

Biochemical and Lung Function Test Accuracy for Predicting the Need for Surfactant Therapy in Preterm Infants: A Systematic Review

Anna Lavizzari^a Chiara Veneroni^b^aNeonatal Intensive Care Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy;^bDipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano University, Milan, Italy

Keywords

Neonatal respiratory distress syndrome · Click test · Stable microbubble test · Lamellar bodies count · Forced oscillation technique

Abstract

Introduction: We evaluate the accuracy of postnatal biochemical and lung function tests performed within 3 h from birth for predicting surfactant need in preterm infants ≤ 34 weeks' gestation receiving noninvasive respiratory support for respiratory distress syndrome (RDS). **Methods:** We systematically searched MEDLINE, Embase, The Cochrane Library, PROSPERO, and clinicaltrials.gov databases for studies published from 2000 to November 10, 2021, cross-referencing relevant literature and contacting experts. We included diagnostic accuracy studies and systematic reviews of biochemical or lung function tests identifying the need for surfactant in preterm neonates ≤ 34 weeks' with RDS not intubated at birth. The authors individually assessed the risk of bias following a tailored QUADAS-2 tool. **Results:** Eight studies, including 810 infants, met the inclusion criteria. Four tests were included: the click test, the stable microbubble test, the lamellar body count on gastric aspirates, and the forced oscillation technique. The reference standards were transparent criteria for distinguishing the infants according to oxygen requirement, which reflected the current criteria

for surfactant therapy. The risk of bias was judged high because of the population selection and exclusion of participants from the analysis. There were no serious concerns regarding blinding and applicability. The individual study sensitivity and specificity range from 0.60 to 1 and from 0.51 to 0.91, respectively. It was not appropriate to combine the accuracy estimates in a meta-analysis because of the heterogeneity of the study characteristics. **Conclusions:** Current evidence is insufficient to recommend biochemical and lung function tests for tailoring surfactant therapy.

© 2022 The Author(s).

Published by S. Karger AG, Basel

Introduction

International recommendations for managing respiratory distress syndrome (RDS) advocate using nasal continuous positive airway pressure and an early rescue surfactant policy as the gold standard [1–4]. Early rescue surfactant administration aims to compensate for the lack of pulmonary surfactant in preterm infants that results in an increased alveolar surface tension, reduced lung compliance and leads to alveolar collapse [5, 6]. The derecruited parts of the lung do not contribute to gas exchange.

Current indications for surfactant therapy rely on the oxygen required to maintain a target peripheral oxygen

saturation [1–3]. Oxygenation-based criteria are straightforward for applicability but assume that hypoxia is a direct marker of surfactant deficiency. On the one hand, this oxygenation-based approach may miss infants requiring surfactant. In fact, either the external respiratory support or the self-generated work of breathing may compensate for the increased surface tension and tendency to lung derecruitment, delaying the increase in oxygen demand [7]. On the other hand, some infants may require oxygen supplementation for reasons other than surfactant deficiency, including hemodynamic impairment [8], poor peripheral perfusion, pulmonary hypertension, inadequate thermoregulation, lung hypoplasia, or insufficient respiratory support [9]. In summary, oxygenation-based criteria may be inaccurate; second, they may not allow a timely treatment [9].

The direct evaluation of the surfactant deficiency, lung mechanics, or recruitment has the potential for more accurate and earlier identification of infants requiring exogenous surfactant. Different techniques have been explored for testing surfactant quantity and biophysical properties. These techniques can be applied to different specimens including amniotic fluid, and gastric or tracheal aspirates [10]. Also, despite the challenges of noninvasive evaluation of lung recruitments and mechanics in the first hours of life, new techniques are becoming available at the bedside, such as the lung ultrasound (LUS) and the forced oscillation technique (FOT). The LUS estimates the parenchymal lung aeration using a semiquantitative score. The FOT measures the respiratory system reactance, which is directly related to lung compliance. All these techniques are noninvasive, relatively easy to perform, and suitable for preterm infants. However, their use has been sparse in various settings, often following different protocols for standardization, and only a few tests have been applied in large clinical trials after pilot studies [11].

The objective of the current systematic review is to evaluate the accuracy of biochemical and lung function tests conducted within 3 h of birth in identifying the need for surfactant in preterm infants ≤ 34 weeks of gestation receiving noninvasive respiratory support for RDS.

Methods

The review protocol was registered on PROSPERO (International Prospective Register of Systematic Reviews; CRD42021279259). We followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses of Diagnostic Test Accuracy (The PRISMA-DTA Statement) guidance for reporting our findings [12].

Literature Search

We conducted a comprehensive literature search in three databases (MEDLINE, Cochrane Library, Embase), restricting to human subjects, English language, and peer-reviewed papers and abstracts. We looked at publications from 2000 to guarantee a preterm infant population that may reflect the current clinical practice. In the last decades, dramatic variations in neonatal care, including the widespread use of antenatal steroids, the reduction in intubation and invasive mechanical ventilation, and the surge of CPAP coupled with surfactant administration by INSURE technique, resulted in a significant change in the preterm infant population. Severe RDS requiring intubation and invasive mechanical ventilation became less frequent in favor of milder disease forms, treated with noninvasive support and early rescue surfactant. In severe RDS requiring intubation, surfactant administration is mandatory and easily assessed by standard clinical tools. Differently, the need for exogenous surfactant in milder RDS forms is less easily predictable at birth and noninvasive tests to early identify surfactant deficiency gain relevant interest. We also looked at two registries (PROSPERO and clinicaltrials.gov) for eligible, unpublished studies, and systematic reviews. To optimize the search strategy for diagnostic accuracy studies, we followed the suggestions by Leeftang et al. [13]. Online supplementary text (for all online suppl. material, see www.karger.com/doi/10.1159/000527670) reports the search syntax used for each database/register. The last search was performed on November 10, 2021. Additional studies were identified by cross-referencing relevant literature and contacting experts in the field. The title and the abstract were manually screened individually by the two authors for relevance to the review before assessing the full-text article for eligibility. Figure 1 represents the literature search and selection by the PRISMA 2020 flow diagram for systematic reviews [14].

Risk of Bias and Applicability Assessment

The two authors individually assessed the risk of bias using QUADAS-2 (a revised tool for the quality assessment of diagnostic accuracy studies) for the four key domains: patient selection, index test, reference standard, and flow and timing. The tool was tailored to the systematic review question as suggested by the members of the QUADAS-2 group [13] and described in the online supplementary text. In the absence of a robust “reference standard” test for defining surfactant need, oxygenation-based criteria for surfactant therapy were considered “reference standard.” Additionally, the authors evaluated the concerns about applicability for the patient selection, index test, and reference standard domains using the same appraisal tool [13]. A.L. and C.V. resolved discrepancies by discussion and consensus.

Population, Index Test, Outcomes, and Type of Evidence

Studies that included preterm infants ≤ 34 weeks of gestation were eligible because they represent the age category at risk of RDS. More mature infants typically present with other causes of respiratory failure. We included spontaneously breathing infants on noninvasive respiratory support since intubated preterm infants with RDS would certainly receive surfactant [1], making the need for predicting tests unwarranted. We excluded publications of preterm infants presenting respiratory failure secondary to causes other than RDS (e.g., congenital malformations, genetic or metabolic disorders, hydrops, heart disease).

The index tests were laboratory and lung function tests performed within 3 h from birth since evidence supports an early res-

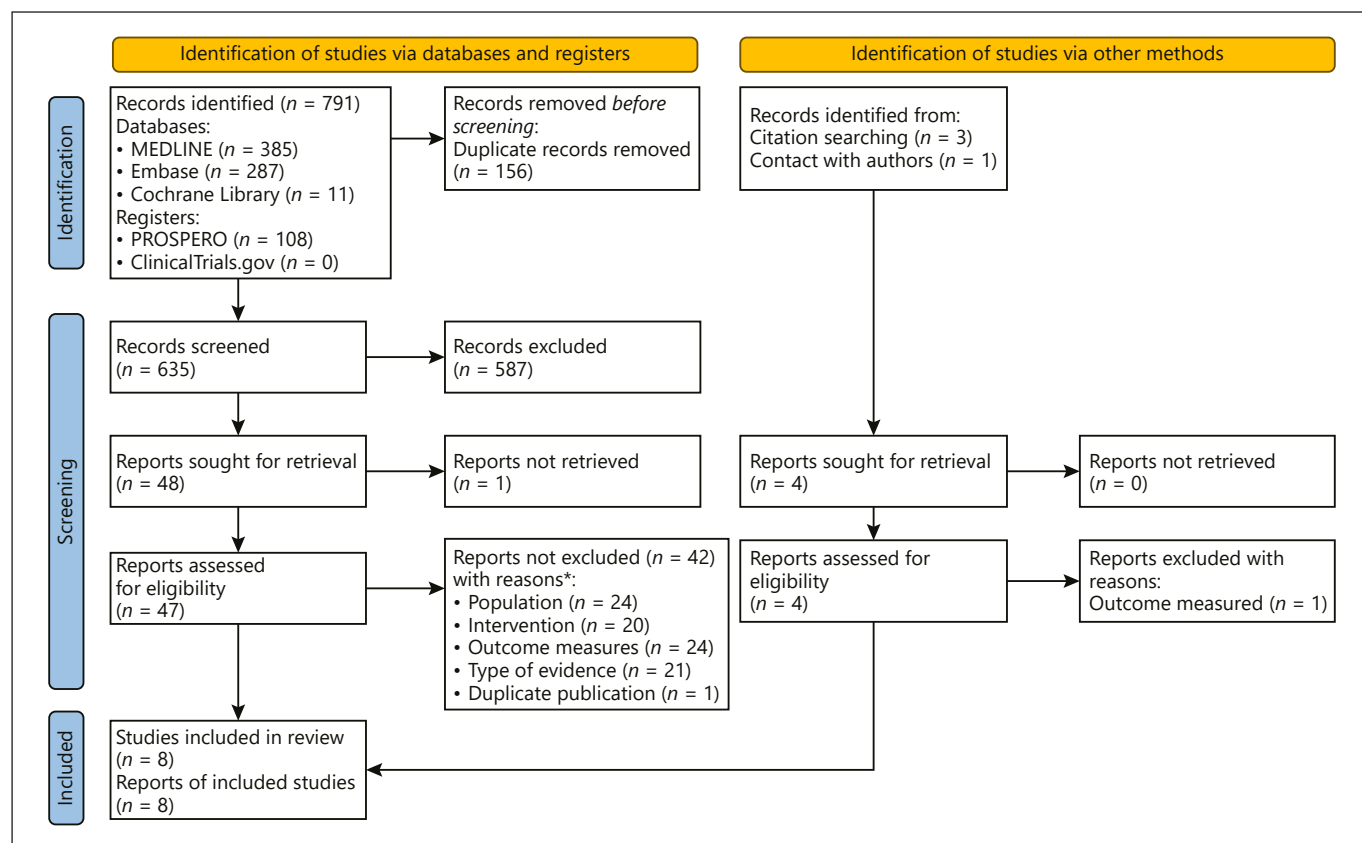


Fig. 1. PRISMA 2020 flow diagram representing the literature search (including searches of databases, registers, and other sources) and selection. Some reports were excluded for more than one reason. Figure adapted from: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, and Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372:n71. doi: 10.1136/bmj.n71.

cue surfactant policy [15]. Studies using imaging and clinical scores were excluded because they were appraised elsewhere [16] or under evaluation in an ongoing systematic review (PROSPERO, CRD42021247888), and for the need for more objective tools than clinical evaluation. The target condition was RDS requiring surfactant, indicative of a moderate to severe disease, whereas infants with mild disease can generally be easily managed with noninvasive respiratory support alone.

We included all diagnostic accuracy studies and systematic reviews aiming at identifying the need for surfactant therapy (or clearly defining the population based on FiO_2 requirement, which reflects current criteria for surfactant therapy-independently from the effective surfactant administration) for RDS. Studies evaluating the diagnostic accuracy for only predicting RDS without providing details on surfactant therapy were excluded.

Data Extraction, Diagnostic Accuracy Measures, and Other Outcomes

Both authors extracted separately the desired data from the reports, including source, setting, time frame for recruitment, characteristic of the study population, sample size, inclusion and exclusion criteria, study design, index test, target condition, and reference standard. The target condition was RDS requiring surfactant. The refer-

ence standard was the criteria for surfactant therapy or transparent criteria for distinguishing the infants-based FiO_2 requirement, which reflected the current criteria for surfactant therapy, independently of the effective surfactant administration. Diagnostic accuracy measures reported were at least one among sensitivity and specificity, area under the curve-receiving operator characteristic (AUC-ROC), positive predictive value and negative predictive value (NPV), and the 95% confidence intervals (95% CI). The number of true positive, false positive, false negative, and true negative were calculated when not directly provided in the report and other estimates for diagnostic accuracy when possible. The Youden Index was computed [17]. Other outcomes included were the time of life the test was performed and the timing for the final results.

Results

Inclusion of Studies and Study Characteristics

The database and register searches identified 791 records. After the initial screening, we performed a full-text review of 47 studies. Four additional studies were found

Table 1. Characteristics of included studies

Author, year, source	Technique	Index test	Setting, country, time frame for enrollment	Study design	Patients recruited, [included in the analysis], <i>n</i>	Inclusion criteria	Exclusion criteria	Target condition and reference standard	Blinding to test results
Fiori et al. 2003 [18] Full-Text	CT on GASP	Presence/absence	Multicenter Brazil; March 1998 to October 1999	Prospective, cohort study	115 [104]	GA: 24–34 weeks, not intubated at birth	Severe congenital hydrops fetalis, infants who received surfactant before sampling	RDS diagnosed <24 h of life and defined as: tachypnea, chest retractions, grunting AND $\text{FiO}_2 \geq 0.40$ to target SpO_2 90–95 % AND reticular granularity and air bronchograms at the chest radiograph	Yes
	SMT on GASP	MB < 15 μm , <i>n</i>			115 [110]				
Fiori et al. 2001 [21] Full-Text	SMT on GASP, computerized analysis	MB < 20 μm , <i>n</i> proportion of MB < 20 μm average MB diameter	Monocenter Brazil; not specified	Prospective, cohort study	82 [80]	GA: 24–34 weeks, not intubated at birth	GASP contaminated with pus	RDS diagnosed <24 h of life and defined as: $\text{FiO}_2 \geq 0.40$ to target SpO_2 90–95 % AND reticular granularity and air bronchograms at the chest radiograph	Yes
Verder et al. 2003 [22] Full-Text	SMT on GASP, computerized analysis	MB < 20 μm , <i>n</i> proportion of MB < 20 μm average MB diameter	Multicenter Denmark, Sweden; January to October 2000	Prospective, cohort study	194 [188]	GA < 32 weeks, on NCPAP	Causes of respiratory failure other than RDS	Moderate/severe RDS defined as: N. symptoms ≥ 2 among the following: need for supplemental O_2 , tachypnea, intercostal retractions, and grunting. Moderate RDS: a $\text{APO}_2 < 0.36$ and >0.14 during the first 72 h Severe RDS: mechanical ventilation within 24 h (main indications: a $\text{APO}_2 < 0.15$ or severe apnea)	Yes
Bhatia et al. 2013 [19] Full-Text	SMT on GASP	MB < 15 μm , <i>n</i>	Monocenter Australia; August 2009 to December 2010	Prospective, cohort study	74 [68]	GA < 30 weeks, not intubated at birth	Congenital abnormality or condition existed that might have an adverse effect on breathing, apart from prematurity	CPAP failure during the first 72 h and defined as: (1) NCPAP 8 cmH_2O and $\text{FiO}_2 \geq 0.4$ for >30 min for SpO_2 88–92 %; (2) apnea unresponsive to stimulation and methylxanthine; (3) arterial pH < 7.25 with $\text{PaCO}_2 > 60$ mm Hg or metabolic acidosis	Yes

Table 1 (continued)

Author, year, source	Technique	Index test	Setting, country, time frame for enrollment	Study design	Patients recruited, included in the analysis], <i>n</i>	Inclusion criteria	Exclusion criteria	Target condition and reference standard	Blinding to test results
Ribeiro et al. 2019 [23] Full-Text	SMT on GASP	MB < 15 μ m, <i>n</i>	Monocenter Brazil; January 2014 to September 2015	Prospective, cohort study	72 [64]	GA <34 weeks	Congenital malformations, genetic syndromes, amniotic fluid with meconium	RDS defined as: tachypnea, chest retractions, grunting AND $F_{iO_2} \geq 0.40$ to target SpO_2 90-95 % AND reticular granularity and air bronchograms at the chest radiograph	Yes
	SMT on OA	MB < 15 μ m, <i>n</i>							
Daniel et al. 2010 [24] Full-Text	SMT on GASP	MB < 15 μ m, <i>n</i>	Monocenter Brazil; October 2007 to June 2008	Prospective, cohort study	63 [63]	GA: 24–34 weeks	Pneumonia, congenital heart disease, CDH, PNX, genetic syndromes that comprise respiratory function, amniotic fluid with meconium	RDS defined as: tachypnea, chest retractions, grunting AND $F_{iO_2} \geq 0.40$ AND reticular granularity and air bronchograms at the chest radiograph	Yes
	LBC on GASP								
Verder et al. 2011 [25] Full-Text	LBC on GASP		Multicenter Denmark; 2000–to 2005	Prospective, cohort study	108 [108]	< 32 weeks on NCPAP; failure other than RDS	Causes of respiratory failure other than RDS	Moderate/severe RDS defined as: N. symptoms ≥ 2 among the following: need for supplemental O_2 , tachypnea, intercostal retractions and grunting. Moderate RDS: $a/APO_2 < 0.35$ and > 0.15 during the first 72 h Severe RDS: mechanical ventilation within 24 h (main indications: $a/APO_2 < 0.15$ or severe apnea)	Unclear
					83 [70]	<30 weeks on NCPAP			
Lavizzari et al. 2021 [20] Full-Text	Oscillatory mechanics assessed by FOT	Xrs	Monocenter Italy; December 2016 to February 2020	Prospective, cohort study	61 [59]	GA 27–34 weeks, not abnormalities, intubated at birth	Major congenital abnormalities, study entry, neonatal ARDS, participation in interventional studies	Surfactant administration during the first 24 h of life: $F_{iO_2} > 0.40$ to target SpO_2 OR because of a sustained increase in F_{iO_2} associated with a Silverman score > 6 after NCPAP initiation	Yes

a/APO_2 , arterial to alveolar oxygen tension ratio; ARDS, acute respiratory distress; CDH, congenital diaphragmatic hernia; CT, click test; F_{iO_2} , fraction of inspired oxygen; FOT, forced oscillation technique; GA, gestational age; GASP, gastric aspirate; LBC, lamellar body count; MB, microbubble; MV, mechanical ventilation; NCPAP, nasal continuous positive airway pressure; O_2 , oxygen; OA, oral aspirate; PNX, pneumothorax; SMT, stable microbubble test; RDS, respiratory distress syndrome; Xrs, respiratory reactance.

Table 2. Characteristics of the study population of included studies

Author, year	GA, weeks	Birth weight, g	Male, N (%)	Antenatal steroids, N (%)	Caesarean, N (%)	Apgar at 5 min
Fiori et al. 2003 [18]	Mean (range), 31.9 (24–34) ¹	Mean, 1,600 ¹	17 (73.9)	4 (17.4)	14 (60.9)	Mean±SD 8.2±1.2
Fiori et al. 2001 [21]	Mean±SD, 32±2	Mean±SD 1,629±429	NA	NA	NA	NA
Verder et al. 2003 [22]	Mean (range), 29 (23–31)	Mean (range) 1,281 (475–2,440)	NA	NA	145 (77)	NA
Bhatia et al. 2013 [19]	Mean±SD, 28.1±1.4	Mean 1,132 ¹	32 (47) ¹	66 (97) ¹	42 (61) ¹	Median [IQR], 8 [7; 9] ¹
Ribeiro et al. 2019 [23]	Median [IQR], 31 [29; 32]	Mean±SD, 1,465±517	33 (52)	2 doses: 42 (66), one dose: 15 (24)	49 (77)	Range, 7–9 ¹
Daniel et al. 2010 [24]	Mean, 31.0 ¹	Mean, 1,454 ¹	NA	NA	NA	NA
Verder et al. 2011 [25]	Study 1: range, 24–32 Study 2: median (range), 27 (24–29)	NA	NA	NA	NA	NA
Lavizzari et al. 2021 [20]	Median [IQR], 31.9 [30.3; 32.9]	Median [IQR], 1,490 [1,230; 1,816]	29 (48)	59 (97)	55 (90)	Median [IQR], 9 [8; 9]

IQR, interquartile range; NA, not available; SD, standard deviation. ¹Calculated.

by cross-referencing the relevant literature and contacting authors. Eight studies matched the systematic review's predefined inclusion/exclusion criteria, including 810 infants in the final analysis. The flow diagram in Figure 1 illustrates the study selection with reasons for exclusion at each stage.

Table 1 presents the characteristics of the selected studies. The study population included preterm infants from 24 to 34 weeks of gestation enrolled in tertiary NICU settings between 1998 and 2020. Due to incomplete demographic reports, it was impossible to fully assess the baseline heterogeneity of the study population (Table 2). Nevertheless, the rate of antenatal steroid use (from 17% [18] to 97% [19, 20]) and cesarean section (from 60% [18, 19] to 90% [20]) suggests that a wide variability was likely. Whereas there was a general agreement in the definition of RDS, encompassing signs of work of breathing, hypoxia, and typical imaging findings, the definition of the reference standard for surfactant therapy varied, despite oxygen requirement as the basis. The majority of the index tests were biochemical tests on the gastric aspirate (GASP) [18, 19, 21–25] in one study only on oral aspirate [23], comprising the click test (CT) [18], the stable microbubble test (SMT) [18, 19, 21–24] and the lamellar body count (LBC) [24, 25]. A single study applied lung function testing as an index test: the FOT, evaluating respiratory mechanics, in particular, lung reactance as a marker of lung recruitment [20]. Online supplementary Table 1 describes technical details of the tests' procedure. The sample processing for SMT was reproducible across the studies. However, we found substantial differences in the analysis for: (1) microbubble (MB) size (from 10 to 25 µm); (2) main parameters considered (absolute number per unit or proportion of MB or average diameter); (3) computerized [21, 22] versus manual counts [18, 19, 23, 24]. The procedures for the LBC analysis also differed. In one study, samples were centrifuged, possibly reducing the number of lamellar bodies, and analyzed fresh or frozen using different automated cell counters [25]. By contrast, the other [24] diluted the samples, systematically froze the material, and used a different cell counter. For most studies, no private funding source was identified except for Verder et al. [25].

Quality Assessment of the Included Studies

Figure 2 synthesizes the methodological quality assessment (risk of bias and concerns regarding applicability) of the included studies following the criteria recommended by QUADAS-2 tool. Online supplementary Table 2 presents a detailed evaluation of the methodological qual-

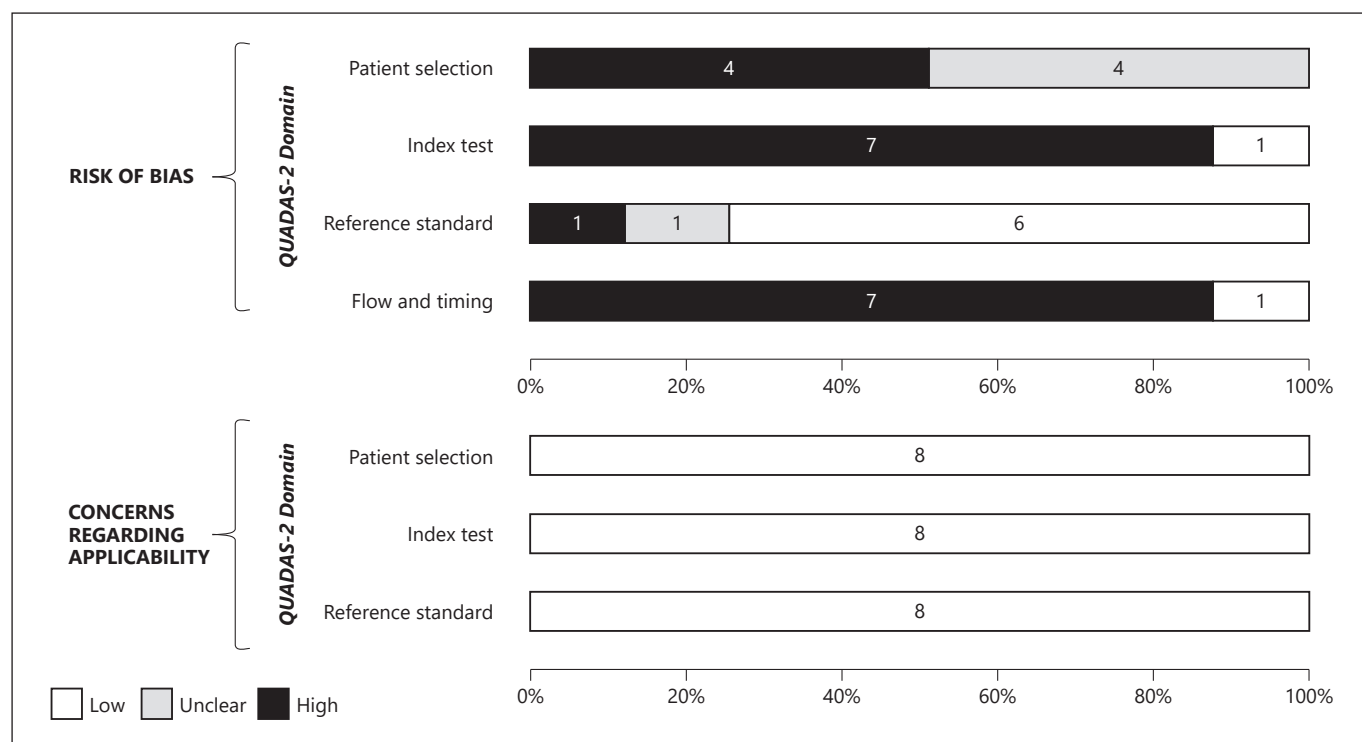


Fig. 2. QUADAS-2 results. Proportion of studies with low, high, or unclear risk of bias and concerns regarding applicability.

ity. The risk of bias was judged high for patients' selection and flow and timing. One study [24] enrolled patients consecutively until the desired sample of both target and control infants was reached. Half of the studies followed a convenience series based on operator availability [19, 20, 23, 25], and three studies did not provide sufficient details on patient selection [18, 21, 22]. All included studies had a prospective cohort trial design. The operators performing the analysis of the index tests were blinded to the patients' clinical outcome, and clinicians provided surfactant therapy (or identified the infants matching the criteria for surfactant) without knowledge of the index test results, except for one study with missing information [25]. In one trial [19], the criteria for surfactant overlapped the criteria for intubation, including apneic spells and metabolic acidosis beyond $\text{FiO}_2 \geq 0.4$, and therefore judged to classify the target condition of surfactant deficiency incorrectly. The index test was performed before the reference standard (identification of criteria for surfactant therapy). The number of patients enrolled often differed from the number of patients included. By contrast, no concerns regarding applicability emerged for any of the included studies.

Diagnostic Accuracy for Surfactant Therapy and Other Secondary Outcomes

Table 3 reports the diagnostic accuracy results of individual studies, including the identified cut-off, a 2×2 data report (true positive, false positive, false negative, true negative), and estimates for accuracy with confidence intervals, the Youden Index, and the AUC-ROC. The cut-off values varied widely within the same kind of index test. Online supplementary Figure shows paired forest plot of sensitivity and specificity of the index tests. Sensitivity analysis revealed overall a moderately high performance (sensitivity > 0.70). Two studies presented an optimal sensitivity of 1 [18, 21]. However, their specificity was poor (0.45 and 0.53, respectively). Other studies showed a sensitivity of 0.80–0.85; nevertheless, the variability – as shown by the confidence interval – was moderately high [18, 20, 21]. Based on the Youden Index, which combines sensitivity and specificity, the best performing tests were the SMT and the LBC in the GASP in Daniel et al. [24]. Finally, according to Swets classification of the AUC-ROC [26], the SMT by Fiori and the LBC by Daniel presented an excellent accuracy (> 0.90), whereas all the other studies had a moderate accuracy ($0.7 < \text{AUC}$

Table 3. Summary of diagnostic accuracy of included studies

Author, year	Technique	Index test	Cut-off	TP	FP	FN	TN	Sensitivity (95% CI) %	Specificity (95% CI) %	Youden index	N. of target patients/N. of included patients (prevalence)	PPV (95% CI), %	NPV (95% CI), %	AUC-ROC (95% CI)
Fiori et al. 2003 [18]	CT	SMT on GASP N. of MB (<15 µm)	Negative result ≤15/mm ²	22 ¹ 19 ¹	45 ¹ 13 ¹	0 ¹ 4 ¹	37 ¹ 74 ¹	1.0 (0.82–1.0) 0.83 (0.62–0.94)	0.45 (0.34–0.57) 0.85 (0.75–0.92)	0.45 ¹ 0.68 ¹	22 ¹ /104 (21%) ¹ 23 ¹ /110 (21%)	0.33 (0.22–0.45) 0.59 (0.41–0.76)	1.00 (0.88–1.0) 0.95 (0.87–0.98)	NA 0.92 (0.86–0.98)
Fiori et al. 2001 [21]	SMT on GASP N. of MB (<20 µm)	SMT on GASP Proportion of MB (<20 µm)	<8/4 mm ² <24%	12 ¹ 10 ¹	8 ¹ 6 ¹	3 ¹ 5 ¹	57 ¹ 59 ¹	0.80 (0.51–0.95) 0.67 (0.39–0.87)	0.88 (0.77–0.94) 0.91 (0.81–0.96)	0.68 0.58	15 ¹ /80 (19%) ¹	0.60 (0.36–0.81) ¹ 0.63 (0.35–0.85) ¹	0.95 (0.86–0.99) ¹ 0.92 (0.83–0.97) ¹	0.85 (0.75–0.95) 0.84 (0.72–0.96)
	SMT on GASP Average bubbles diameter	>45 µm		13 ¹	14 ¹	2 ¹	51 ¹	0.87 (0.58–0.98)	0.78 (0.65–0.87)	0.65		0.48 (0.29–0.68) ¹	0.96 (0.87–1.00) ¹	0.83 (0.71–0.95)
	SMT on GASP N. of MB if total count <60 and MB proportion if total count ≥60	<8/4 mm ² and <24%		14 ¹	10 ¹	1 ¹	55 ¹	0.93 (0.66–0.99)	0.85 (0.73–0.92)	0.78		0.58 (0.37–0.78) ¹	0.98 (0.90–1.00) ¹	NA
Verder et al. 2003 [22]	SMT on GASP N. of MB (<20 µm)	SMT on GASP Proportion of MB (<20 µm)	≤50/3.1 mm ² ≤70%	68 ¹ 68 ¹	49 ¹ 42 ¹	19 ¹ 19 ¹	52 ¹ 59 ¹	0.78 (0.69–0.87) 0.78 (0.69–0.87)	0.51 (0.41–0.60) 0.58 (0.49–0.68)	0.29 ¹ 0.36 ¹	87/188 (46%)	0.58 (0.49–0.67) 0.62 (0.53–0.71)	0.73 (0.62–0.84) 0.76 (0.66–0.86)	NA NA
	SMT on GASP Average bubbles diameter	≥24 µm		52 ¹	33 ¹	35 ¹	68 ¹	0.60 (0.49–0.70)	0.69 (0.60–0.78)	0.29 ¹		0.63 (0.52–0.74) 0.63 (0.52–0.74)	0.67 (0.58–0.76) 0.67 (0.58–0.76)	NA NA
Bhatia et al. 2013 [19]	SMT on GASP N. of MB (<15 µm)	<8/mm ²		28	19	0	21	1.0 (0.88–1.0)	0.53 (0.36–0.69)	0.53 ¹	28/68 (41%)	0.60 (0.44–0.74)	1.00 (0.84–1.00)	0.80 (0.70–0.90)
Ribeiro et al. 2019 [23]	SMT on OA N. of MB <15 µm	N. of MB ≤25/mm ²		17 ¹	6 ¹	4 ¹	37 ¹	0.81 (0.67–0.90)	0.86 (0.65–0.95)	0.67 ¹	21/64 (33%)	0.74 (0.52–0.90) ¹	0.90 (0.77–0.97) ¹	0.89 (0.81–0.97)
	SMT on GASP N. of MB <15 µm	Pre-Specified N. of MB ≤15/mm ² N. of MB ≤25/mm ²		20 ¹ 19 ¹	16 ¹ 12 ¹	1 ¹ 2 ¹	27 ¹ 31 ¹	0.96 (0.85–0.99) 0.91 (0.78–0.96)	0.62 (0.42–0.79) 0.71 (0.50–0.86)	0.79 ¹ 0.62 ¹		0.56 (0.38–0.72) ¹ 0.61 (0.42–0.78) ¹	0.96 (0.82–1.00) ¹ 0.94 (0.78–0.99) ¹	0.88 (0.80–0.96) 0.88 (0.80–0.96)
Daniel et al. 2010 [24]	SMT on GASP N. of MB (<15 µm)	<14/mm ²		32 ¹	4 ¹	2 ¹	25 ¹	0.94 (0.77–1.0)	0.86 (0.76–0.95)	0.80 ¹	34/63 (54%) as per the study design	0.89 (0.74–0.97) ¹	0.93 (0.76–1.00) ¹	0.89 (0.81–0.94)
	LBC on GASP	<42,000 lamellar bodies/µL		31 ¹	4 ¹	3 ¹	25 ¹	0.92 (0.73–1.0)	0.86 (0.77–0.95)	0.78 ¹		0.89 (0.73–0.97) ¹	0.89 (0.72–0.98) ¹	0.92 (0.85–0.99)
Verder et al. 2011 [25]	LBC on GASP	<2,000 lamellar bodies/µL <8,000 lamellar bodies/µL		40 ¹ 18	16 ¹ 13	15 ¹ 6	37 ¹ 33	0.73 (0.59–0.84) ² 0.75 (0.53–0.90)	0.70 (0.56–0.82) 0.72 (0.57–0.84)	0.43 0.47 ¹	55/108 (51%) 24/70 (34%)	0.71 (0.58–0.83) ² 0.58 (0.39–0.75)	0.71 (0.60–0.83) ² 0.85 (0.69–0.94)	NA NA
Lavizzari et al. 2021 [20]	FOT Xrs	<–33.4 cmH ₂ O's/L		11	8	2	38	0.85 (0.55–0.98)	0.83 (0.69–0.92)	0.68 ¹	13/59 (22%)	0.58 (0.34–0.80) ¹	0.95 (0.83–0.99) ¹	0.86 (0.76–0.96)

AUC-ROC, area under the curve-receiving operator characteristic; 95% CI, 95% confidence interval; CT, click test; FP, false positive; FN, false negative; GASP, gastric aspirate; LBC, lamellar body count; MB, microbubble; N., number; NPV, negative predictive value; OA, oral aspirate; PPV, positive predictive value; SMT, stable microbubble test; Xrs, respiratory reactance; TP, true positive; TN, true negative. The reported/computed PPV and NPV are relative to the prevalence of the target condition found in the respective study. ¹ Calculated values. ² CI calculated.

≤ 0.9). The positive predictive value and negative predictive value were reported only in 50% of studies and calculated considering the prevalence of surfactant need in the study population that not always reflected the incidence of surfactant need in preterm infants. Online supplementary Table 1 reports other secondary outcomes.

Conclusion

We evaluated the accuracy of biochemical and lung function tests to assess the need for surfactant. Eight studies [18–25] addressed the objective of the systematic review. They included four different tests (CT, SMT, LBC on GASP, and FOT) performed in the early postnatal period (within 3 h from birth) to evaluate the need for surfactant in preterm infants ≤ 34 weeks gestation who did not require intubation at birth. The overall quality of the evidence showed a high risk of bias but no concerns about applicability. The individual accuracy of the tests evaluated varied from moderate to high. However, the overall interpretation of the study findings might be problematic due to high heterogeneity in the population, index test, and comparators; therefore, it was not appropriate to conduct a meta-analysis, and additional evidence is required before addressing recommendations for clinical practice.

Even though we limited the study search from 2000, we were able to find research up to the present, demonstrating the compelling and long-lasting need for clinicians to tailor surfactant therapy better. The demand for an accurate method to identify infants requiring surfactant has clinical, ethical, and economic implications. Clinicians aim to find a method with high sensitivity (minimizing the false negative) not to miss any infant with surfactant deficiency. Meta-analysis demonstrated that surfactant therapy reduces critical neonatal outcomes such as mortality, the composite outcome of death, bronchopulmonary dysplasia, and air leaks [27]. Additionally, a high specificity is also mandatory (minimizing the false positive). Surfactant is an animal-derived, expensive drug, and administration through direct laryngoscopy may present several complications: trauma, bleeding, bradycardia, desaturations, changes in cerebral blood flow, pain, need for sedation and analgesia, and selective administration to the right main bronchus. The risk of drawbacks secondary to surfactant administration has led to developing less-invasive methods for application [28–30]. Efforts for better tailoring the drug indications are also imperative.

In recent years, the role of LUS for individualizing surfactant therapy has emerged [16]. The LUS is a bedside,

radiation-free tool allowing the estimation of parenchymal aeration using a semiquantitative score. A meta-analysis showed a sensitivity of 0.88 (0.80–0.93) and specificity of 0.82 (0.74–0.89) in evaluating the need for surfactant or mechanical ventilation [16]. Nevertheless, it presents limitations. It is a semiquantitative operator-dependent measure. LUS score is subject to variations in technique and interpretations, depends on the used probe, and requires specific training. LUS can assess only a portion (the more superficial one) but not the entirety of the lung and it is not a direct measure of the lung functional status. Whether evaluating the surfactant properties or lung function might be more accurate in identifying the need for exogenous supplementation remains to be determined. Also, the different techniques mentioned above might be complementary. Comparison among various techniques remains difficult because of the limitation of the currently used reference standard.

The individual test accuracy varied (from moderately low to high) among the different methods and within studies evaluating the same test. This heterogeneity might be due to differences in (1) the study population (more mature infants, higher use of antenatal steroids, variation in the definition of RDS, diverse level of care) or (2) methods for sampling, processing, and analyzing data. For instance, the studies on SMT considered different MB sizes (from 10 to 25 μm), various parameters for defining the cut-off (absolute number, proportion, or average diameter of MB for a given sampling area) [18, 19, 21–24]; sometimes the analysis was performed by an operator [18, 19, 23, 24], sometimes by a computerized program [21, 22]; some studies used fresh [19, 24] while others frozen samples [18, 21–23]. However, above all, one of the study's significant limitations was the lack of a proper gold standard. The standard references were different criteria based on oxygen demand. There is no standard reference diagnostic test for many clinical conditions, which is why new tests are actively sought [31]. Therefore, the inadequacy of current oxygenation-based criteria for surfactant therapy represents both the systematic review's rationale and important limitation. This limitation prevents the definitive interpretation of the results of the single studies. The high accuracy of the index test would only indicate a good correlation with a high FiO_2 requirement. In contrast, a test showing a lower accuracy may express a poor agreement with the current criteria but not necessarily signify an inappropriate ability to detect the need for surfactant. Nevertheless, comparing the current oxygenation-based standard reference is a mandatory first step before applying new tests for identifying surfactant-

deficient subjects in clinical practice. The new tests' overall moderate to high accuracy would encourage further research on their clinical application.

Another essential aspect related to surfactant therapy is the timing of administration. We considered only tests performed within 3 h of birth since evidence suggests this is the critical time frame to maximize surfactant benefits [15]. In an observational trial, Dargaville et al. [32] found preterm infants born between 25 and 28 weeks and 29 and 32 weeks fail NCPAP – requiring intubation and surfactant – on average at 8 and 18 h after birth, respectively. During this interval delay before surfactant administration, infants are exposed to increased muscle fatigue, risk of air leaks, and self-imposed lung tissue stress [33]. All the studies included were able to perform the test within 3 h. One study was excluded because sampling LBC within 6 h [34] and showed worse test accuracy in predicting surfactant need, possibly (but not exclusively) due to change in the GASP after application of CPAP and gastric fluid digestion. By contrast, even though all studies claimed the ability to perform a fast, point-of-care test, most of the analyses were performed afterward [18, 20–23], highlighting the “pilot” nature of the studies.

When choosing a diagnostic test, considerations of feasibility and cost-effectiveness might be relevant to clinical practice. All tests included in the review were (1) highly feasible (a very small percentage of patients presented technical issues), (2) performed at the bedside, (3) noninvasive, and (4) did not require patient cooperation; therefore, they were all suitable for the NICU environment. However, the SMT and the CT are more operator-dependent and require specific expertise (unless performed by a computer) than the LBC and FOT. Overall, all tests were low-cost and suitable for resource-limited settings except the FOT, which requires a specific mechanical ventilator. Nevertheless, the ventilator can be used as the standard respiratory support for infants in the NICU. Also, FOT is repeatable, whereas if the sample of GASP or oral aspirate is missed or critically contaminated, the analysis cannot be performed. Moreover, FOT may be more versatile, offering relevant clinical information beyond the first surfactant dose (e.g., repeated surfactant doses, optimization of the respiratory support, lung function follow-up) [35–38].

We acknowledge the limitations of the current systematic review. First, there was a risk of publication bias due to both English language restriction and a tendency to publish studies with positive findings rather than nonsignificant results. We extended the search to trial registers to minimize this issue. Second, the need for defining rig-

orous inclusion criteria might have excluded studies providing relevant information to the topic, such as the role of the surfactant adsorption test [39] or the spectroscopy of GASP [40–42]. Third, the risk of bias was judged high, even though primarily due to the population selection (convenience rather than consecutive series) and time and flow domain (not all subjects were included in the analysis). The limitations due to the study design, standard reference, and heterogeneity have already been addressed.

Nevertheless, the moderate to high accuracy of the new tests, supported by a high prevalence of blinding (87.5%), should encourage future trial designs. Few studies further evaluated the tests included in the review [11, 43]. Whether the new tests alone or in combination with oxygenation-based criteria or imaging tests may improve important clinical neonatal outcomes (including the need for mechanical ventilation, duration of respiratory support, rate of chronic lung disease, long-term respiratory morbidity) deserves further investigation. Given the limitation of the reference standard for surfactant administration, randomized-controlled trials are advocated to compare surfactant administration performed following alternative criteria to the one currently in use.

In conclusion, the systematic review appraised four different tests for predicting surfactant therapy, performed on 810 infants that showed individual moderate to high accuracy. Nevertheless, due to the high heterogeneity, risk of bias, and inconsistency among the studies, the strength of evidence was judged low. Further research is needed before the new approaches explored in our study for tailoring surfactant therapy can be recommended in clinical practice.

Acknowledgment

We acknowledge Prof Susan Smith for her precious encouragement and support to this project.

Statement of Ethics

An ethics statement is not applicable because this study is based exclusively on published literature.

Conflict of Interest Statement

Anna Lavizzari was a consultant for Getinge, Chiesi S.p.A, and Vyaire and received travel grants from Vapotherm and Fisher & Paykel. Chiara Veneroni has nothing to declare.

Funding Sources

The study was partially funded by Italian Ministry of Health.

Data Availability Statement

All data generated or analyzed during this study are included in this article. Further inquiries can be directed to the corresponding author.

Author Contributions

Anna Lavizzari and Chiara Veneroni conceived the study design, realized the data extraction, performed the data analysis, and critically interpreted the data; Anna Lavizzari drafted the manuscript and Chiara Veneroni critically revised it.

References

- 1 Sweet DG, Carnielli V, Greisen G, Hallman M, Ozek E, Te Pas A, et al. European consensus guidelines on the management of respiratory distress syndrome: 2019 update. *Neonatology*. 2019;115(4):432–50.
- 2 Banerjee S, Fernandez R, Fox GF, Goss KCW, Mactier H, Reynolds P, et al. Surfactant replacement therapy for respiratory distress syndrome in preterm infants: United Kingdom national consensus. *Pediatr Res*. 2019; 86(1):12–4.
- 3 Ng EH, Shah V. Guidelines for surfactant replacement therapy in neonates. *Paediatr Child Health*. 2021;26(1):35–41.
- 4 Polin RA, Carlo WA, Papile LA, Polin RA, Carlo W, Tan R, et al. Surfactant replacement therapy for preterm and term neonates with respiratory distress. *Pediatrics*. 2014;133(1): 156–63.
- 5 Macklem PT, Proctor DF, Hogg JC. The stability of peripheral airways. *Respir Physiol*. 1970;8(2):191–203.
- 6 Perez-Gil J, Weaver TE. Pulmonary surfactant pathophysiology: current models and open questions. *Physiology*. 2010;25(3):132–41.
- 7 Stocks J. Respiratory physiology during early life. *Monaldi Arch Chest Dis*. 1999;54(4): 358–64.
- 8 de Boode WP, van der Lee R, Horsberg Eriksen B, Nestaas E, Dempsey E, Singh Y, et al. The role of neonatologist performed echocardiography in the assessment and management of neonatal shock. *Pediatr Res*. 2018; 84(S1):57–67.
- 9 De Luca D, Autilio C, Pezza L, Shankar-Aguilera S, Tingay DG, Carnielli VP. Personalized medicine for the management of RDS in preterm neonates. *Neonatology*. 2021;118(2): 127–38.
- 10 Hentschel R, Bohlin K, van Kaam A, Fuchs H, Danhaive O. Surfactant replacement therapy: from biological basis to current clinical practice. *Pediatr Res*. 2020;88(2):176–83.
- 11 Verder H, Ebbesen F, Fenger-Grøn J, Henriksen TB, Andreasson B, Bender L, et al. Early surfactant guided by lamellar body counts on gastric aspirate in very preterm infants. *Neonatology*. 2013;104(2):116–22.
- 12 McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies the PRISMA-DTA statement. *JAMA*. 2018;319(4):388–96.
- 13 Leeftang MMG, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149(12): 889–98.
- 14 Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372: n71.
- 15 Bahadue FL, Soll R. Early versus delayed selective surfactant treatment for neonatal respiratory distress syndrome. *Cochrane Database Syst Rev*. 2012;11(11):CD001456.
- 16 Razak A, Faden M. Neonatal lung ultrasonography to evaluate need for surfactant or mechanical ventilation: a systematic review and meta-analysis. *Arch Dis Child Fetal Neonatal Ed*. 2019;105(2):164–71.
- 17 Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32–5.
- 18 Fiori HH, Varela I, Justo AL, Fiori RM. Stable microbubble test and click test to predict respiratory distress syndrome in preterm infants not requiring ventilation at birth. *J Perinat Med*. 2003;31(6):509–14.
- 19 Bhatia R, Morley CJ, Argus B, Tingay DG, Donath S, Davis PG. The stable microbubble test for determining continuous positive airway pressure (CPAP) success in very preterm infants receiving nasal CPAP from birth. *Neonatology*. 2013;104(3):188–93.
- 20 Lavizzari A, Veneroni C, Beretta F, Ottaviani V, Fumagalli C, Tossici M, et al. Oscillatory mechanics at birth for identifying infants requiring surfactant: a prospective, observational trial. *Respir Res*. 2021;22(1):314–1.
- 21 Fiori HH, Linderholm B, Fiori RM, Robertson B. Computerized image analysis of bubbles in gastric aspirate for prediction of respiratory distress syndrome. *Acta Paediatr*. 2001; 90(12):1402–4.
- 22 Verder H, Ebbesen F, Linderholm B, Robertson B, Eschen C, Arrøe M, et al. Prediction of respiratory distress syndrome by the microbubble stability test on gastric aspirates in newborns of less than 32 weeks' gestation. *Acta Paediatr*. 2003;92(6):728–33.
- 23 Ribeiro MADS, Fiori HH, Luz JH, Garcia PCR, Fiori RM. Rapid diagnosis of respiratory distress syndrome by oral aspirate in premature newborns. *J Pediatr*. 2019;95(4):489–94.
- 24 Daniel IWBD, Fiori HH, Piva JP, Munhoz TP, Nectoux AV, Fiori RMH. Lamellar body count and stable microbubble test on gastric aspirates from preterm infants for the diagnosis of respiratory distress syndrome. *Neonatology*. 2010;98(2):150–5.
- 25 Verder H, Ebbesen F, Brandt J, Dahl M, Esberg G, Eschen C, et al. Lamellar body counts on gastric aspirates for prediction of respiratory distress syndrome. *Acta Paediatr*. 2011; 100(2):175–80.
- 26 Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988;240(4857):1285–93.
- 27 Soll R, Özek E. Prophylactic animal derived surfactant extract for preventing morbidity and mortality in preterm infants. *Cochrane Database Syst Rev*. 1997;4:CD000511.
- 28 Kribs A, Roll C, Göpel W, Wieg C, Groneck P, Laux R, et al. Nonintubated surfactant application vs. conventional therapy in extremely preterm infants: a randomized clinical trial. *JAMA Pediatr*. 2015;169(8):723–30.
- 29 Abdel-Latif ME, Davis PG, Wheeler KI, De Paoli AG, Dargaville PA. Surfactant therapy via thin catheter in preterm infants with or at risk of respiratory distress syndrome. *Cochrane Database Syst Rev*. 2021;5(5): CD011672.
- 30 Gaertner VD, Thomann J, Bassler D, Ruegger CM. Surfactant nebulization to prevent intubation in preterm infants: a systematic review and meta-analysis. *Pediatrics*. 2021;148(5): e2021052504.
- 31 Greenhalgh T. How to read a paper: papers that report diagnostic or screening tests. *BMJ*. 1997;315(7107):540–3.

- 32 Dargaville PA, Aiyappan A, De Paoli AG, Dalton RGB, Kuschel CA, Kamlin CO, et al. Continuous positive airway pressure failure in preterm infants: incidence, predictors and consequences. *Neonatology*. 2013;104(1):8–14.
- 33 Brochard L, Slutsky A, Pesenti A. Mechanical ventilation to minimize progression of lung injury in acute respiratory failure. *Am J Respir Crit Care Med*. 2017;195(4):438–42.
- 34 Raschetti R, Centorrino R, Letamendia E, Benachi A, Marfaing-Koka A, De Luca D. Estimation of early life endogenous surfactant pool and CPAP failure in preterm neonates with RDS. *Respir Res*. 2019;20(1):75–8.
- 35 Veneroni C, Wallström L, Sindelar R, Dellacà RL. Oscillatory respiratory mechanics on the first day of life improves prediction of respiratory outcomes in extremely preterm newborns. *Pediatr Res*. 2019;85(3):312–7.
- 36 Wallström L, Veneroni C, Zannin E, Dellacà RL, Sindelar R. Forced oscillation technique for optimising PEEP in ventilated extremely preterm infants. *Eur Respir J*. 2020;55(5):1901650.
- 37 Raffaeli G, Veneroni C, Ghirardello S, Lavizzari A, Passera S, Mosca F, et al. Role of lung function monitoring by the forced oscillation technique for tailoring ventilation and weaning in neonatal ECMO: new insights from a case report. *Front Pediatr*. 2018;6:332.
- 38 Veneroni C, Mercadante D, Lavizzari A, Colnaghi M, Mosca F, Dellacà RL. Changes in respiratory mechanics at birth in preterm infants: a pilot study. *Pediatr Pulmonol*. 2020;55(7):1640–5.
- 39 Autilio C, Echaide M, Benachi A, Marfaing-Koka A, Capoluongo ED, Pérez-Gil J, et al. A noninvasive surfactant adsorption test predicting the need for surfactant therapy in preterm infants treated with continuous positive airway pressure. *J Pediatr*. 2017;182:66–73.e1.
- 40 Verder H, Heiring C, Clark H, Sweet D, Jessen TE, Ebbesen F, et al. Rapid test for lung maturity, based on spectroscopy of gastric aspirate, predicted respiratory distress syndrome with high sensitivity. *Acta Paediatr*. 2017;106(3):430–7.
- 41 Schousboe P, Verder H, Jessen TE, Heiring C, Bender L, Ebbesen F, et al. Predicting respiratory distress syndrome at birth using fast test based on spectroscopy of gastric aspirates. 1. Biochemical part. *Acta Paediatr*. 2020;109(2):280–4.
- 42 Heiring C, Verder H, Schousboe P, Jessen TE, Bender L, Ebbesen F, et al. Predicting respiratory distress syndrome at birth using a fast test based on spectroscopy of gastric aspirates: 2. Clinical part. *Acta Paediatr*. 2020;109(2):285–90.
- 43 Fiori HH, Fritscher CC, Fiori RM. Selective surfactant prophylaxis in preterm infants born at < or =31 weeks' gestation using the stable microbubble test in gastric aspirates. *J Perinat Med*. 2006;34(1):66–70.