Diagnostic accuracy of lung ultrasound to predict weaning outcome: a systematic review and meta-analysis

Zhiyang Zhang[1,2†], Li Guo[3†], Huawei Wang[2], Ze Zhang[2], Limin Shen[2] and Heling Zhao[1,2]*

[1]Department of Critical Care Medicine, Hebei Medical University, Shijiazhuang, China, [2]Department of Intensive Care Unit, Hebei General Hospital, Shijiazhuang, China, [3]Department of Neonatal, Shijiazhuang Fourth Hospital, Shijiazhuang, China

Background: This systematic review and meta-analysis aim to systematically assess the diagnostic accuracy of lung ultrasound in predicting weaning failure from mechanical ventilation in critically ill patients.

Methods: We searched the relevant literature up to January 2024 in the databases Web of Science, Cochrane Library, Embase, and PubMed. Two researchers independently screened eligible studies and extracted data; disagreements, if any, were resolved through discussion or consultation with a third-party expert. The quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool. Statistical analyses were performed using Review Manager version 5.3 and Stata version 18.0, applying bivariate random-effects models to estimate sensitivity, specificity, diagnostic odds ratios, and their 95% confidence intervals, as well as to summarize receiver operating characteristic curves. Inter-study heterogeneity was assessed using the I-squared statistic, and potential sources of heterogeneity were explored by meta-regression analysis. The study follows the guidelines for Preferred Reporting Items for Systematic Reviews and Meta-Analyses in reporting.

Results: Fourteen studies were included in the systematic review, of which 13 studies (totaling 988 patients) were included in the meta-analysis. The meta-analysis revealed an overall sensitivity of 0.86 (95% confidence interval: 0.77−0.91) and a specificity of 0.75 (95% confidence interval: 0.66−0.83) for lung ultrasound in predicting extubation failure. The area under the receiver operating characteristic curve was 0.87 (95% confidence interval: 0.84−0.89). Meta-regression analysis identified lung ultrasound thresholds, reference standards (extubation outcomes), and study flow and time bias as significant factors influencing diagnostic accuracy.

Conclusion: This systematic review and meta-analysis demonstrated that lung ultrasound has high diagnostic accuracy in predicting extubation failure in mechanically ventilated critically ill patients. Despite some study heterogeneity, lung ultrasound proved to be a reliable predictive tool for extubation failure. Future research should focus on standardizing the definition of extubation failure, exploring the impact of different thresholds on the predictive ability of lung ultrasound, and validating its application in various clinical settings to enhance its utility and accuracy in clinical practice.

Systematic review registration: This systematic review and meta-analysis was registered with PROSPERO (registration number: CRD42024555909). The study adhered to the guidelines set by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Details of the PROSPERO protocol can be found in Supplementary Table 1.

Introduction

One of the major challenges faced by the scientific community is determining the optimal timing for extubation. This involves assessing the patient's current health and predicting potential post-extubation complications, making the decision-making process particularly complex (1). Although mechanical ventilation (MV) is widely used in intensive care units (ICUs) and saves countless lives daily, prolonged reliance on mechanical ventilation significantly increases the risk of mortality and complications (2–5). If extubation is performed prematurely, the patient may require reintubation due to failure to wean from the ventilator, exposing them to unnecessary hemodynamic and respiratory stress. Conversely, delayed extubation prolongs mechanical ventilation, increasing the risk of ventilator-associated pneumonia, tracheal injuries, and barotrauma (6–8). Thus, both premature and delayed extubation can elevate mortality, prolong ICU stays, and increase healthcare costs (9). To optimize the timing of ventilator withdrawal, healthcare professionals rely on various indices to assess a patient's readiness for extubation. These indices include physiological measures such as minute ventilation and inspiratory pressure, along with clinical indicators like the rapid shallow breathing index and tracheal obstruction pressure. This comprehensive approach helps to navigate the complexities involved in extubation decision-making. However, despite these methods, studies have shown that approximately 20% of patients still experience challenges during weaning and extubation (10, 11).

Ultrasound technology has proven to be a highly reliable and accurate tool for monitoring internal organ abnormalities, often surpassing traditional radiological methods (12–14). In particular, ultrasound assessments of diaphragm function, such as diaphragm excursion (DE) and diaphragm thickening fraction (DTF), have been shown to predict weaning success (15, 16). Meta-analyses also support their role in predicting weaning failure (17–20). A study by Llamas-Álvarez et al. (19) suggested that the lung ultrasound score (LUS) may serve as a valid predictor of extubation outcomes, although the study included only five trials, and the results were less conclusive.

Lung ultrasound (LUS) scoring, as initially described by Bouhemad et al. (21), quantitatively assesses pulmonary ventilation by examining different lung regions bilaterally, determining whether lung ventilation is adequate or compromised. This method has been clinically applied to monitor the pulmonary status of mechanically ventilated patients, guiding decisions regarding readiness for weaning. Over time, the use of LUS has enabled clinicians to track changes in lung conditions, aiding in informed decision-making about the continuation or cessation of mechanical support. Despite its potential, variability in LUS score thresholds across studies—used to determine successful spontaneous breathing trials (SBT) and extubation—has led to inconsistent conclusions. Therefore, this systematic review and meta-analysis aims to delve deeper into the diagnostic accuracy of lung ultrasound in predicting extubation failure among mechanically ventilated critically ill patients, providing more precise and practical insights for clinical practice.

Methods

Study protocol

This systematic review and meta-analysis aimed to evaluate the diagnostic accuracy of the lung ultrasound score (LUS) in predicting extubation failure among mechanically ventilated, critically ill patients. Before initiating the study, a detailed protocol was developed and registered with PROSPERO (registration number: CRD42024555909) (see Supplementary Table 1 for the PROSPERO Protocol). The study adhered strictly to the guidelines set by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (22).

Literature search and selection

Two independent reviewers conducted comprehensive literature searches in databases, including Web of Science, Cochrane Library, Embase, and PubMed, up until January 7, 2024. To ensure thorough coverage, references from relevant studies were also examined, without imposing restrictions on study design or language. EndNote software was used for managing and screening the literature. The search process continued until no additional relevant studies were found. Search terms included Medical Subject Headings (MeSH) and other standardized terms, and a concept-based approach was employed to capture terms related to "extubation," "weaning," "ultrasound," "ultrasonography," and "echography." The complete search strategy is available in Supplementary Table 2. In cases of disagreement between reviewers, a third researcher was consulted to reach a consensus.

Inclusion and exclusion criteria

Studies were included if they met the following criteria: (1) critically ill adult patients (aged ≥18 years) on mechanical ventilation in the ICU who were deemed ready for withdrawal, and (2) patients who had undergone lung ultrasonography with complete data on withdrawal outcomes. Extubation failure was defined as the need for reintubation (invasive or noninvasive mechanical ventilation),

tracheostomy, or death within 72 h of withdrawal, or failure of the spontaneous breathing trial (SBT). Success of withdrawal was defined as the absence of extubation failure criteria. Exclusion criteria were: (1) duplicate or overlapping studies, (2) studies lacking sufficient data to construct a 2×2 table, and (3) unpublished data.

Rationale for excluding patients with noninvasive ventilation post-extubation

Patients who received noninvasive ventilation post-extubation were excluded, as these patients might require reintubation, which could confound the assessment of extubation outcomes. While noninvasive ventilation is commonly used to enhance lung recruitment and reduce respiratory muscle workload, its use could obscure the true outcomes of weaning failure. This exclusion was intended to reduce variability and ensure that our study results accurately reflect the diagnostic accuracy of lung ultrasound scores in predicting extubation failure. We acknowledge that this decision may limit the generalizability of our findings but believe it is essential for maintaining scientific rigor and ensuring the practical utility of the results.

Data extraction

Two researchers independently extracted data from the included studies, with a third researcher resolving any disagreements. The extracted data included the study ID (first author and year of publication), country, study design, setting, inclusion and exclusion criteria, ultrasound probe frequency, lung regions examined, the scoring system used, critical score values, methods for spontaneous breathing trials, and ultrasound parameters predicting extubation failure (categorized as true positives, true negatives, false positives, and false negatives).

Quality assessment and publication bias

The methodological quality of each study was evaluated using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool (23), which assesses four areas of bias: patient selection, index tests, reference standards, and process and timing. Publication bias was assessed using funnel plots, and potential bias was determined by examining graphical asymmetry (24).

Statistical analysis

Statistical analyses were conducted using Review Manager 5.3 and Stata 18.0 software, utilizing the "metandi" function (25). Bivariate meta-analysis models (26) were applied to estimate sensitivity, specificity, positive and negative likelihood ratios, and diagnostic odds ratios, with corresponding 95% confidence intervals calculated. Summary receiver operating characteristic (SROC) curves and their areas under the curve (AUC) were generated to evaluate the accuracy of lung ultrasound scores in predicting extubation failure (27). The closer the SROC curve is to the upper left corner, the larger the AUC, indicating higher overall test accuracy. Non-threshold heterogeneity was assessed using the

Q-value and I^2 statistic (28). Fagan plots were applied to assess the clinical applicability of lung ultrasound in diagnosing weaning outcomes.

Meta-regression analysis

Meta-regression analyses were performed to explore study characteristics associated with the diagnostic accuracy of diaphragmatic ultrasound and to identify potential sources of heterogeneity in sensitivity and specificity. These characteristics included the cut-off value of the lung ultrasound score (LUS), index test, reference standard, process and timing, patient selection, and study quality.

Results

Study identification and selection

A detailed flowchart of the literature search is shown in Figure 1. The initial database search yielded 2,617 references. After the removal of duplicates, 2,068 records remained. Screening of titles and abstracts resulted in 57 studies, of which 43 were excluded after full-text review (see Figure 1 for details on exclusion reasons). Ultimately, 14 studies published between 2012 and 2023 were included in the analysis (29–42); one study was included in the qualitative analysis (42), and the remaining 13 were included in the quantitative analysis, evaluating 988 patients for lung ultrasound scores.

Baseline characteristics of included studies

The baseline characteristics of the included studies are presented in Tables 1, 2; Supplementary Table 3. All the studies were prospective cohort studies. Among the 14 studies, 4 were conducted in Egypt (29, 31, 34, 41), 2 in France (30, 39), 2 in the United States (33, 35), 2 in China (38, 40), and the remaining 4 in Spain (36), Brazil (42), Nepal (37), and India (32).

Most of the studies were conducted in general ICUs, but three were carried out in medical and surgical ICUs (29, 36, 37), and one in an emergency ICU (35). Thirteen studies performed a quantitative assessment of lung ultrasound and provided scores, while one study used a qualitative assessment (42). Ten of the quantitative studies restricted the scans to the anterior, posterior, and lateral lung regions, defining the lung ultrasound score range as 0–36. Three additional studies restricted the scans to the anterior and lateral lung regions, with a lung ultrasound score range of 0–24 (33, 35, 36). Further details regarding lung ultrasound score definitions and LUS cut-offs for each study can be found in Supplementary Table 3.

The definition of "failure to extubate" varied among the studies and included one or more of the following: inability to maintain spontaneous respiration without ventilatory support for more than 48 h, the need for noninvasive ventilation, high-flow nasal cannula, reintubation, terminal extubation, tracheostomy, delayed extubation, and/or failure of the spontaneous breathing trial (SBT). Patient inclusion criteria and modes of extubation also differed across studies, as detailed in Tables 1, 2.

FIGURE 1
Flow diagram of the literature and selection in the meta-analysis.
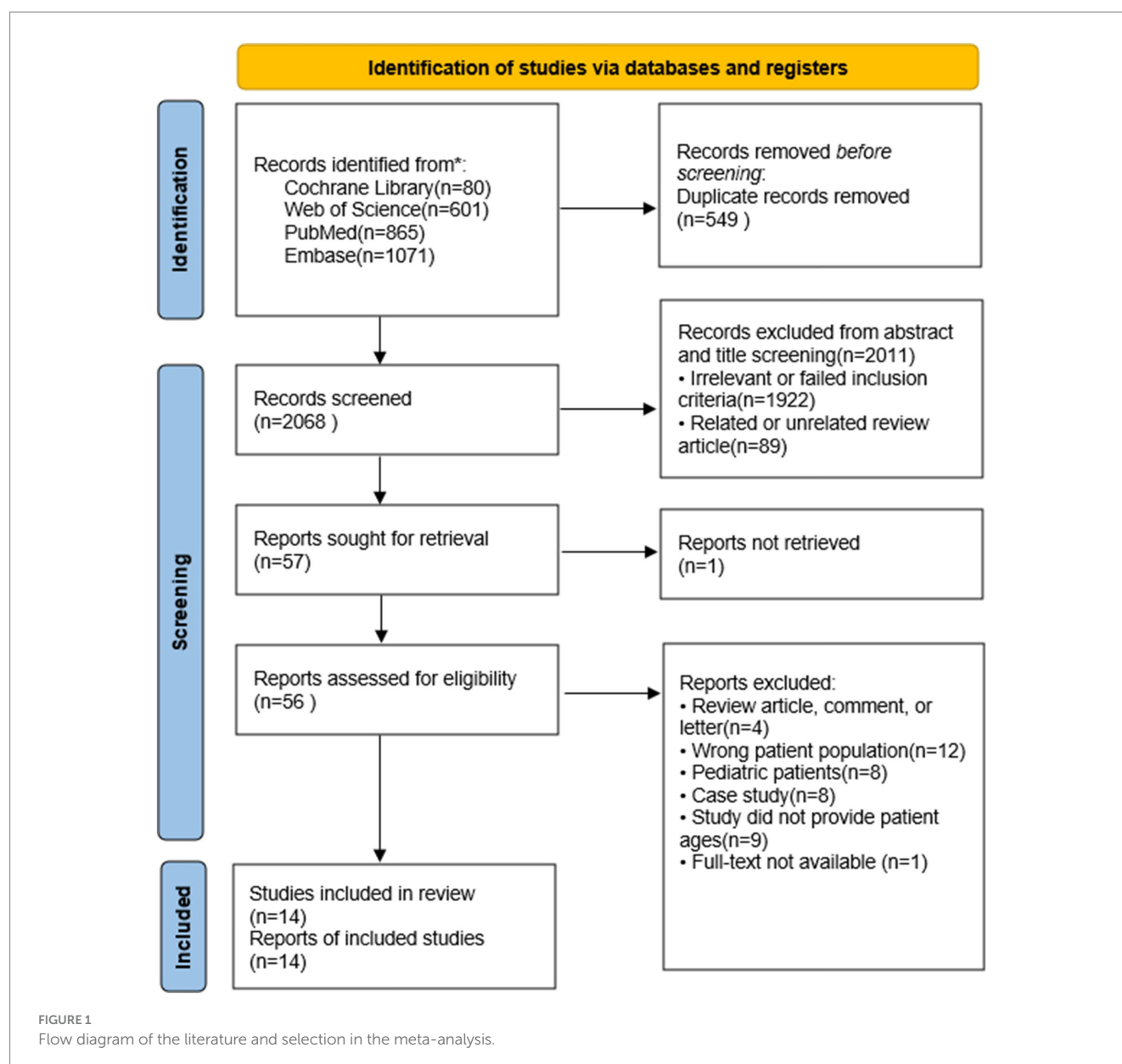
Assessment of methodologic quality

The results of the Cochrane risk of bias assessment for the studies included in this systematic review are shown in Figure 2. Most studies demonstrated an unclear or high risk of bias regarding the index test and reference standards. The majority of the studies [except three (32, 39, 40)] did not report whether the lung ultrasound (index test) was interpreted independently of the extubation outcomes. Nearly half of the studies did not disclose the identity of the ultrasound operators. Lung ultrasounds were typically performed before or during the spontaneous breathing trials (SBT), but two studies (39, 40) only considered individuals who successfully passed these tests (see Table 2). Most studies [except for three (30, 39, 40)] provided varying definitions of extubation outcomes (reference standard), and these definitions were not always adequately described. One study defined extubation failure solely as a failure in the spontaneous breathing trial, which presented a high risk of bias.

Diagnostic accuracy and heterogeneity

The meta-analysis included 13 studies, excluding one study due to its qualitative design, which involved a different range of ultrasound scans and lacked a quantitative score.

Figure 3 displays the forest plots for the sensitivity and specificity of ultrasound scores across the included studies. Figure 4 shows the summary receiver operating characteristic (SROC) curves, illustrating the pooled points and 95% confidence intervals for the sensitivity and specificity of the lung ultrasound score (LUS) (AUC = 0.87, 95% CI 0.84–0.89). The pooled sensitivity and specificity for LUS were 0.86 (95% CI 0.77–0.91) and 0.75 (95% CI 0.66–0.83), respectively. Pooled diagnostic odds ratios and likelihood ratios are presented in Supplementary Figure 1. Heterogeneity in sensitivity and specificity for LUS was significant, with the Q-test = 60.68, $p < 0.01$, $I^2 = 80.23\%$ for sensitivity, and Q-test = 34.21, $p < 0.01$, $I^2 = 64.92\%$ for specificity.

TABLE 1 Characteristics of included studies.

Study/year	Country	Design	Definition of weaning failure	Ultrasound probe frequency
Osman et al./2017	Egypt	Prospective observational	Reintubation within 48 h	3.5 MHz
Binet et al./2014	France	Prospective observational	Reintubation, nonscheduled NIMV, or death within 48 h	NR
Soliman et al./2019	Egypt	Prospective observational	Failed SBT or the need for reintubation within 48 h following extubation	2–4 MHz
Banerjee et al./2018	India	Prospective observational	Reintubation within 48 h	NR
Gok et al./2021	America	Prospective observational	Reintubated within 48 h, need for noninvasive mechanical ventilation	2–4 MHz
Shoaeir et al./2016	Egypt	Prospective observational	Need for reintubation and mechanical ventilation, need for noninvasive positive pressure ventilation	2–4 MHz
Gosai et al./2021	America	Prospective observational	Reintubation and mechanical ventilation after 48 h	3.5–6.5 MHz
Tenza-Lozano et al./2018	Spain	Prospective observational	The occurrence of reintubation, non-invasive ventilatory support or death within 48 h following extubation.	2–4 MHz
Antonio et al./2018	Brazil	Prospective observational	An inability to tolerate a T-piece SBT of 30–120 min	2–4 MHz
Rajbanshi et al./2023	Nepal	Prospective observational	Require any form of invasive or noninvasive respiratory support within 48 h of extubation	2–4 MHz
Wang et al./2022	China	Prospective observational	Failed SBT or the need for reintubation within 48 h following extubation	3.5 MHz
Soummer et al./2012	France	Prospective observational	Ventilatory support (either noninvasive or invasive ventilation) within 48 h after extubation	2–4 MHz
Xu et al./2020	China	Prospective observational	The requirement for NPPV or reintubation within 48 h	NR
Gu et al./2020	Egypt	Prospective observational	Reintubation and mechanical ventilation after 48 h	2–4 MHz

NIMV, noninvasive mechanical ventilation; SBT, spontaneous breathing trial; NPPV, noninvasive positive-pressure ventilation.

Fagan plots were used to demonstrate the shift in probabilities of extubation failure based on lung ultrasound outcomes (Supplementary Figure 2). The initial pre-test probability was set at 50%, representing the median risk in similar clinical populations and typical uncertainties in extubation decisions. For patients with a positive LUS result (exceeding the cut-off value), the post-test probability of extubation failure increased to 78%, indicating a high likelihood of needing continued mechanical support. Conversely, a negative LUS result (below the cut-off) reduced the post-test probability to 16%, suggesting a lower risk and potentially safer conditions for extubation. These probability adjustments were based on likelihood ratios derived from the sensitivity and specificity values of LUS in predicting extubation failure.

Meta-regression analysis

The results of the meta-regression analysis for LUS sensitivity and specificity are shown in Supplementary Figures 3–6. Factors such as thresholds, risk of bias in article quality, reference

standards, and study flow and timing, as identified in the QUADAS-2 assessment, were analyzed as potential sources of heterogeneity in the diagnostic accuracy of lung ultrasound. A threshold of >15 points significantly increased the sensitivity of LUS in predicting extubation failure ($p = 0.02$). Studies with low risk in QUADAS-2 domains were categorized as high quality. However, there was no significant difference in diagnostic accuracy between high-and low-quality studies. A low risk of bias in the reference standard (extubation outcomes) significantly increased the sensitivity of LUS ($p = 0.01$), while a low risk of bias in patient flow and timing significantly increased the specificity of LUS ($p = 0.01$).

Sensitivity analysis and publication bias

Sensitivity analysis, including impact analysis and outlier detection, identified two outliers for LUS (see Supplementary Figure 7). Excluding these outliers in subgroup analysis resulted in a slightly reduced sensitivity (0.81, 95% CI

TABLE 2 Participant characteristics.

Study	Patients	Ageᵃ	Setting	Inclusion	Type of SBT
Osman et al.	68	56 (45–65)	MICU SICU	Patients ready for SBT	T-tube
Binet et al.	48	59 ± 16	MD ICU	Ventilated >48 h	PS
Soliman et al.	100	57 ± 14	ICU	Ventilated >48 h	PS
Banerjee et al.	53	Female 55.52 ± 6.39 Male 55.07 ± 5.2	ICU	Patients ready for SBT	T-tube/PS
Gok et al.	62	57.6 ± 14.1	ICU	Ventilated >48 h	T-tube/PS
Shoaeir et al.	50	SG 47.52 ± 14.60 FG 51.89 ± 14.58	ICU	Patients ventilated >48 h and ready for a first SBT	NR
Gosai et al.	100	62 (26–72)	EICU	Patients ready for SBT	NR
Tenza-Lozano et al.	69	66 (53, 78)	ICU	Ventilated >24 h; ready for weaning	T-tube/PS
Antonio et al.	250	SG 75 (60–83) FG 66 (47–81)	MICU SICU	Undergone invasive mechanical ventilation for 24 h	NR
Rajbanshi et al.	102	43.52 ± 19.50	MICU SICU	Mechanically ventilated for a duration exceeding 24 h and were considered ready for the weaning process	NR
Wang et al.	92	SG 60.25 ± 13.64 FG 66.92 ± 14.53	ICU	Patients who met the diagnostic standard of ARDS (the Berlin definition). Patients ready for SBT	PS
Soummer et al.	86	60 ± 15	2 MD ICU	Patients ventilated >48 h who successfully passed a first SBT	T-tube
Xu et al.	105	57.98 ± 15.20	ICU	Successful SBT and considered ready for extubation were recruited	NR
Gu et al.	53	68 (53–77)	ICU	Acute moderate and severe respiratory failure, Patients ventilated >48 h and ready for a first SBT	T-tube

ICU, intensive care unit; MICU, medical intensive care unit; SICU, surgical intensive care unit; MD, multidisciplinary; EICU, emergency intensive care unit; SG, success group; FG, failure group; SBT, spontaneous breathing trial; ARDS, acute respiratory distress syndrome; PS, low pressure support trial, NR, not reported.
ᵃAge is expressed according to data extracted from each study as mean ± standard deviation or median (interquartile range).

0.74–0.86) and an almost unchanged specificity (0.75, 95% CI 0.66–0.82). The fitness and bivariate normality analyses confirmed the moderate robustness of the random-effects bivariate model used in this meta-analysis.

The Deeks funnel plot test did not detect significant publication bias for LUS (see Supplementary Figure 8).

Discussion

This systematic review and meta-analysis aimed to assess the diagnostic accuracy of lung ultrasound (LUS) scoring in predicting extubation failure in mechanically ventilated, critically ill patients. The LUS score has recently gained recognition as a simple and reliable method, particularly for diagnosing lung disease in the intensive care unit (ICU) setting (43). LUS provides a wealth of information about morphological changes in the lungs, aiding physicians in assessing the progression of lung disease by observing ultrasound signs such as abnormal pleural lines, increased B-lines, and solid lung lesions. The total LUS score in this study reflects the severity of ventilation loss in both lungs (44).

The LUS score, known for integrating respiratory and cardiac assessments, functions as a composite tool rather than merely a diagnostic indicator, offering a comprehensive view of patient status. This dual utility is especially crucial when diastolic dysfunction alters LUS scores post-extubation, transitioning from a low to high score. These findings underscore the importance of nuanced interpretation of LUS score changes over time.

Our results support the utility of LUS as a valuable predictor of extubation failure, with a pooled sensitivity of 0.86, specificity of 0.75, and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.87, indicating high reliability. These results align with those
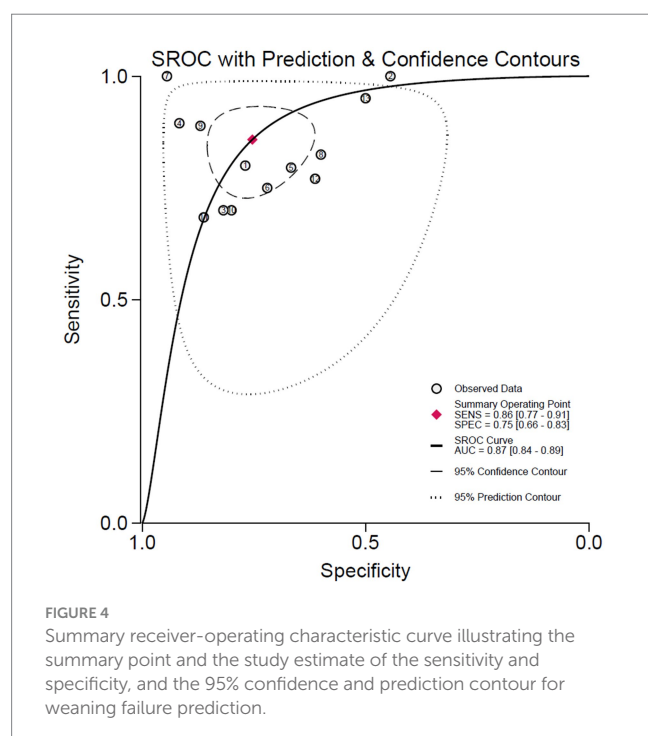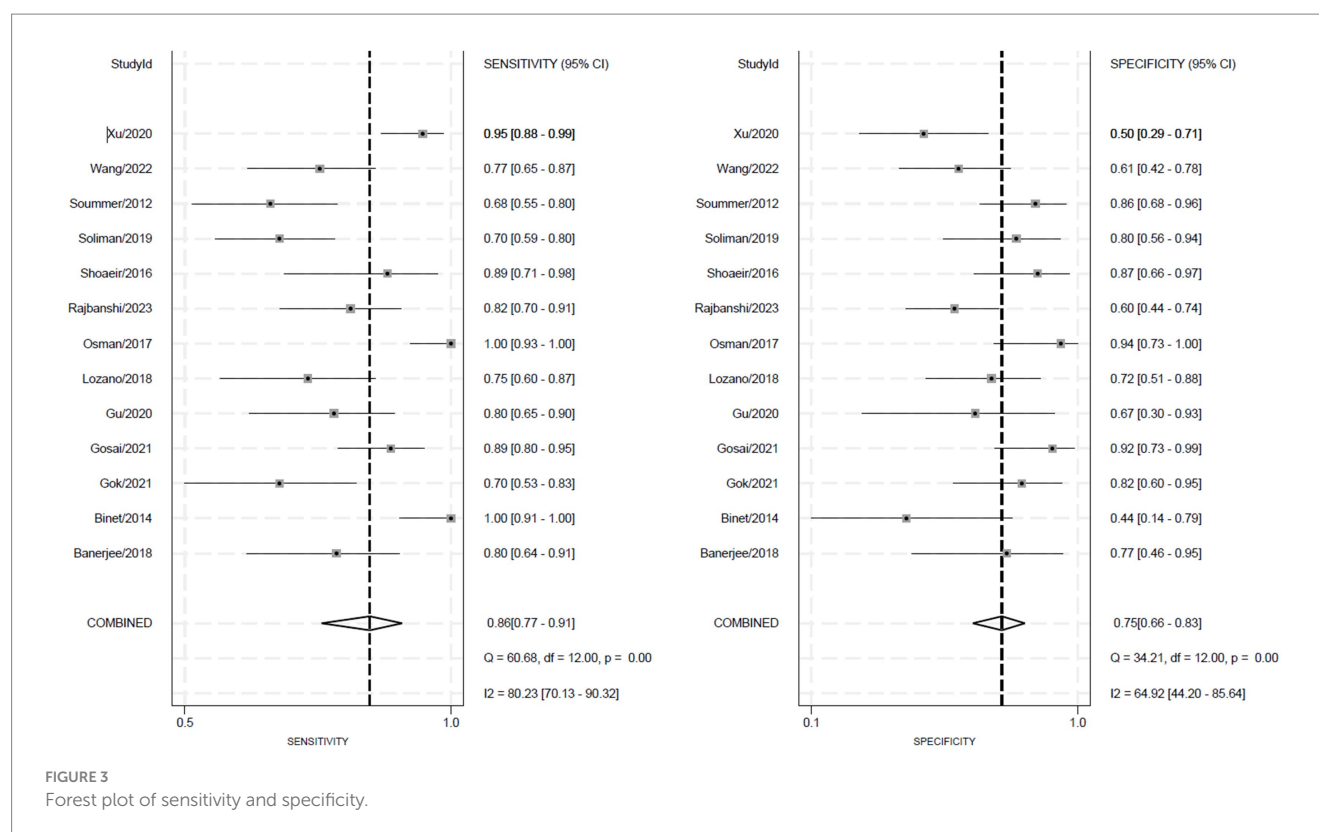
FIGURE 2
Risk of bias and applicability concerns assessment with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS).

from Mohsen's et al. (45) neonatal systematic evaluation (AUC of 0.87) and are consistent with findings by Biasucci et al. (46), who also noted that LUS is an accurate and useful diagnostic tool across critically ill patients of all ages. These consistent findings reinforce the value of LUS in predicting extubation failure in mechanically ventilated, critically ill patients.

Banerjee et al. (32) reported that patients with LUS scores between 1 and 10 had a high rate of successful extubation, while those with scores between 16 and 32 had a higher rate of extubation failure. This is consistent with our meta-regression analysis, which showed that using a threshold greater than 15 points significantly increased the sensitivity of LUS in predicting extubation failure. These results

highlight the importance of selecting appropriate LUS thresholds in clinical practice, aligning with findings from Amara et al. (47), who observed significant differences in LUS between simple and complex extubation scenarios, further demonstrating the utility of LUS in nuanced clinical assessments.

Additionally, the rapid acquisition and ease of use of LUS make it a time-efficient tool that, when used alongside other clinical assessments, significantly enhances the accuracy and speed of decision-making. This efficiency is particularly valuable in high-pressure ICU environments, such as those managing COVID-19 patients. For example, Meroi et al. (48) demonstrated that in COVID-19 ICUs, LUS allowed for significant time savings compared

FIGURE 3
Forest plot of sensitivity and specificity.

FIGURE 4
Summary receiver-operating characteristic curve illustrating the summary point and the study estimate of the sensitivity and specificity, and the 95% confidence and prediction contour for weaning failure prediction.

to traditional radiology, with assessments taking approximately 4.2 min per patient. This rapid evaluation enables timely monitoring of pulmonary conditions, minimizing exposure to healthcare workers and conserving critical resources. Furthermore, extubation failure is not always related to lung pathology but can also result from upper airway issues, such as laryngeal edema, which may require additional

evaluation methods like the leak test. Moreover, LUS has been shown to reduce healthcare costs by decreasing the need for patient transport to radiology for chest CT scans and reducing X-ray exposure (49). The cost savings from fewer diagnostic imaging procedures and complications during patient transfers position LUS as a cost-effective alternative in respiratory care.

Previous systematic reviews (17) have noted that inconsistencies in defining extubation failure across studies may affect the accuracy of diaphragm ultrasound in predicting extubation outcomes. This issue also applies to using LUS for predicting extubation failure. Despite focusing on different anatomical areas, both diaphragm ultrasound and LUS have potential value in assessing readiness for extubation. Therefore, our findings emphasize the need for strict and consistent criteria when using LUS to predict extubation failure. Standardizing definitions of extubation failure will improve the sensitivity and specificity of predictions, suggesting that a uniform and standardized definition is essential for improving prediction accuracy in both diaphragmatic ultrasound and LUS-based assessments.

Our study also found that reducing bias in patient flow and timing significantly increased the specificity of LUS. This underscores the importance of reducing bias in study design and implementation to improve the accuracy of LUS in predicting successful extubation. Variations in patient inclusion and exclusion criteria, as well as inconsistencies in ultrasound operators and operation times, introduced potential biases that may have affected the specificity of LUS. Thus, conducting rigorous and standardized studies is crucial to ensuring the accuracy and reliability of LUS scores in predicting extubation failure. This includes clearly defining patient inclusion and exclusion criteria, standardizing ultrasound procedures, and controlling for potential confounders during analysis. By adhering to rigorously designed studies, we can ensure the validity and reliability

of LUS scores in predicting extubation failure in mechanically ventilated, critically ill patients, ultimately improving patient outcomes.

Strengths and limitations

In recent years, significant attention has been given to the diagnostic accuracy of lung ultrasound (LUS) scores in predicting extubation failure in mechanically ventilated, critically ill patients. Compared to earlier studies from Llamas-Álvarez et al. (19), this systematic review added nine new relevant studies, expanding the meta-analysis to include more uniform and clinically relevant data. This enhancement increases the reliability and representativeness of the results. Additionally, we thoroughly explored the sources of sensitivity and specificity heterogeneity in LUS scores. Through meta-regression analyses, we identified potential factors influencing the accuracy of LUS, providing valuable guidance for its application in different clinical settings and future studies.

However, there are some limitations to this systematic review and meta-analysis. First, unpublished data were excluded to maintain study quality, which may have resulted in the omission of relevant studies, potentially impacting the completeness of the findings. Despite our in-depth analysis of heterogeneity, significant differences remained between studies regarding patient populations, definitions of extubation failure, and the application of ultrasound scoring systems. These differences could affect the consistency and generalizability of the results. Moreover, some studies did not specify the identity of the ultrasound operators, which could introduce operator bias and influence the accuracy and reliability of LUS scoring.

Despite these limitations, our analysis indicates that the LUS score is a valid tool for predicting extubation failure in critically ill patients on mechanical ventilation. With further research and standardization of methods, the LUS score has the potential to become an important auxiliary diagnostic tool in clinical practice. Future research should focus on improving study design and implementation, as well as standardizing the definition of extubation failure, to enhance the accuracy and reliability of LUS in predicting extubation outcomes. Additionally, studies should explore the impact of different cut-off values on the predictive ability of LUS scores and validate its application in various clinical settings.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: this study is a meta-analysis, and the data analyzed were derived from previously published literature. Detailed references are provided within the article, and there are no direct links to unpublished original datasets involved.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1486636/full#supplementary-material

References

1. Vetrugno L, Guadagnin GM, Brussa A, Orso D, Garofalo E, Bruni A, et al. Mechanical ventilation weaning issues can be counted on the fingers of just one hand: part 1. *Ultrasound J.* (2020) 12:9. doi: 10.1186/s13089-020-00161-y

2. Levine S, Nguyen T, Taylor N, Friscia ME, Budak MT, Rothenberg P, et al. Rapid disuse atrophy of diaphragm fibers in mechanically ventilated humans. *N Engl J Med.* (2008) 358:1327–35. doi: 10.1056/NEJMoa070447

3. Peñuelas O, Frutos-Vivar F, Fernández C, Anzueto A, Epstein SK, Apezteguía C, et al. Characteristics and outcomes of ventilated patients according to time to liberation from mechanical ventilation. *Am J Respir Crit Care Med.* (2011) 184:430–7. doi: 10.1164/rccm.201011-1887OC

4. Powers SK, Kavazis AN, Levine S. Prolonged mechanical ventilation alters diaphragmatic structure and function. *Crit Care Med.* (2009) 37:S347–53. doi: 10.1097/CCM.0b013e3181b6e760

5. Vetrugno L, Castaldo N, Fantin A, Deana C, Cortegiani A, Longhini F, et al. Ventilatory associated barotrauma in COVID-19 patients: a multicenter observational case control study (COVI-MIX-study). *Pulmonology.* (2023) 29:457–68. doi: 10.1016/j.pulmoe.2022.11.002

6. Caroleo S, Agnello F, Abdallah K, Santangelo E, Amantea B. Weaning from mechanical ventilation: an open issue. *Minerva Anestesiol.* (2007) 73:417–27.

7. Funk GC, Anders S, Breyer MK, Burghuber OC, Edelmann G, Heindl W, et al. Incidence and outcome of weaning from mechanical ventilation according to new categories. *Eur Respir J*. (2010) 35:88–94. doi: 10.1183/09031936.00056909

8. Lemaire F, Teboul JL, Cinotti L, Giotto G, Abrouk F, Steg G, et al. Acute left ventricular dysfunction during unsuccessful weaning from mechanical ventilation. *Anesthesiology*. (1988) 69:171–9. doi: 10.1097/00000542-198808000-00004

9. Jubran A, Lawm G, Duffner LA, Collins EG, Lanuza DM, Hoffman LA, et al. Post-traumatic stress disorder after weaning from prolonged mechanical ventilation. *Intensive Care Med*. (2010) 36:2030–7. doi: 10.1007/s00134-010-1972-8

10. Esteban A, Anzueto A, Frutos F, Alía I, Brochard L, Stewart TE, et al. Characteristics and outcomes in adult patients receiving mechanical ventilation: a 28-day international study. *JAMA*. (2002) 287:345–55. doi: 10.1001/jama.287.3.345

11. Yang KL, Tobin MJ. A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation. *N Engl J Med*. (1991) 324:1445–50. doi: 10.1056/NEJM199105233242101

12. Lichtenstein D. Lung ultrasound in the critically ill. *Curr Opin Crit Care*. (2014) 20:315–22. doi: 10.1097/MCC.0000000000000096

13. 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. doi: 10.1007/s00134-012-2513-4

14. Vetrugno L, Biasucci DG, Deana C, Spadaro S, Lombardi FA, Longhini F, et al. Lung ultrasound and supine chest X-ray use in modern adult intensive care: mapping 30 years of advancement (1993–2023). *Ultrasound J*. (2024) 16:7. doi: 10.1186/s13089-023-00351-4

15. Dres M, Goligher EC, Dubé BP, Morawiec E, Dangers L, Reuter D, et al. Diaphragm function and weaning from mechanical ventilation: an ultrasound and phrenic nerve stimulation clinical study. *Ann Intensive Care*. (2018) 8:53. doi: 10.1186/s13613-018-0401-y

16. Vassilakopoulos T, Petrof BJ. A stimulating approach to ventilator-induced diaphragmatic dysfunction. *Am J Respir Crit Care Med*. (2017) 195:281–2. doi: 10.1164/rccm.201608-1619ED

17. Le Neindre A, Philippart F, Luperto M, Wormser J, Morel-Sapene J, Aho SL, et al. Diagnostic accuracy of diaphragm ultrasound to predict weaning outcome: a systematic review and meta-analysis. *Int J Nurs Stud*. (2021) 117:103890. doi: 10.1016/j.ijnurstu.2021.103890

18. Li C, Li X, Han H, Cui H, Wang G, Wang Z. Diaphragmatic ultrasonography for predicting ventilator weaning: a meta-analysis. *Medicine (Baltimore)*. (2018) 97:e10968. doi: 10.1097/MD.0000000000010968

19. Llamas-Álvarez AM, Tenza-Lozano EM, Latour-Pérez J. Diaphragm and lung ultrasound to predict weaning outcome: systematic review and Meta-analysis. *Chest*. (2017) 152:1140–50. doi: 10.1016/j.chest.2017.08.028

20. Mahmoodpoor A, Fouladi S, Ramouz A, Shadvar K, Ostadi Z, Soleimanpour H. Diaphragm ultrasound to predict weaning outcome: systematic review and meta-analysis. *Anaesthesiol Intensive Ther*. (2022) 54:164–74. doi: 10.5114/ait.2022.117273

21. Bouhemad B, Mongodi S, Via G, Rouquette I. Ultrasound for "lung monitoring" of ventilated patients. *Anesthesiology*. (2015) 122:437–47. doi: 10.1097/ALN.0000000000000558

22. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. (2009) 6:e1000097. doi: 10.1371/journal.pmed.1000097

23. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. (2011) 155:529–36. doi: 10.7326/0003-4819-155-8-201110180-00009

24. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. (2005) 58:882–93. doi: 10.1016/j.jclinepi.2005.01.016

25. Harbord R, Whiting P, Sterne J. Meta-analysis in Stata: an updated collection from the Stata journal. *2nd* ed. College Station, TX: Strata Press (2009).

26. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. (2005) 58:982–90. doi: 10.1016/j.jclinepi.2005.02.022

27. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. (2007) 8:239–51. doi: 10.1093/biostatistics/kxl004

28. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Assess*. (2005) 9:1–113. doi: 10.3310/hta9120

29. Osman AM, Hashim RM. Diaphragmatic and lung ultrasound application as new predictive indices for the weaning process in ICU patients. *Egypt J Radiol Nucl Med*. (2017) 48:61–6. doi: 10.1016/j.ejrnm.2017.01.005

30. Binet C, Neron L, Rochart N, Cousson J, Floch T, Charbit B, et al. Validation d'un indice échographique prédictif d'échec de sevrage de la ventilation mécanique. *Ann Fr Anesth Reanim*. (2014) 33:A383. doi: 10.1016/j.annfar.2014.07.653

31. Soliman SB, Ragab F, Soliman RA, Gaber A, Kamal A. Chest ultrasound in predication of weaning failure. *Open Access Maced J Med Sci*. (2019) 7:1143–7. doi: 10.3889/oamjms.2019.277

32. Banerjee A, Mehrotra G. Comparison of lung ultrasound-based weaning indices with rapid shallow breathing index: are they helpful? *Indian J Crit Care Med*. (2018) 22:435–40. doi: 10.4103/ijccm.IJCCM_331_17

33. Gok F, Mercan A, Kilicaslan A, Sarkilar G, Yosunkaya A. Diaphragm and lung ultrasonography during weaning from mechanical ventilation in critically ill patients. *Cureus*. (2021) 13:e15057. doi: 10.7759/cureus.15057

34. Shoaeir M, Noeam K, Mahrous A, Alaa A. Lung aeration loss as a predictor of reintubation using lung ultrasound in mechanically ventilated patients. *Biolife*. (2016) 4:514–20. doi: 10.5281/zenodo.7332900

35. Gosai JB, Sangani SV, Patel CJ, Meghraj MA, Agarwal R, Parmar R. Lung and diaphragm ultrasound as a predictor of successful weaning from mechanical ventilation: a prospective observational study of 100 ICU patients. *J Cardiovasc Dis Res*. (2021) 12:642–9.

36. Tenza-Lozano E, Llamas-Alvarez A, Jaimez-Navarro E, Fernández-Sánchez J. Lung and diaphragm ultrasound as predictors of success in weaning from mechanical ventilation. *Crit Ultrasound J*. (2018) 10:12. doi: 10.1186/s13089-018-0094-3

37. Rajbanshi LK, Bajracharya A, Devkota D. Prediction of successful spontaneous breathing trial and extubation of trachea by lung ultrasound in mechanically ventilated patients in intensive care unit. *Indian J Crit Care Med*. (2023) 27:482–7. doi: 10.5005/jp-journals-10071-24487

38. Wang R, Qi B, Zhang X, Meng L, Wu X. Prophetic values of lung ultrasound score on post-extubation distress in patients with acute respiratory distress syndrome. *Eur J Med Res*. (2022) 27:27. doi: 10.1186/s40001-022-00652-9

39. Soummer A, Perbet S, Brisson H, Arbelot C, Constantin JM, Lu Q, et al. Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress. *Crit Care Med*. (2012) 40:2064–72. doi: 10.1097/CCM.0b013e31824e68ae

40. Xu X, Wu R, Zhang YJ, Li HW, He XH, Wang SM. Value of combination of heart, lung, and diaphragm ultrasound in predicting weaning outcome of mechanical ventilation. *Med Sci Monit*. (2020) 26:e924885. doi: 10.12659/MSM.924885

41. Gu C, Xie Y, Zheng T, Ren H, Wang G, Ren J. Predictive value of lung ultrasound score on weaning outcome in patients with intro-abdominal infection undergoing mechanical ventilation. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue*. (2020) 32:94–8. doi: 10.3760/cma.j.cn121430-20191127-00017

42. Antonio ACP, Knorst MM, Teixeira C. Lung ultrasound prior to spontaneous breathing trial is not helpful in the decision to wean. *Respir Care*. (2018) 63:873–8. doi: 10.4187/respcare.05817

43. Eman G, Marjan I, Shwe S, Aluko AH. Thoracic ultrasound in COVID-19: a prospective study using lung and diaphragm ultrasound in evaluating dyspnea in ICU survivors with COVID-19 in a post-ICU clinic. *Chest*. (2022) 161:A198. doi: 10.1016/j.chest.2021.12.230

44. Lichtenstein DA, Mezière GA. Relevance of lung ultrasound in the diagnosis of acute respiratory failure: the BLUE protocol. *Chest*. (2008) 134:117–25. doi: 10.1378/chest.07-2800

45. Mohsen N, Solis-Garcia G, Jasani B, Nasef N, Mohamed A. Accuracy of lung ultrasound in predicting extubation failure in neonates: a systematic review and meta-analysis. *Pediatr Pulmonol*. (2023) 58:2846–56. doi: 10.1002/ppul.26598

46. Biasucci DG, Loi B, Centorrino R, Raschetti R, Piastra M, Pisapia L, et al. Ultrasound-assessed lung aeration correlates with respiratory system compliance in adults and neonates with acute hypoxemic restrictive respiratory failure: an observational prospective study. *Respir Res*. (2022) 23:360. doi: 10.1186/s12931-022-02294-1

47. Amara V, Vishwas P, Maddani SS, Natarajan S, Chaudhuri S. Evaluation of abdominal expiratory muscle thickness pattern, diaphragmatic excursion, and lung ultrasound score in critically ill patients and their association with weaning patterns: a prospective observational study. *Indian J Crit Care Med*. (2022) 26:307–13. doi: 10.5005/jp-journals-10071-24125

48. Meroi F, Orso D, Vetrugno L, Bove T. Lung ultrasound score in critically ill COVID-19 patients: a waste of time or a time-saving tool? *Acad Radiol*. (2021) 28:1323–4. doi: 10.1016/j.acra.2021.06.008

49. Cammarota G, Vetrugno L, Longhini F. Lung ultrasound monitoring: impact on economics and outcomes. *Curr Opin Anaesthesiol*. (2023) 36:234–9. doi: 10.1097/ACO.0000000000001231