CHEST

The Predictive Value of Lung Ultrasound Scores in Developing Bronchopulmonary Dysplasia

A Prospective Multicenter Diagnostic Accuracy Study

Almudena Alonso-Ojembarrena, MD, PhD; Iker Serna-Guerediaga, MD; Victoria Aldecoa-Bilbao, MD; Rebeca Gregorio-Hernández, MD, PhD; Paula Alonso-Quintela, MD, PhD; Ana Concheiro-Guisán, MD, PhD; Alicia Ramos-Rodríguez, MD; Mónica de las Heras-Martín, MD; Lorena Rodeño-Fernández, MD; and Ignacio Oulego-Erroz, MD, PhD

BACKGROUND: Different lung ultrasound (LUS) scanning protocols have been used, and the results in terms of diagnostic accuracy are heterogeneous.

RESEARCH QUESTIONS: What is the diagnostic accuracy of the LUS score to predict moderate to severe bronchopulmonary dysplasia (msBPD)? Does scanning of posterior lung fields improve the diagnostic accuracy?

STUDY DESIGN AND METHODS: This was a multicenter prospective, observational study in six centers. Two LUS aeration scores, one involving only anterolateral lung fields and the other adding the posterior fields were obtained at birth, on the third day of life (DOL), on the seventh DOL, on the 14th DOL, and on the 21st DOL. The diagnostic accuracy of both scores to predict msBPD was assessed at each time point.

RESULTS: Eight hundred thirty-two LUS examinations in 298 infants were included. Both LUS score using anterolateral and posterior fields and LUS score using only anterolateral fields showed a similar moderate diagnostic accuracy to predict msBPD on the third DOL (area under the receiver operating characteristic curve [AUC] 95% CI, 0.68-0.85 vs 0.68-0.85; $P = .97$), seventh DOL (AUC 95% CI, 0.74-0.85 vs 0.74-0.84; $P = .26$), and 21st DOL (AUC 95% CI, 0.72-0.86 vs 0.74-0.88; $P = .17$). The LUS score using anterolateral and posterior fields was slightly more accurate at 14th DOL (AUC 95% CI, 0.69-0.83 vs 0.66-0.80; $P = .01$). A cutoff of 8 points in the LUS score using only anterolateral fields on the seventh DOL provided a sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio of 70%, 79%, 3.3, and 0.38, respectively, to predict msBPD. Adding gestational age (GA) and sex improved the discriminative value without significant differences compared with a predictive model based on multiple clinical variables: AUC 95% CI, 0.77-0.88 vs 0.80-0.91 ($P = .52$).

INTERPRETATION: The LUS score is able to predict msBPD from the third DOL with a moderate diagnostic accuracy. Scanning posterior lung fields slightly improved diagnostic accuracy only at the 14th DOL. Adding GA and sex improves the diagnostic accuracy of the LUS scores. The LUS score is useful to stratify BPD risk early after birth.

CHEST 2021; 160(3):1006-1016

KEY WORDS: biomarker; bronchopulmonary dysplasia; lung diagnostic imaging; preterm; ultrasound

FOR EDITORIAL COMMENT, SEE PAGE 799

ABBREVIATIONS: AUC = area under the receiver operating characteristic curve; BPD = bronchopulmonary dysplasia; DOL = day of life; GA = gestational age; IQR = interquartile range; LUS = lung ultrasound; msBPD = moderate to severe bronchopulmonary dysplasia; RDS = respiratory distress syndrome

AFFILIATIONS: From the Neonatal Intensive Care Unit (A. Alonso-Ojembarrena), the Research Unit (A. Alonso-Ojembarrena), Biomedical Research and Innovation Institute of Cádiz (INiBICA), the Pediatric Department (A. Ramos-Rodríguez), Puerta del Mar University Hospital, Cádiz, the Neonatal Intensive Care Unit

Take-home Points

Study Question: Does scanning posterior lung fields improve lung ultrasound (LUS) diagnostic accuracy in moderate to severe bronchopulmonary dysplasia (msBPD)?

Results: LUS shows similar diagnostic accuracy in msBPD using anterolateral only fields (AUC at seventh day of life, 0.80 [95% CI, 0.74-0.85]) or adding posterior fields (AUC at seventh day of life, 0.79 [95% CI, 0.74-0.84]).

Interpretation: LUS is useful to stratify msBPD risk early after birth, even though we study only anterolateral fields.

Bronchopulmonary dysplasia (BPD) is the most common sequelae of prematurity.[1] BPD still causes significant mortality and has a negative impact on lung function and quality of life, which extends to adulthood. BPD is diagnosed at 36 weeks' postmenstrual age, when structural changes may be irreversible. Finding early biomarkers of developing BPD is needed to stratify individual risk soon after birth and to implement preventive and therapeutic strategies when clinicians can still alter the pathologic process.[2,3] A myriad of biochemical biomarkers have been studied with this purpose.[4,5] However, diagnostic accuracy has been only modest, and most of them remain research tools not available in daily practice.

(I. Serna-Guerediaga, M. de las Heras-Martín, and L. Rodeño-Fernández), Basurto Universitary Hospital, Bilbao, the Neonatology Department (V. Aldecoa-Bilbao), Hospital Clínic Barcelona, BCNatal—Barcelona Center for Maternal-Fetal and Neonatal Medicine, Barcelona, the Neonatal Intensive Care Unit (R. Gregorio-Hernández), Virgen de la Salud Hospital, Toledo, the Neonatal Intensive Care Unit (P. Alonso-Quintela), Complejo Asistencial Universitario de León, the Biomedicine Institute of León (IBIOMED) (P. Alonso-Quintela and I. Oulego-Erroz), University of León, the Pediatric Intensive Care Unit (I. Oulego-Erroz), Complejo Asistencial Universitario de León, León, the Neonatal Intensive Care Unit (A. Concheiro-Guisán), Alvaro Cunqueiro Hospital, the Sanitary Investigation Institute Galicia Sur (IIS Galicia Sur) (A. Concheiro-Guisán), Vigo, and the Working Group on Bedside Ultrasound of the Spanish Society of Pediatric Intensive Care (SECIP) (I. Oulego-Erroz), Madrid, Spain.

Presented at the 5th Annual Meeting of the Spanish Society of Neonatology (SENEO), October 15, 2020, Virtual conference (https://jornadaneonatologia2020.com/).

FUNDING/SUPPORT: The authors have reported to *CHEST* that no funding was received for this study.

CORRESPONDENCE TO: Almudena Alonso-Ojembarrena, MD, PhD; email: almudena.alonso.sspa@juntadeandalucia.es

DOI: https://doi.org/10.1016/j.chest.2021.02.066

Lung ultrasound (LUS) is a widely available, inexpensive, and useful imaging method to assess neonatal respiratory disease. Its main advantages are high interobserver agreement[6] and high reliability and sensitivity. LUS facilitates the differential diagnosis of acute neonatal respiratory failure[7,8] and may aid in clinical management. LUS represents a change in clinical practice and may reduce radiation exposure in neonates.[9,10] LUS findings in BPD are a thickened pleural line with subjacent scattered small consolidations and a nonhomogeneous interstitial syndrome that varies from coalescent B-lines interposed with spared areas to a so-called white lung pattern.[11]

Beyond its use in the acute setting, the potential of LUS to predict outcomes is a promising field of research. Based on the ability of LUS to evaluate the regional distribution of lung aeration, the LUS score has been developed. The LUS score accurately predicts failure of noninvasive ventilation in preterm infants with respiratory distress syndrome (RDS) and shortens the time to surfactant administration.[12-15] This LUS score also has been used to predict the development of BPD.[16-18] However, unlike BPD, RDS is a homogeneously distributed disease caused by surfactant deficiency. At the time RDS is evaluated and treated, the effect of patient positioning (dependent vs nondependent disease distribution) plays no role. As a result, the original LUS score described by Brat et al[13] for RDS explored only anterolateral fields that are easily accessible in ventilated infants. In contrast, in preterm infants who demonstrate BPD, the posterior lung fields generally are less aerated.[19] This phenomenon of dependent atelectasis impairs oxygenation and increases the LUS scores. However, whether it contributes to lung injury and increases the risk of BPD is not known. If this were the case, adding posterior lung fields to the LUS score may improve the prediction of BPD. A single-center study including the assessment of posterior lung fields[16] found a higher diagnostic accuracy (area under the receiver operating characteristic curve [AUC] at the seventh DOL, 0.94) than in the other studies (AUC from the third DOL to the eighth week of life, 0.63-0.94).[17,18] However, a recent multicenter study did not find added value in exploring posterior lung fields to predict BPD.[20] The main objective of this multicenter study was to assess the diagnostic accuracy of LUS performed in the first week of life to predict the development of moderate to severe BPD (msBPD) and to test the prespecified hypothesis that examination of posterior lung fields will improve diagnostic accuracy.

Methods

This was a prospective, observational, multicenter diagnostic accuracy study at six neonatal centers. Infants born before 32 weeks' gestation from January 2017 through June 2020 were eligible. The exclusion criteria were major malformations or chromosomopathies, palliative care since birth, and death or transfer to another center before 36 weeks' postmenstrual age. Respiratory care was provided according to the latest update of the European guidelines for respiratory support in very-low-birth-weight infants at all participating centers.[21]

Our primary outcome was the development of msBPD using the physiologic definition proposed by Walsh et al.[22] Our secondary outcomes were any grade of BPD (defined as the need for oxygen supplementation for at least 28 postnatal days),[23] number of days of mechanical ventilation, the need for postnatal systemic corticosteroids, and hospital discharge with supplemental oxygen (e-Appendix 1).

LUS Protocol

The study protocol was described in a previous publication.[17] We performed LUS and calculated the LUS score as described originally[13] using only anterolateral fields,[17] and we also added the examination of one posterior field to obtain an additional score. For scanning the posterior lung fields, the infant was placed in a partial lateral decubitus position, and the transducer was positioned longitudinally over the mid paravertebral area. All LUS procedures were performed by one or two neonatologists with extensive experience in LUS. Every LUS examination was obtained after at least 1 h in the supine position. Each center calculated their patient's LUS scores independently, and interobserver agreement was calculated using 20 anonymous LUS images.

Statistical Analysis

Demographic data and outcomes were summarized using descriptive statistics. Diagnostic accuracy based on AUC analysis was calculated for LUS score using only anterolateral fields and LUS score using anterolateral and posterior fields at different time points. Because multiple comparisons were performed, we used Bonferroni's correction. The AUCs of the different scores at each time point were compared using the DeLong test. Optimal cutoff points were selected using the maximum Youden index.

Multivariate logistic regression was used to test whether the LUS scores obtained on the third and seventh DOL provided added diagnostic value compared with msBPD prediction based on clinical variables. Only clinical covariates that are readily available at the time of prediction were considered (e-Appendix 1). The discriminative capacity of the selected models was assessed by the AUC and adjusted R^2 value. ORs (with 95% CIs) were calculated for each LUS score.

According to our hypothesis that the diagnostic accuracy of LUS will improve using the posterior approach, the study sample size was calculated based on comparative AUC analysis between the LUS score using only anterolateral fields and LUS score using anterolateral and posterior fields on the seventh DOL (which we considered to be the optimal time point for BPD prediction). Considering an AUC of 0.8 (from studies using the LUS score using only anterolateral fields), an AUC of 0.94 from the study using a LUS score using anterolateral and posterior fields, and an estimated overall prevalence of msBPD in our cohort of 25%, we needed a sample of 288 patients (α-error and β-error of 5% and 15%, respectively). Weighted κ scores and intraclass correlation coefficients (ICCs) were calculated to assess interobserver agreement in LUS interpretation. All tests used were considered statistically significant if P values were less than .05. All analyses were performed using STATA version 14.2 software (StataCorp).

Ethics

The study protocol was approved by each regional ethics committee with the code number LUS-NEO-17-01, and parents provided written informed consent. The study was conducted following the Strengthening the Reporting of Observational Studies in Epidemiology statement guidelines.[24]

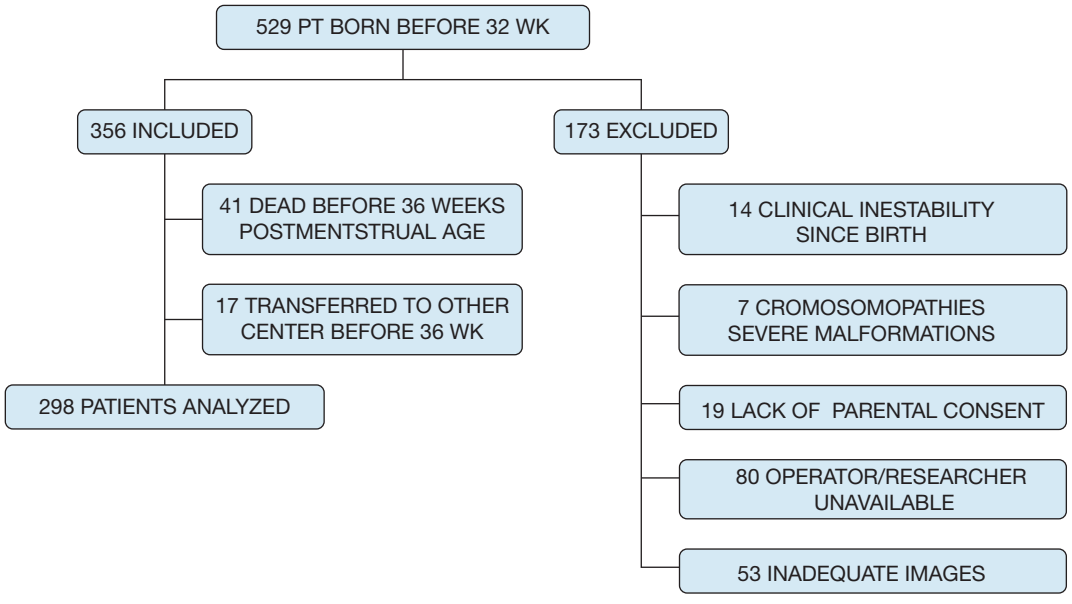

Figure 1 – *Flow diagram showing patient inclusion. PT = preterm infant.*

Results

We included 832 LUS examinations involving 298 patients (Fig 1). The median birth weight and gestational age (GA) of the included infants were 1100 g (interquartile range [IQR], 859-1340 g) and 29 weeks (IQR, 26-30 weeks), respectively; 174 patients were male (58%). msBPD was diagnosed in 73 infants (24.5%), 43 (63%) of whom had a GA of < 28 weeks. The patients' clinical characteristics are described in Table 1.

The median number of LUS examinations per patient was 4 (IQR, 2-4). The LUS score using only anterolateral fields and LUS score using anterolateral and posterior fields adjusted either by GA or postnatal age were highly correlated ($r = 0.97$ and $P < .001$ in both instances). Both the LUS score using only anterolateral fields and LUS score using anterolateral and posterior fields were higher in infants with msBPD than in the rest of the cohort at every time point ($P < .001$ in all instances). The comparison of LUS scores between the study groups is shown in Figure 2 and e-Figure 1. LUS scores on the third and seventh DOL were correlated with the severity of BPD ($P < .001$, linear trend test) (Fig 3).

Diagnostic Accuracy of the LUS Score

The LUS score using only anterolateral fields had a moderate diagnostic accuracy to predict msBPD on the third, seventh, 14th, and 21st DOL, with AUCs of 0.77 (95% CI, 0.68-0.85), 0.79 (95% CI, 0.74-0.84), 0.73 (95% CI, 0.66-0.80), and 0.80 (95% CI, 0.72-0.86), respectively. We found no significant differences between the AUCs of the LUS score using only anterolateral fields and the LUS score using anterolateral and posterior fields in predicting msBPD at birth ($P = .36$), at the third DOL ($P = .97$), at the seventh DOL ($P = .26$), or at the 21st DOL ($P = .17$), as shown in Table 2 and e-Table 1. In contrast, the LUS score using anterolateral and posterior fields showed a significantly higher AUC on the 14th DOL than the LUS score using only anterolateral fields (AUC, 0.77 [95% CI, 0.69-0.83] vs 0.73 [95% CI, 0.66-0.80]; $P = .01$). The multivariate model including LUS score using only anterolateral fields at seventh DOL, sex, and GA (model 3) provided a diagnostic accuracy similar to the model including multiple clinical variables (model 2): AUC, 0.83 (95% CI, 0.77-0.88) vs 0.85 (95% CI, 0.80-0.91; $P = .52$, DeLong test). The best model on the seventh DOL included both clinical variables and LUS-al (model 4): AUC of 0.87 (95% CI, 0.82-0.92) (Table 3). Substituting the LUS score using only anterolateral fields for the LUS score using anterolateral and posterior fields in multivariate analysis did not improve the model's discrimination. The receiver operating characteristic curves for the LUS score using only anterolateral fields and LUS score using anterolateral and posterior fields on the third and seventh DOL, as well as those derived from multivariate models, are shown in Figure 4.

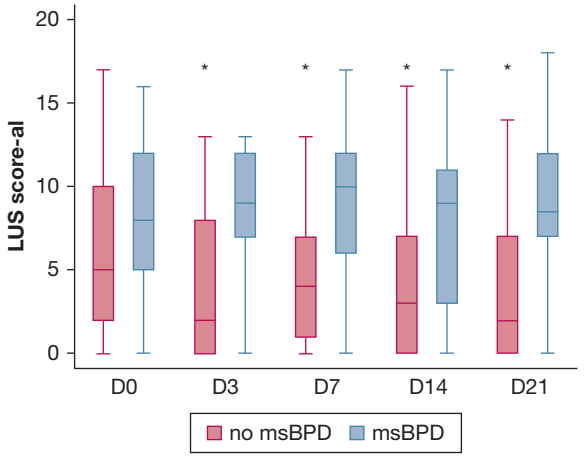

Figure 2 – *Boxplot showing LUS score-al at D0, at D3, at D7, at D14, and at D21 in patients with msBPD and without msBPD. *P < .05 compared with D0 (Bonferroni adjustment for multiple comparisons). D0 = birth; D3 = 3 days of life; D7 = 7 days of life; D14 = 14 days of life; D21 = 21 days of life; LUS score-al = lung ultrasound score using only anterolateral fields; msBPD = moderate to severe bronchopulmonary dysplasia.*

Secondary Outcomes

We found a significant linear trend of both the LUS scores with the respiratory support required by our patients ($P < .001$ for both scores) (Fig 5). We also found positive correlations between LUS scores and the duration of mechanical ventilation at each time point ($P < .01$ in all instances) (e-Table 2). The LUS score on the third and seventh DOL showed a modest diagnostic accuracy to predict treatment with systemic corticosteroids (AUC, 0.76 [95% CI, 0.66-0.84] and 0.76 [95% CI, 0.69-0.81], respectively), as well as oxygen dependency (AUC, 0.72 [95% CI, 0.62-0.81] and 0.73 [95% CI, 0.67-0.79], respectively) (e-Tables 3 and 4). The LUS score on the third and seventh DOL showed a good diagnostic accuracy to predict any grade of BPD (AUC > 0.8 at any time) (e-Table 5 and 6).

LUS Score Agreement Analysis

The median weighted κ score between the observers was 0.82 (IQR, 0.74-0.87). The pooled interobserver intraclass correlation coefficient was 0.98 (95% CI, 0.96-0.99).

Discussion

In this multicenter study, we found that the LUS score predicts the development of msBPD with

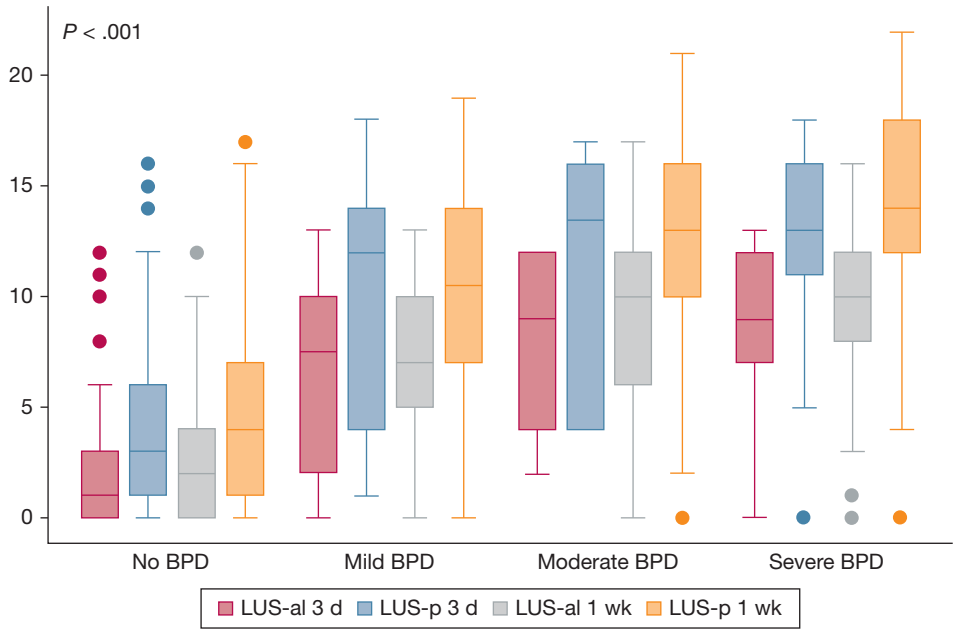

Figure 3 – *Boxplot showing both LUS scores at the third and seventh day of life according to the degree of BPD. Both LUS scores show significative linear tendency using the Jonckheere-Terpstra test at both times. BPD = bronchopulmonary dysplasia; LUS score-al = lung ultrasound score using only anterolateral fields; LUS score-p = lung ultrasound score using anterolateral and posterior fields.*

TABLE 1] Main Descriptive Variables in the Sample

Variable	msBPD (n = 73)	No msBPD (n = 225)	*P* Value
Gestational age, wk	26 (25-28)	29 (27-30)	< .001
Birth weight, g	860 (640-1,000)	1,215 (970-1,400)	< .001
Male sex	49 (67)	125 (56)	.08
White	56 (77)	183 (82)	.86
Twins	18 (25)	66 (29)	.44
In-born	66 (90)	204 (91)	.21
Antenatal steroids	66 (90)	207 (92)	.86
Cesarean delivery	49 (67)	160 (71)	.52
Chorioamnionitis	22 (30)	51 (23)	.20
CRIB-II	11 (9-13)	7 (4-10)	< .001
SNAPPE-II	27 (18-41)	10 (5-26)	< .001
Invasive mechanical ventilation, d	10 (1-40)	0 (0-2)	< .001
Noninvasive mechanical ventilation, d	30 (17-39)	7 (2-23)	< .001
Surfactant dose	...	...	<. 001
1	32 (68)	83 (65)	
2	12 (26)	8 (6)	
3	1 (2)	1 (1)	
Postnatal steroids	25 (34)	9 (4)	< .001
Oxygen at 28 d	73 (100)	85 (38)	< .001
Significant patent ductus arteriosus	21 (29)	12 (5)	< .001
Home oxygen at discharge	27 (39)	3 (1)	< .001
Length of admission, d	92 (76-109)	51 (36-71)	< .001

Data are presented as No. (%) or median (interquartile range), unless otherwise indicated. CRIB-II = clinical risk index for babies II; msBPD = moderate to severe bronchopulmonary dysplasia; SNAPPE II = score for neonatal acute physiology with perinatal extension II.

TABLE 2] Predictive Ability of LUS score-al for msBPD at Different Moments of the Study

Variable	Cutoff Point	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Likelihood Ratio	
						Positive	Negative
LUS score-al 3-d AUC, 0.77 (95% CI, 0.68-0.85)	2	95% (76%-99%)	37% (28%-47%)	24% (16%-35%)	97% (86%-100%)	1.5 (1.2-1.8)	0.14 (0.02-0.94)
	7	**80% (58%-92%)**	**71% (61%-79%)**	**37% (29%-47%)**	**94% (87%-98%)**	**2.8 (1.9-4.1)**	**0.28 (0.17-0.68)**
	13	10% (2.8%-30%)	99% (94%-100%)	67% (21%-94%)	84% (76%-89%)	9.3 (0.9-98)	0.91 (0.8-1.1)
LUS score-al 1-wk AUC, 0.79 (95% CI, 0.74-0.84)	3	91% (81%-96%)	44% (37%-51%)	35% (28%-43%)	93% (86%-97%)	1.6 (1.4-1.9)	0.2 (0.1-0.47)
	8	**70% (58%-80%)**	**79% (72%-84%)**	**53% (45%-61%)**	**89% (84%-92%)**	**3.3 (2.4-4.5)**	**0.38 (0.26-0.56)**
	9	67% (54%-77%)	82% (76%-87%)	55% (44%-66%)	88% (82%-92%)	3.7 (2.6-5.2)	0.41 (0.3-0.58)
	14	8% (3.4%-17.3)	100% (98%-100%)	100% (57%-100%)	76% (71%-81%)	> 100	0.92 (0.86-0.99)
LUS score-al 2-wk AUC, 0.73 (95% CI, 0.66-0.80)	2	88% (73%-95%)	40% (32%-49%)	29% (21%-39%)	92% (82%-97%)	1.5 (1.2-1.8)	0.3 (0.12-0.78)
	8	**63% (47%-78%)**	**80% (71%-86%)**	**47% (36%-58%)**	**89% (83%-92%)**	**3.1 (2.0-4.8)**	**0.46 (0.29-0.73)**
	9	55% (38%-70%)	84% (76%-89%)	49% (33%-64%)	87% (79%-92%)	3.4 (2-5.6)	0.54 (0.37-0.8)
	17	3% (0.5%-15%)	100% (97%-100%)	100% (21%-100%)	79% (71%-84%)	> 100	1.0 (0.97-1.1)
LUS score-al 3-wk AUC, 0.80 (95% CI, 0.72-0.86)	6	90% (74%-97%)	65% (56%-74%)	42% (31%-54%)	96% (89%-99%)	2.6 (1.9-3.4)	0.15 (0.05-0.45)
	7	**77% (59%-88%)**	**74% (65%-81%)**	**45% (36%-54%)**	**92% (85%-96%)**	**2.9 (2.0-4.2)**	**0.32 (0.16-0.61)**
	13	20% (9%-37%)	98% (93%-100%)	75% (41%-93%)	81% (74%-87%)	11 (2.3-49.8)	0.8 (0.7-0.98)
	16	7% (1.9%-21%)	100% (97%-100%)	100% (34%-100%)	79% (72%-85%)	> 100	0.93 (0.85-1.03)

Boldface values indicate the optimal cutoff point. AUC = area under the receiver operating characteristic curve; LUS score-al = lung ultrasound score using only anterolateral fields; msBPD = moderate to severe bronchopulmonary dysplasia.

TABLE 3] Multivariate Analysis to Predict msBPD at 3 and 7 d of Age

Variable	OR (95% CI)	*P* Value	AUC	Adjusted R^2
Model 1			0.77	0.247
GA	0.64 (0.56-0.73)	< .001		
Sex	0.66 (0.36-1.21)	.18		
Model 2 (d 3)			0.80	0.235
GA	0.75 (0.56-1)	.06		
Sex	0.51 (0.16-1.6)	.25		
Prenatal corticosteroids	1.61 (0.65-4.0)	.30		
IMV at d 3	2.37 (0.58-9.6)	.58		
Model 3 (d 3)			0.81	0.289
GA	0.79 (0.61-1.0)	.06		
Sex	0.32 (0.17-1.7)	.29		
LUS score-al d 3	1.20 (1.04-1.4)	.01		
Model 4 (d 3)			0.82	0.31
GA	0.78 (0.58-1.04)	.09		
Sex	0.58 (0.17-1.9)	.37		
Prenatal corticosteroids	1.79 (0.68-4.70)	.24		
IMV at d 3	0.85 (0.16-4.48)	.85		
LUS score-al d 3	1.23 (1.03-1.47)	.02		
Model 2 (d 7)			0.85	0.43
GA	0.76 (0.63-0.90)	.002		
Sex	0.56 (0.27-1.18)	.129		
Prenatal corticosteroids	1.73 (0.96-3.12)	.066		
IMV at d 7	2.46 (1.04-5.80)	.040		
Surfactant	3.20 (1.35-7.54)	.008		
PDA	1.75 (1.15-2.67)	.009		
LOS before d 7	11.0 (0.89-137)	.062		
Model 3 (d 7)			0.83	0.35
GA	0.79 (0.67-0.93)	.006		
Sex	0.54 (0.27-1.08)	.081		
LUS score-al d 7	1.22 (1.11-1.35)	< .001		
Model 4 (d 7)			0.87	0.46
GA	0.87 (0.71-1.06)	.159		
Sex	0.54 (0.25-1.16)	.115		
Prenatal corticosteroids	1.68 (0.87-3.26)	.122		
IMV at d 7	2.21 (0.88-5.54)	.091		
Surfactant	2.30 (0.94-5.65)	.068		
PDA	1.68 (1.07-2.65)	.025		
LOS before d 7	9.52 (0.79-114)	.091		
LUS score-al d 7	1.17 (1.06-1.30)	.003		

All models showed a Hemershow-Lemeshow test with $P > .05$. AUC = area under the receiver operating characteristic curve; GA = gestational age; IMV = invasive mechanical ventilation; LOS = late-onset sepsis; LUS score-al = lung ultrasound score using only anterolateral fields; msBPD = moderate to severe bronchopulmonary dysplasia; PDA = clinically significant persistent ductus arteriosus.

moderate diagnostic accuracy from the third to 21st DOL. Adding the LUS score on the seventh DOL to sex and GA provided a similar diagnostic performance compared with the assessment of multiple BPD clinical risk factors. Globally, the assessment of posterior lung fields did not improve diagnostic accuracy.

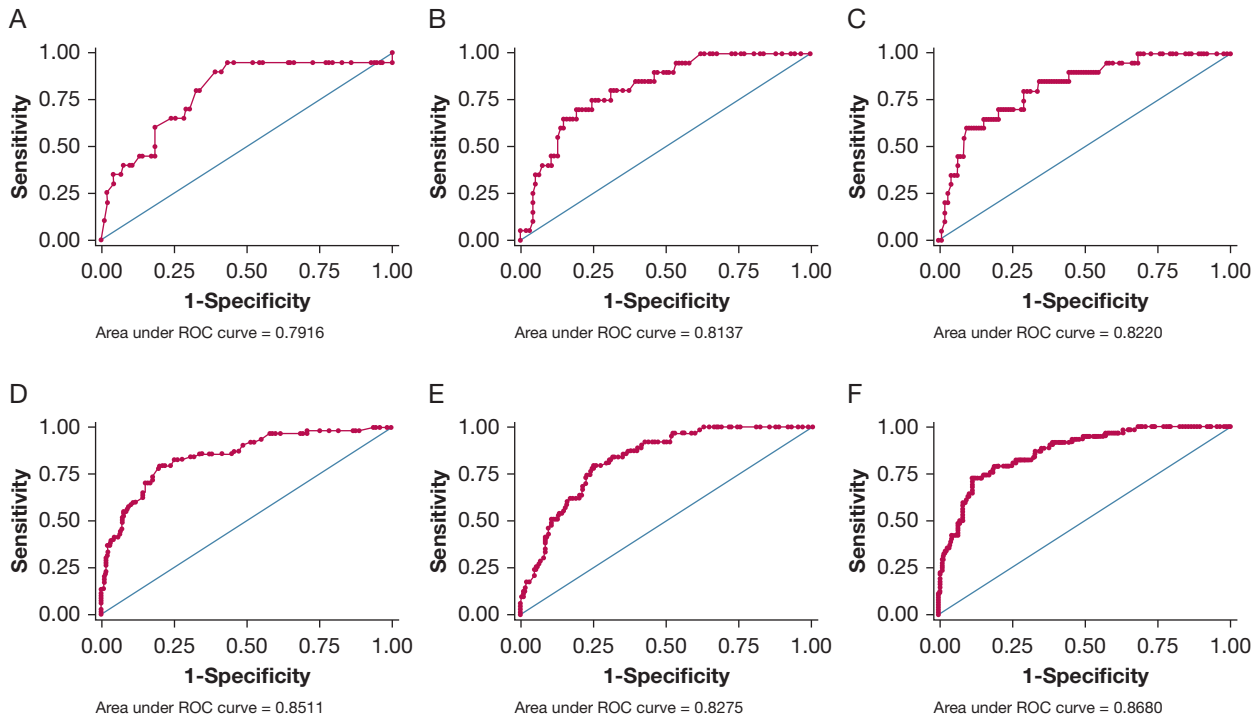

Figure 4 – *A-F, ROC curves from different multivariate models calculated to predict moderate to severe bronchopulmonary dysplasia in infants born before 32 weeks' gestation: (A) model 2 at 3 days, (B) model 4 at 3 days, (C) model 6 at 3 days, (D) model 2 at 1 week, (E) model 4 at 1 week, and (F) model 6 at 1 week. ROC = receiver operating characteristic.*

The first study published using the LUS score used the LUS score using only anterolateral fields[13] to predict the need for surfactant replacement in RDS, and subsequent publications have continued the same methodology to monitor other lung diseases.[14,25,26] Specifically, the LUS score using only anterolateral fields has been shown to be a good predictor of BPD at 1 week of age in small single centers.[17,18] A recent prospective multicenter study confirmed that the LUS score using only anterolateral fields adjusted by GA accurately predicts BPD at 36 weeks' postmenstrual age.[20] Some researchers have incorporated the scanning of posterior lung fields based on the assumption that assessment of the dependent to nondependent distribution of lung aeration may add to the predictive capability of the LUS score.[26,27]

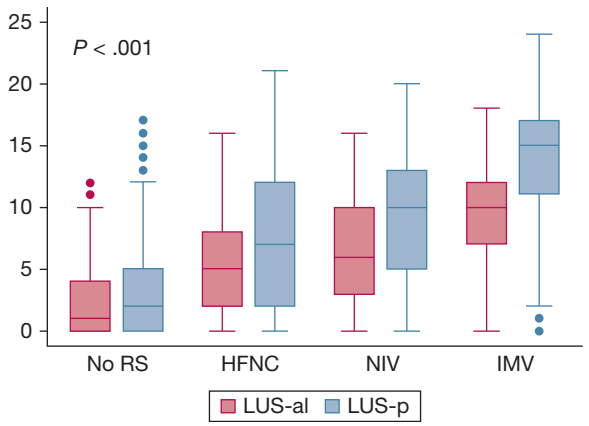

Figure 5 – *Box-and-whisker plot showing both LUS scores according to the type of respiratory support. HFNC = high-flow nasal cannula; IMV = invasive mechanical ventilation; LUS score-al = lung ultrasound score using only anterolateral fields; LUS score-p = lung ultrasound score using anterolateral and posterior fields; NIV = noninvasive ventilation; RS = respiratory support.*

To the best of our knowledge, this is one of the first studies that has compared the original LUS score using only anterolateral fields with the LUS score using anterolateral and posterior fields to predict the development of BPD and related respiratory outcomes. It seems that the diagnostic accuracy of the LUS score using anterolateral and posterior fields is higher than that of the LUS score using only anterolateral fields when LUS is performed at a later stage, although in our study, the difference was not large enough to have clinical relevance. Although worse LUS scores usually are found in dependent areas (mainly posterior lung fields) at any time, perhaps it is the persistence of dependent lung collapse, which may be associated with ventilator-induced lung injury[28] and higher BPD risk.[29] Specific research on the impact of patient positioning scheduling protocols on BPD risk may be worth

attempting, and LUS would be an ideal tool to track the effect of switching body position on lung aeration at the bedside.[30] Until then, it should be emphasized that scanning posterior lung fields compels mobilization of a highly vulnerable preterm infant and may cause harm.

Recently, international evidence-based guidelines on point-of-care ultrasound by the Point of Care Ultrasound Working Group of the European Society of Pediatric and Neonatal Intensive Care were published and provide recommendations for the use of LUS in lung diseases,[31] but they do not offer guidance on which specific scanning protocol or timings should be used in different neonatal lung diseases.

The main benefit of using LUS in BPD would be obtained if prediction can be made early in postnatal life, when changes in therapy have a greater chance of altering the pathologic process and minimizing secondary lung injury.[32] Many authors have identified this "window of opportunity" in the first 2 weeks of life.[3] Latter prediction, although informative, may be less useful from a therapeutic perspective. We have shown that combining the LUS score on the third and seventh DOL with sex and GA improved diagnostic performance, but no additional discrimination was gained compared with the multivariate model that included all clinical risk factors. However, the diagnostic performance of our LUS and clinical models were comparable with the online predictor tool developed by the National Institute of Child Health and Human Development Neonatal Research Network,[33] which is considered to be the best predictive model currently available.[34] Beyond the complexity of constructing predictive models based on multiple clinical risk factors, the main disadvantage of this approach is that many risk factors for BPD are related to therapies that are not completely standardized (eg, ventilator settings, extubation criteria) or actually are comorbidities of prematurity that lack widely accepted definitions (eg, hemodynamically significant patent ductus arteriosus), and therefore are subjected to high variability between centers and over time. These facts make external validation of these models extremely challenging. Conversely, using protocolized LUS examinations combined with basic and uncontroversial clinical variables such as sex and GA provides an important advantage in terms of external validity and implementation in clinical practice.

However, the ability of any of the LUS scores in isolation to predict msBPD in our cohort is not as good as that reported in the multicenter study by Loi et al.[20] However, in that study, GA-adjusted LUS scores were used, so the results are difficult to compare with our data. We demonstrated better results using any grade of BPD as the end point in a previous study,[17] and these results have been replicated with a larger sample. It seems that early prediction of advanced grades of BPD is challenging. Most msBPD cases occur in preterm infants younger than 28 weeks' gestation. We previously showed that these more immature infants maintain naturally higher LUS scores during the first weeks of life compared with infants older than 28 weeks' gestation, even when they do not finally demonstrate msBPD, raising concerns about an age dependency of LUS scores.[17,35] This fact may be related to pulmonary insufficiency of prematurity, excessive fluid overload, frequent lung atelectasis, or other factors.[17,36] This phenomenon of persistent white lung can affect the predictive capacity of early LUS by a selection bias when cohorts include a wide range of GAs. We think that specific research in more premature infants should be accomplished to test the usefulness of the LUS score to predict BPD according to GA because subgroup analysis for this purpose may be misleading. Another factor that may explain our lower predictive capability is the decision to exclude dead infants. Most infants who do not survive are at great risk of advanced BPD and may have high LUS scores. The aforementioned study by Oulego-Erroz et al,[16] which reported the highest diagnostic accuracy of the LUS score (AUC of 0.94 on the seventh DOL), used a combined outcome of msBPD, death, or both. Composite outcomes are an attractive solution to limit sample size and at the same time increase statistical efficiency when individual outcomes are rare or behave as competing events. However, to be used properly, all components of the composite outcome should be equally clinically relevant, a condition that often is not met. As a result, the use of these composite outcomes has been criticized.[37] Moderate BPD accounts for most cases of msBPD. These patients usually follow a favorable course, and in our opinion, inclusion in the same category as dead infants is debatable.

The LUS score reflects decreased lung aeration, which may be caused by increased water content[38] and inflammatory lung edema[39] and has been shown to be highly correlated with oxygenation indexes.[13] In our study, we also showed that the LUS score increases with increasing respiratory support and correlates with the duration of mechanical ventilation, which is in concordance with previous studies.[16,18] The need for

postnatal systemic corticosteroids (a surrogate indicator of severe respiratory course) was predicted with modest accuracy early after birth, although these results should be taken with caution because this treatment was not standardized and was prescribed at the neonatologist's discretion. Another important respiratory outcome is the need for home oxygen in BPD patients. Infants may pass the oxygen reduction test at 36 weeks' postmenstrual age and be classified as having mild BPD. However, some of these infants ultimately are discharged home with supplemental oxygen and follow a clinical course similar to that of infants with moderate BPD. We showed that the LUS score can help to predict this outcome from the third DOL, but with modest diagnostic accuracy. A recent study showed that later LUS on the 28th DOL more accurately predicts oxygen dependency at discharge.[40]

Our study has several limitations. As with all observational studies, the main limitation is the presence of uncontrolled confounding factors, such variability in respiratory and general management among participating centers. Thus, although all LUS examinations were performed after at least 1 h in the supine position, the precise time that each patient may have been in a supine or prone position before LUS was unknown, and this may affect LUS findings, especially regarding the assessment of dependent lung zones. Finally, we used the physiological definition of BPD (a modification of the original National Institute of Child Health and Human Development workshop definition[23]), which is still the most commonly used in daily practice. However, this definition is becoming obsolete as modes of respiratory support have evolved (eg, generalization of high-flow oxygen cannula). It may be worth exploring the potential of the LUS score using new definitions such as the one by Jensen et al,[41] which has been shown to be associated more closely with mortality and long-term outcomes than classic BPD definitions.

LUS merits all the attributes to be considered an imaging biomarker of lung disease[42]: it can be measured objectively and accurately, it correlates with lung disease severity,[20] and it permits the assessment of treatment response (recruitment maneuvers,[43] fluid restriction and diuretics,[44] and so forth) and predicts relevant outcomes (need for surfactant,[13] extubation failure,[45] or BPD[20]). However, BPD is a very complex disease with different phenotypes,[46] and many influencing factors make it highly unlikely that LUS would be the so-called silver bullet for prediction. Further research should explore the combination of LUS with other clinical variables and biological biomarkers.[47]

Interpretation

The LUS score predicts the development of msBPD with moderate diagnostic accuracy on the third and seventh DOL. Adding sex and GA to the LUS score increases diagnostic accuracy and provides a diagnostic performance comparable with the assessment of multiple clinical risk factors. Adding examination of posterior lung zones does not seem to improve the early prediction of msBPD.

Acknowledgments

Author contributions: A. A.-O. takes full responsibility of the content of the manuscript including data and analysis. A. A.-O. conceived and designed the study, analyzed the data, and drafted the manuscript. I. S.-G., V. A.-B., R. G.-H., P. A.-Q., A. C.-G., A. R.-R., M. H. M., and L. R.-F. acquired data, analyzed data, and critically reviewed the manuscript. I. O.-E. acquired data, edited the first and subsequent drafts, and critically reviewed the manuscript for relevant intellectual content. All authors gave their approval for the final version of the manuscript.

Financial/nonfinancial disclosures: None declared.

Additional information: The e-Appendix, e-Figure, and e-Tables can be found in the Supplemental Materials section of the online article.

References

1. Lui K, Lee SK, Kusuda S, et al. Trends in outcomes for neonates born very preterm and very low birth weight in 11 high-income countries. *J Pediatr*. 2019;215:32-40.e14.
2. Alvarez-Fuente M, Moreno L, Lopez-Ortego P, et al. Exploring clinical, echocardiographic and molecular biomarkers to predict bronchopulmonary dysplasia. *PloS One*. 2019;14(3):e0213210.
3. Álvarez-Fuente M, Moreno L, Mitchell JA, et al. Preventing bronchopulmonary dysplasia: new tools for an old challenge. *Pediatr Res*. 2019;85(4):432-441.
4. Förster K, Sass S, Ehrhardt H, et al. Early identification of bronchopulmonary dysplasia using novel biomarkers by proteomic screening. *Am J Respir Crit Care Med*. 2018;197(8):1076-1080.
5. Rivera L, Siddaiah R, Oji-Mmuo C, Silveyra GR, Silveyra P. Biomarkers for bronchopulmonary dysplasia in the preterm infant. *Front Pediatr*. 2016;4:33.
6. Brusa G, Savoia M, Vergine M, Bon A, Copetti R, Cattarossi L. Neonatal lung sonography: interobserver agreement between physician interpreters with varying levels of experience. *J Ultrasound Med*. 2015;34(9):1549-1554.
7. Vergine M, Copetti R, Brusa G, Cattarossi L. Lung ultrasound accuracy in respiratory distress syndrome and transient tachypnea of the newborn. *Neonatology*. 2014;106(2):87-93.
8. Raimondi F, Yousef N, Rodriguez Fanjul J, et al. A multicenter lung ultrasound study on transient tachypnea of the neonate. *Neonatology*. 2019;115(3):263-268.
9. Puch-Kapst K, Juran R, Stoever B, Wauer RR. Radiation exposure in 212

very low and extremely low birth weight infants. *Pediatrics*. 2009;124(6):1556-1564.

10. Escourrou G, De Luca D. Lung ultrasound decreased radiation exposure in preterm infants in a neonatal intensive care unit. *Acta Paediatr*. 2016;105(5):e237-e239.
11. Kurepa D, Zaghloul N, Watkins L, Liu J. Neonatal lung ultrasound exam guidelines. *J Perinatol*. 2018;38(1):11-22.
12. Raschetti R, Yousef N, Vigo G, et al. Echography-guided surfactant therapy to improve timeliness of surfactant replacement: a quality improvement project. *J Pediatr*. 2019;212:137-143.e1.
13. Brat R, Yousef N, Klifa R, Reynaud S, Shankar Aguilera S, De Luca D. Lung ultrasonography score to evaluate oxygenation and surfactant need in neonates treated with continuous positive airway pressure. *JAMA Pediatr*. 2015;169(8):e151797.
14. De Martino L, Yousef N, Ben-Ammar R, Raimondi F, Shankar-Aguilera S, De Luca D. Lung ultrasound score predicts surfactant need in extremely preterm neonates. *Pediatrics*. 2018;142(3):e20180463.
15. Aldecoa-Bilbao V, Balcells-Esponera C, Herranz Barbero A, et al. Lung ultrasound for early surfactant treatment: development and validation of a predictive model. *Pediatr Pulmonol*. 2021;56(2):433-441.
16. Oulego-Erroz I, Alonso-Quintela P, Terroba-Seara S, Jiménez-González A, Rodríguez-Blanco S. Early assessment of lung aeration using an ultrasound score as a biomarker of developing bronchopulmonary dysplasia: a prospective observational study. *J Perinatol*. 2021;41(1):62-68.
17. Alonso-Ojembarrena A, Lubián-López SP. Lung ultrasound score as early predictor of bronchopulmonary dysplasia in very low birth weight infants. *Pediatr Pulmonol*. 2019;54(9):1404-1409.
18. Abdelmawla M, Louis D, Narvey M, Elsayed Y. A lung ultrasound severity score predicts chronic lung disease in preterm infants. *Am J Perinatol*. 2019;36(13):1357-1361.
19. Louis D, Belen K, Farooqui M, et al. Prone versus supine position for lung ultrasound in neonates with respiratory distress. *Am J Perinatol*. 2021;38(2):176-181.
20. Loi B, Vigo G, Baraldi E, et al. Lung ultrasound to monitor extremely preterm infants and predict BPD: multicenter longitudinal cohort study. *Am J Respir Crit Care Med*. 2021;203(11):1398-1409.
21. Sweet DG, Carnielli V, Greisen G, et al. European Consensus Guidelines on the management of respiratory distress syndrome—2016 update. *Neonatology*. 2017;111(2):107-125.
22. Walsh MC, Yao Q, Gettner P, et al. Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics*. 2004;114(5):1305-1311.
23. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2001;163(7):1723-1729.
24. Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med*. 2007;147(8):W163-W194.
25. Perri A, Riccardi R, Iannotta R, et al. Lung ultrasonography score versus chest X-ray score to predict surfactant administration in newborns with respiratory distress syndrome. *Pediatr Pulmonol*. 2018;53(9):1231-1236.
26. Rodriguez-Fanjul J, Jordan I, Balaguer M, Batista-Muñoz A, Ramon M, Bobillo-Perez S. Early surfactant replacement guided by lung ultrasound in preterm newborns with RDS: the ULTRASURF randomised controlled trial. *Eur J Pediatr*. 2020;179(12):1913-1920.
27. Gregorio-Hernández R, Arriaga-Redondo M, Pérez-Pérez A, Ramos-Navarro C, Sánchez-Luna M. Lung ultrasound in preterm infants with respiratory distress: experience in a neonatal intensive care unit. *Eur J Pediatr*. 2020;179(1):81-89.
28. Beitler JR, Malhotra A, Thompson BT. Ventilator-induced lung injury. *Clin Chest Med*. 2016;37(4):633-646.
29. Keszler M, Sant'Anna G. Mechanical ventilation and bronchopulmonary dysplasia. *Clin Perinatol*. 2015;42(4):781-796.
30. Liu J, Chen S-W, Liu F, Li Q-P, Kong X-Y, Feng Z-C. The diagnosis of neonatal pulmonary atelectasis using lung ultrasonography. *Chest*. 2015;147(4):1013-1019.
31. Singh Y, Tissot C, Fraga MV, et al. International evidence-based guidelines on point of care ultrasound (POCUS) for critically ill neonates and children issued by the POCUS Working Group of the European Society of Paediatric and Neonatal Intensive Care (ESPNIC). *Crit Care*. 2020;24(1):65.
32. Thébaud B, Goss KN, Laughon M, et al. Bronchopulmonary dysplasia. *Nat Rev Dis Primer*. 2019;5(1):78.
33. Laughon MM, Langer JC, Bose CL, et al. Prediction of bronchopulmonary dysplasia by postnatal age in extremely premature infants. *Am J Respir Crit Care Med*. 2011;183(12):1715-1722.
34. Onland W, Debray TP, Laughon MM, et al. Clinical prediction models for bronchopulmonary dysplasia: a systematic review and external validation study. *BMC Pediatr*. 2013;13:207.
35. Alonso-Ojembarrena A, Francesco R. Using LU score to predict extubation failure in preterm infants should consider gestational age at birth. *Pediatr Pulmonol*. 2020;55(12). 3228-3228.
36. Steinhorn R, Davis JM, Göpel W, et al. Chronic pulmonary insufficiency of prematurity: developing optimal endpoints for drug development. *J Pediatr*. 2017;191:15-21.e1.
37. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ*. 2010;341:c3920.
38. Zong H, Guo G, Liu J, Bao L, Yang C. Using lung ultrasound to quantitatively evaluate pulmonary water content. *Pediatr Pulmonol*. 2020;55(3):729-739.
39. Yousef N, Vigo G, Shankar-Aguilera S, De Luca D. Semiquantitative ultrasound assessment of lung aeration correlates with lung tissue inflammation. *Ultrasound Med Biol*. 2020;46(5):1258-1262.
40. Hoshino Y, Arai J, Miura R, et al. Lung ultrasound for predicting the respiratory outcome in patients with bronchopulmonary dysplasia [published online ahead of print December 29, 2020]. *Am J Perinatol*. https://doi.org/10.1055/s-0040-1721848.
41. Jensen EA, Dysart K, Gantz MG, et al. The diagnosis of bronchopulmonary dysplasia in very preterm infants. An evidence-based approach. *Am J Respir Crit Care Med*. 2019;200(6):751-759.
42. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95.
43. Shady NMA, Awad HAS, Kamel DR, Fouda EM, Ahmed NT, Dawoud MO. Role of lung ultrasound in the assessment of recruitment maneuvers in ventilated preterm neonates with respiratory distress syndrome and its correlation with tracheal IL-6 levels: a randomized controlled trial [published online ahead of print December 3, 2020]. *J Neonatal-Perinat Med*. https://doi.org/10.3233/NPM-200605.
44. Alonso-Ojembarrena A, Lechuga-Sancho AM, Morales-Arandojo P, Acuñas-Soto S, López-de-Francisco R, Lubián-López SP. Lung ultrasound score and diuretics in preterm infants born before 32 weeks: a pilot study. *Pediatr Pulmonol*. 2020;55(12):3312-3318.
45. El Amrousy D, Elgendy M, Eltomey M, Elmashad AE. Value of lung ultrasonography to predict weaning success in ventilated neonates. *Pediatr Pulmonol*. 2020;55(9):2452-2456.
46. Wang S-H, Tsao P-N. Phenotypes of bronchopulmonary dysplasia. *Int J Mol Sci*. 2020;21(17):6112.
47. Rodriguez-Blanco S, Oulego-Erroz I, Gautreaux-Minaya S, Perez-Muñuzuri A, Couce-Pico ML. Early NT-proBNP levels as a screening tool for the detection of hemodynamically significant patent ductus arteriosus during the first week of life in very low birth weight infants. *J Perinatol Off J Calif Perinat Assoc*. 2018;38(7):881-888.