, Rabb D. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care. 2012;28(2):138–44. https://doi.org/10.1017/s0266462312000086.
11. Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. BMC Med Res Methodol. 2017;17(1):1–11. https://doi.org/10.1186/s12874-017-0347-z.
12. Nussbaumer-Streit B, Klerings I, Dobrescu AI, Persad E, Stevens A, Garritty C, Kamel C, Affengruber L, King VJ, Gartlehner G. Excluding non-English publications from evidence-syntheses did not change conclusions: a meta-epidemiological study. J Clin Epidemiol. 2020;118:42–54. https://doi.org/10.1016/j.jclinepi.2019.10.011.
13. Dobrescu A, Nussbaumer-Streit B, Klerings I, Wagner G, Persad E, Sommer I, Herkner H, Gartlehner G. Restricting evidence syntheses of interventions to English-language publications is a viable methodological shortcut for most medical topics: a systematic review. J Clin Epidemiol. 2021;137:209–17. https://doi.org/10.1016/j.jclinepi.2021.04.012.
14. Daabis R, Banawan L, Rabea A, Elnakedy A, Sadek A. Relevance of chest sonography in the diagnosis of acute respiratory failure: comparison with current diagnostic tools in intensive care units. Egypt J Chest Dis Tubercul. 2014;63(4):979–85. https://doi.org/10.1016/j.ejcdt.2014.05.005.
15. Bass CM, Sajed DR, Adedipe AA, West TE. Pulmonary ultrasound and pulse oximetry versus chest radiography and arterial blood gas analysis for the

Boumans *et al. Critical Care*   (2024) 28:224

Page 15 of 15

diagnosis of acute respiratory distress syndrome: a pilot study. Crit Care. 2015;19(1):1–11. https://doi.org/10.1186/s13054-015-0995-5.

16. Huang D, Ma H, Xiao Z, Blaivas M, Chen Y, Wen J, Guo W, Liang J, Liao X, Wang Z, Li H, Li J, Chao Y, Wang X, Wu Y, Qin T, Su K, Wang S, Tan N. Diagnostic value of cardiopulmonary ultrasound in elderly patients with acute respiratory distress syndrome. BMC Pulm Med. 2018;18(1):1–9. https://doi.org/10.1186/s12890-018-0666-9.

17. See KC, Ong V, Tan YL, Sahagun J, Taculod J. Chest radiography versus lung ultrasound for identification of acute respiratory distress syndrome: a retrospective observational study. Crit Care. 2018;22(1):1–9. https://doi.org/10.1186/s13054-018-2105-y.

18. Pisani L, Vercesi V, van Tongeren PS, Lagrand WK, Leopold SJ, Huson MA, Henwood PC, Walden A, Smit MR, Riviello ED, Pelosi P, Dondorp AM, Schultz MJ. The diagnostic accuracy for ARDS of global versus regional lung ultrasound scores—a post hoc analysis of an observational study in invasively ventilated ICU patients. Intensive Care Med Exp. 2019;7(S1):1–11. https://doi.org/10.1186/s40635-019-0241-6.

19. Costamagna A, Pivetta E, Goffi A, Steinberg I, Arina P, Mazzeo AT, Del Sorbo L, Veglia S, Davini O, Brazzi L, Ranieri VM, Fanelli V. Clinical performance of lung ultrasound in predicting ARDS morphology. Ann Intensive Care. 2021;11(1):1–8. https://doi.org/10.1186/s13613-021-00837-1.

20. Baid H, Vempalli N, Kumar S, Arora P, Walia R, Chauhan U, Shukla K, Verma A, Chawang H, Agarwal D. Point of care ultrasound as initial diagnostic tool in acute dyspnea patients in the Emergency Department of a Tertiary Care Center: diagnostic accuracy study. Int J Emerg Med. 2022;15(1):27. https://doi.org/10.1186/s12245-022-00430-8.

21. Chaitra S, Hattiholi VV. Diagnostic accuracy of bedside lung ultrasound in emergency protocol for the diagnosis of acute respiratory failure. J Med Ultrasound. 2022;30(2):94. https://doi.org/10.4103/jmu.jmu_25_21.

22. Arthur M, Pichamuthu K, Turaka VP, Putta T, Jeeyavudeen MS, Zachariah A, Sathyendra S, Hansdak SG, Iyadurai R, Karuppusami R, Sudarsanam TD. Bedside Lung ultrasonography: comparison with chest radiography (Blur), a diagnostic study in a developing country. Postgrad Med J. 2022;99(1173):724–30. https://doi.org/10.1136/pmj-2021-141343.

23. Smit MR, Hagens LA, Heijnen NF, Pisani L, Cherpanath TG, Dongelmans DA, de Grooth H-JS, Pierrakos C, Tuinman PR, Zimatore C, Paulus F, Schnabel RM, Schultz MJ, Bergmans DC, Bos LD, Bos LD, Hagens LA, Schultz MJ, Smit MR, van der Ven FLIM, Bergmans DCJJ, Gietema HA, Gerretsen SC, Heijnen NFL, Schnabel RM, Geven I, Nijsen TME, Verschueren AR. Lung ultrasound prediction model for acute respiratory distress syndrome: a multicenter prospective observational study. Am J Respir Crit Care Med. 2023;207(12):1591–601. https://doi.org/10.1164/rccm.202210-1882OC.

24. Rubenfeld GD, Caldwell E, Granton J, Hudson LD, Matthay MA. Interobserver variability in applying a radiographic definition for ARDS. Chest. 1999;116(5):1347–53. https://doi.org/10.1378/chest.116.5.1347.

25. House DR, Amatya Y, Nti B, Russell FM. Lung ultrasound training and evaluation for proficiency among physicians in a low-resource setting. Ultrasound J. 2021;13(1):34. https://doi.org/10.1186/s13089-021-00236-4.

26. Copetti R, Soldati G, Copetti P. Chest sonography: a useful tool to differentiate acute cardiogenic pulmonary edema from acute respiratory distress syndrome. Cardiovasc Ultrasound. 2008;6(1):1–10. https://doi.org/10.1186/1476-7120-6-16.

27. Heldeweg MLA, Haaksma ME, Smit JM, Smit MR, Tuinman PR. Lung ultrasound to discriminate non-cardiogenic interstitial syndrome from cardiogenic pulmonary edema: Is "Gestalt" as good as it gets? J Crit Care. 2023;73: 154180. https://doi.org/10.1016/j.jcrc.2022.154180.

28. Heldeweg ML, Smit MR, Kramer-Elliott SR, Haaksma ME, Smit JM, Hagens LA, Heijnen NF, Jonkman AH, Paulus F, Schultz MJ, Girbes AR, Heunks LM, Bos LD, Tuinman PR. Lung ultrasound signs to diagnose and discriminate interstitial syndromes in ICU patients: a diagnostic accuracy study in two cohorts*. Crit Care Med. 2022;50(11):1607–17. https://doi.org/10.1097/ccm.0000000000005620.

29. Plantinga C, Klompmaker P, Haaksma ME, Mousa A, Blok SG, Heldeweg MLA, Paulus F, Schultz MJ, Tuinman PR. Use of lung ultrasound in the new definitions of acute respiratory distress syndrome increases the occurrence rate of acute respiratory distress syndrome. Crit Care Med. 2023. https://doi.org/10.1097/ccm.0000000000006118.

30. Heldeweg MLA, Lieveld AWE, de Grooth HJ, Heunks LMA, Tuinman PR, ALIFE Study Group. Determining the optimal number of lung ultrasound zones to monitor COVID-19 patients: Can we keep it ultra-short and ultra-simple? Intensive Care Med. 2021;47(9):1041–3. https://doi.org/10.1007/s00134-021-06463-6.

31. Vercesi V, Pisani L, van Tongeren PS, Lagrand WK, Leopold SJ, Huson MM, Henwood PC, Walden A, Smit M, Riviello ED, Pelosi P, Dondorp AM, Schultz MJ. External confirmation and exploration for the Kigali modification for diagnosing moderate or severe ARDS. Intensive Care Med. 2018;44(4):523–4. https://doi.org/10.1007/s00134-018-5048-5.

32. Leopold SJ, Ghose A, Plewes KA, Mazumder S, Pisani L, Kingston HW, Paul S, AnupamBarua M, Sattar A, Huson MAM, Walden AP, Henwood PC, Riviello ED, Schultz MJ, Day NPJ, Dutta AK, White NJ, Dondorp AM. Point-of-care lung ultrasound for the detection of pulmonary manifestations of malaria and sepsis: an observational study. PLoS ONE. 2018;13(12): e0204832. https://doi.org/10.1371/journal.pone.0204832.

33. Pisani L, De Nicolo A, Schiavone M, Adeniji AO, De Palma A, Di Gennaro F, Emuveyan EE, Grasso S, Henwood PC, Koroma AP, Leopold S, Marotta C, Marulli G, Putoto G, Pisani E, Russel J, Neto AS, Dondorp AM, Hanciles E, Koroma MM, Schultz MJ. Lung ultrasound for detection of pulmonary complications in critically ill obstetric patients in a resource-limited setting. Am J Trop Med Hyg. 2021;104(2):478–86. https://doi.org/10.4269/ajtmh.20-0996.

34. Arntfield R, VanBerlo B, Alaifan T, Phelps N, White M, Chaudhary R, Ho J, Wu D. Development of a convolutional neural network to differentiate among the etiology of similar appearing pathological B lines on lung ultrasound: a deep learning study. BMJ Open. 2021;11(3): e045120. https://doi.org/10.1136/bmjopen-2020-045120.

35. Nuesch E, Trelle S, Reichenbach S, Rutjes AW, Tschannen B, Altman DG, Egger M, Juni P. Small study effects in meta-analyses of osteoarthritis trials: Meta-epidemiological study. BMJ. 2010;341(jul16 1):c3515–c3515. https://doi.org/10.1136/bmj.c3515.

36. Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. Clin Chem. 2012;58(12):1636–43. https://doi.org/10.1373/clinchem.2012.182576.

37. van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. BMC Med Res Methodol. 2014;14(1):1–11. https://doi.org/10.1186/1471-2288-14-70.

Publisher's Note