Received: 21 October 2020 | Revised: 23 December 2020 | Accepted: 7 January 2021

DOI: 10.1002/ppul.25279

ORIGINAL ARTICLE: NEONATAL LUNG DISEASE

PEDIATRIC PULMONOLOGY WILEY

Development of a severity scale to assess chronic lung disease after extremely preterm birth

Hugh M. O'Brodovich MD[1] | Robin Steinhorn MD[2] | Robert M. Ward MD[3] | Mikko Hallman MD[4] | Ethan J. Schwartz BS[5] | Magdalena Vanya PhD[6] | Ellen M. Janssen PhD[5] | Alexandra Mangili MD, MPH[7] | Linda Han MD, MPH[8] | Sujata P. Sarda PhD[9]

[1]Department of Pediatrics, Stanford University School of Medicine, Stanford, California, USA

[2]Department of Pediatrics, University of California San Diego, San Diego, California, USA

[3]Department of Pediatrics, University of Utah, Salt Lake City, Utah, USA

[4]Oulu University Hospital, University of Oulu, Oulu, Finland

[5]Patient Centered Outcomes, ICON, Gaithersburg, Maryland, USA

[6]Patient Centered Outcomes, ICON, South San Francisco, California, USA

[7]Global Clinical Development, Rare Metabolic Diseases, Takeda, Zurich, Switzerland

[8]Global Clinical Development, Rare Metabolic Diseases, Takeda, Cambridge, Massachusetts, USA

[9]Global Evidence and Outcomes, Takeda, Lexington, Massachusetts, USA

Correspondence
Sujata P. Sarda, PhD, Global Evidence and Outcomes, Takeda, 300 Shire Way, Lexington, MA 02421, USA.
Email: jaysujata@gmail.com

Funding information
Takeda

Abstract

Objective: Chronic lung disease of prematurity (CLDP) is a frequent complication of prematurity. We aimed to identify what clinicians believe are the most important factors determining the severity of CLDP in extremely preterm infants (<28 weeks gestational age) after discharge from the neonatal intensive care unit (NICU) through 12 months corrected age (CA), and to evaluate how these factors should be weighted for scoring, to develop a CLDP severity scale.

Study design: Clinicians completed a three-round online survey utilizing Delphi methodology. Clinicians rated the importance of various factors used to evaluate the severity of CLDP, from 0 (*not at all important*) to 10 (*very important*) for the period between discharge home from the NICU and 12 months CA. Fourteen factors were considered in Round 1; 13 in Rounds 2 and 3. The relative importance of factors was explored via a set of 16 single-profile tasks (i.e., hypothetical patient profiles with varying CLDP severity levels).

Results: Overall, 91 clinicians from 11 countries who were experienced in treating prematurity-related lung diseases completed Round 1; 88 completed Rounds 2 and 3. Based on Round 3, the most important factors in determining CLDP severity were mechanical ventilation (mean absolute importance rating, 8.89), supplemental oxygen ≥2 L/min (8.49), rehospitalizations (7.65), and supplemental oxygen <2 L/min (7.56). Single-profile tasks showed that supplemental oxygen had the greatest impact on profile classification.

Conclusion: The most important factors for clinicians assigning CLDP severity during infancy were mechanical ventilation, supplemental oxygen ≥2 L/min, and respiratory-related rehospitalizations.

KEYWORDS

bronchopulmonary dysplasia, chronic lung disease, Delphi panel, extremely premature, preterm

1 | INTRODUCTION

Bronchopulmonary dysplasia (BPD) is diagnosed at 36 weeks postmenstrual age (PMA) and is a frequent complication of extremely premature birth (<28 weeks gestational age). The reported global incidence of BPD is 10%–89%.[1] Across regions, reported estimates range from 18%–89% in North America, 10%–73% in Europe, 30%–62% in Oceania, and 18%–82% in Asia.[1–3] BPD is associated with high healthcare costs and an increased risk of mortality and rehospitalization among preterm infants during the first year of life.[4,5] BPD is also a leading cause of chronic lung disease of prematurity (CLDP),[6] which can manifest in patients with or without a prior diagnosis of BPD (using current criteria). No new drugs for the prevention and treatment of BPD have been approved in recent decades.[7] Recombinant human insulin-like growth factor 1 complexed with its binding protein (rhIGF-1/rhIGFBP-3) is currently under investigation for the prevention of complications of prematurity among extremely preterm infants. A 2019 phase 2 trial evaluating rhIGF-1/rhIGFBP-3 supplementation in extremely preterm infants reported substantial reductions in the incidence of severe BPD.[8] A phase 2b trial is currently ongoing to evaluate if rhIGF-1/rhIGFBP-3 can reduce the burden of CLDP in extremely preterm infants (NCT03253263).

Currently, the assessment of long-term pulmonary morbidity is associated with the diagnosis of BPD and is driven by the amount of respiratory support (e.g., the requirement for oxygen, continuous positive airway pressure [CPAP], and mechanical ventilation) administered at 36 weeks PMA or at the time of discharge from the neonatal intensive care unit (NICU).[9–11] This approach focuses on short-term, rather than long-term, morbidity and is heavily impacted by variations in NICU respiratory support practices.[10] In addition, infants who do not meet the diagnostic criteria for BPD may at a later date exhibit clinically important respiratory disease.[12] For these reasons, at least in part, the US Food and Drug Administration now requires a long-term assessment of these infants when determining the effectiveness of respiratory interventions.[13] BPD as currently defined is considered an imperfect biomarker for long-term pulmonary outcomes[14] and is not accepted as an endpoint for regulatory decision making.

Based on a survey of North American academic health science center-based pulmonologists, Gage et al.[15] developed a chronic lung disease severity score, which assessed CLDP severity in very low birth weight infants at 4–9 months corrected age (CA). The objective of the current study was to build on the work of Gage et al. by identifying additional important factors believed by physicians to be measured for the severity of CLDP during the months following extremely preterm birth and to evaluate how these factors should be weighted for scoring in a CLDP severity scale (CLDPSS). The CLDPSS is being used as a secondary outcome measure in the global phase 2b clinical trial, being conducted by Takeda, which is evaluating if rhIGF-1/rhIGFBP-3 can decrease the morbidity of CLDP through 12 months CA (NCT03253263).

2 | MATERIALS AND METHODS

2.1 | Participants

Three Delphi rounds were conducted; participants on the panels were required to have (1) a general medical license/registration; (2) board certification, or equivalent, in neonatology, pediatric pulmonology, and/or pediatrics; (3) ≥2 years' post-fellowship/residency experience treating prematurity-related chronic lung disease (i.e., CLDP); (4) currently treating at least two premature infants with CLDP per year in the outpatient setting; and (5) fluency in English. Clinicians were recruited by a third-party market research vendor (Global Perspectives; Norwich, United Kingdom), via online physician research panels. Potentially eligible participants were emailed a secure link to complete an initial screening questionnaire to determine eligibility. In addition to the screening questions, potential participants were directed to an online "consent" screen, which included study information and participant requirements, with an option to opt in to participate in the study. Eligible participants could then proceed to complete Round 1 of the survey. Survey respondents who completed all three rounds of the survey received $300 in honoraria. IRB approval was not required because this was a non-interventional study.

2.2 | Variables

In the Round 1 survey, to reach a consensus on the importance of different variables in the determination of CLDP severity at 1 year CA, clinicians considered a comprehensive set of prespecified factors related to the severity of chronic lung disease during infancy. The factors selected for assessment were based on the previous work by Gage et al.[15] and were supplemented with factors identified by four clinical experts (HMO'B, RS, RMW, and MH).

Fourteen factors were evaluated in Round 1 of the survey; 13 factors were evaluated in Rounds 2 and 3. Intermittent administration of pulmonary vasodilator was removed following Round 1, based on clinical expert feedback (Table 1). The factors considered in the Delphi Rounds included the use of home mechanical ventilation, including bilevel positive airway pressure (BiPAP) and nasal intermittent positive pressure ventilation (NIPPV); supplemental oxygen (thresholds of <2 L/min or ≥2 L/min, the latter of which includes CPAP); respiratory-related rehospitalizations after NICU discharge; respiratory-related emergency department (ED) visits without hospitalization; and use of pulmonary medications (i.e., bronchodilators, corticosteroids [inhaled and systemic], diuretics, and pulmonary vasodilators [intermittent administration only included in Round 1]).

TABLE 1 Factors considered in the Delphi survey

Home mechanical ventilation, including BiPAP and NIPPV
Supplemental oxygen via nasal cannula at ≥2 L/min, including CPAP
Supplemental oxygen via nasal cannula at <2 L/min
Respiratory-related rehospitalization after NICU discharge
Respiratory-related ED visits without hospitalization
Daily (≥3 days/week) administration of a bronchodilator
Daily (≥3 days/week) administration of inhaled corticosteroid
Daily (≥3 days/week) administration of diuretic
Daily (every day) administration of pulmonary vasodilator[a]
Intermittent administration of a bronchodilator
Intermittent administration of inhaled corticosteroid
Intermittent administration of systemic corticosteroid
Intermittent administration of diuretic
Intermittent administration of pulmonary vasodilator[a]

Abbreviations: BiPAP, bilevel positive airway pressure; CPAP, continuous positive airway pressure; ED, emergency department; NICU, neonatal intensive care unit; NIPPV, nasal intermittent positive pressure ventilation.

[a]Included in Round 1, but not included in Rounds 2 and 3, based on feedback from clinical expert consultants.

2.3 | Iterative surveys

2.3.1 | Delphi survey

This study used a modified three-round online Delphi survey, designed to explore the most important and relevant factors in determining the severity of lung disease among extremely preterm infants after discharge from the NICU. The Delphi approach was adopted because it represents an established method of obtaining an expert opinion and evaluating the degree of consensus on a given topic.[16] The Delphi method is a structured communication technique that involves participants (in this case, selected clinicians) who answer a questionnaire anonymously in an iterative manner after being provided with a summary of group responses.[16]

In Round 1 of the Delphi survey, clinicians rated the importance of respiratory-related factors used to evaluate the severity of CLDP, from 0 (*not at all important*) to 10 (*very important*) for the period between discharge home from the NICU and 12 months CA. Clinicians also ranked the relative importance of factors (i.e., relative to others) in determining severity. Clinicians had the opportunity to identify and rate additional attributes (not already included) via free-text responses. To facilitate response, clinicians were first asked to select and rank-order the five most important factors, then to select and rank-order the next five most important, and so on. In Rounds 2 and 3, clinicians were presented with anonymized aggregate results from the previous round and were given the opportunity to accept or change their prior response. A fourth Delphi round was not required because sufficient consensus was achieved in Round 3.

2.3.2 | Discrete choice experiment

We also conducted an exploratory discrete choice experiment (DCE) to explore the relative importance and weighting of attributes included in the survey. A DCE is a methodology used to elicit preferences and evaluate the relative importance of aspects related to health outcomes among participants (in this case, selected clinicians).[17] In a DCE, participants are presented with a series of hypothetical clinical profiles that are composed of a fixed set of treatment characteristics, which are presented using systematically varied levels for each treatment characteristic. Participants are then asked to make a choice for each hypothetical profile.

In the current study, the DCE was conducted through a set of 16 single-profile choice tasks based on hypothetical patient profiles with varying CLDP severity levels, representative of an infant born extremely preterm aged 12 months CA who had been diagnosed with BPD at 36 weeks PMA. Clinicians were presented with eight respiratory-related treatment attributes with a possible two to four levels each (Table S1). Clinicians were then asked to rate the severity of CLDP in relation to the infant profiles, based solely on recent respiratory treatment utilization presented in the profile (see Table S2, for example).

A D-efficient experimental design was generated in Ngene 1.1.2 (ChoiceMetrics) to systematically vary the attribute levels and generate 12 single-profile choice tasks. A D-efficient design is commonly used to maximize the statistical efficiency in measuring the main effects.[17] Clinical expert input was used to ensure that these profiles were clinically possible. Four additional choice tasks, developed by a clinical expert (HMO'B), were included within the survey to assess internal validity. The DCE was conducted in full in Round 2; however, in Round 3, only the four choice tasks developed by the clinical expert (HMO'B) were included.

2.4 | Data analysis

Analyses were performed using SAS version 9.4 (SAS Institute Inc) or R version 3.3.3 or higher (R Core Team).

Following each round, descriptive analyses were conducted for each survey question. If ≥75% of clinicians indicated that an attribute had no importance (i.e., a rating of 0; overall, at discharge, or at 12 months CA), it was excluded from future survey rounds. In Round 1, each attribute included the percentage of "no importance" ratings and the mean, standard deviation, and interquartile range (IQR) of the importance ratings. For absolute importance ratings (or weight; on a scale from 0 to 10), a higher score indicated greater importance. For relative importance rankings, a lower score (e.g., a ranking of 1) indicated greater importance; relative rankings were therefore adjusted such that a higher score indicated greater importance. Overall

importance scores were calculated by multiplying the absolute importance ratings by the adjusted relative importance rankings.

The first 12 experimental design–generated DCE choice tasks were analyzed descriptively and via multinomial logistic regression. The final four clinical expert–generated DCE choice tasks were analyzed by examining choice frequencies for each severity indication. The results are presented as relative risk ratios and predicted probabilities, which represent the effect of each attribute level on severity classification, independently from the other attributes.

The outcome variable was modeled as a choice among the four severity classifications (asymptomatic/minimal, mild, moderate, or severe), with the asymptomatic/minimal classification as the baseline category. The independent, predictive variables were the attribute levels included in the choice profiles, treated as dummy variables. For each attribute, the least severe indication was treated as the reference category. Relative risk ratios were estimated to show how each attribute level affected the choice of severity indication. Predicted probabilities were calculated to explore the predicted probability of selecting each severity indication at each level of the different attributes, holding the other attributes at their means or, alternatively, at the lowest (least severe) levels. In Rounds 2 and 3, the severity assessments for the four additional clinician-generated profiles were summarized descriptively.

2.5 | Consensus

In Rounds 2 and 3 of the survey, the between-clinician consensus was assessed by examining the IQR of the importance ratings and rankings.[16] Based on a 2012 review, an IQR <2 was considered good consensus when assessing responses to a scale of 0–10.[16] After Round 3 of the survey, an analysis of variance (ANOVA) for repeated measures with equal variance, including a test of homogeneity of variance, was run for all remaining factors. We studied the consensus between Round 2 and 3 responses using ANOVA. A separate test was performed for each factor. The ANOVA included the absolute importance rating values for the factor as the dependent variable and clinicians as the independent variable. An F value ≥4 for the clinician variable indicated that the between-clinician variability was substantially larger than the within-clinician variability.

2.6 | Weighting and scoring

The importance scores of attributes included in the final survey round were used to develop the final attribute weights on an integer scale of 0–100. We identified the factors associated with the most severe chronic lung disease; these mutually exclusive "worst" factors were set to correspond to a total score of 100 and included home mechanical ventilation, at least one respiratory-related rehospitalization, at least one respiratory-related ED visit without hospitalization, daily pulmonary vasodilator use, daily diuretic use, daily bronchodilator use, intermittent systemic corticosteroid use, and daily inhaled corticosteroid use (Table S3). The absence of any factors was set to correspond to a total score of 0.

3 | RESULTS

3.1 | Participant characteristics

A total of 91 participants (51 pediatric pulmonologists, 20 pediatricians, and 20 neonatologists) completed Round 1 of the CLDP severity survey; 88 of these participants completed Rounds 2 and 3. Participants resided in 11 countries across North America, Europe, Asia, and South America (Table 2). They had a mean of 16 years' post-fellowship/residency experience in caring for premature infants with CLDP and treated a mean of ~32 premature infants with CLDP in an outpatient setting per year. When asked if they had co-authored peer-reviewed publications, spoken at conferences, or acted as a principal investigator in a neonatal clinical trial pertaining to CLDP, 35.2% answered yes.

TABLE 2 Characteristics of CLDP survey participants

Characteristic	Round 1 (N = 91)	Rounds 2 and 3 (N = 88)
Clinician type, *n* (%)[a]		
Pediatric pulmonologist	51 (56.0)	50 (56.8)
Pediatrician	20 (22.0)	19 (21.6)
Neonatologist	20 (22.0)	19 (21.6)
Country of practice, *n* (%)		
United States	22 (24.2)	22 (25.0)
Canada	15 (16.5)	14 (15.9)
Germany	10 (11.0)	9 (10.2)
United Kingdom	9 (9.9)	9 (10.2)
Italy	9 (9.9)	8 (9.1)
Spain	8 (8.8)	8 (9.1)
South Korea	5 (5.5)	5 (5.7)
France	4 (4.4)	4 (4.5)
Japan	4 (4.4)	4 (4.5)
Brazil	3 (3.3)	3 (3.4)
Mexico	2 (2.2)	2 (2.3)
Years of experience, mean (*SD*)[b]	15.6 (7.3)	15.6 (7.4)
Patients treated per year[c]		
Median (Q1–Q3)	30.0 (10.0–50.0)	27.5 (10.0–50.0)
Range	3.0–95.0	3.0–95.0

Abbreviations: CLDP, chronic lung disease of prematurity; *SD*, standard deviation; Q, quarter.

[a]Some participants had multiple specialties.

[b]Years of experience in caring for premature infants with CLDP.

[c]Premature infants with CLDP treated per year on an outpatient basis.

3.2 | Delphi survey

3.2.1 | Round 1

In the first-round survey, the most important factors, in terms of mean absolute importance (on a scale from 0 [*not at all important*] to 10 [*very important*]), were determined to be home mechanical ventilation (8.38), supplemental oxygen at ≥2 L/min (8.14), rehospitalizations (7.97), and ED visits without hospitalization (7.82) (Figure 1). The factors ranked most important relative to the others were home mechanical ventilation (3.48), supplemental oxygen ≥2 L/min (3.70), rehospitalizations (5.38), and supplemental oxygen <2 L/min (7.07) (Figure S1).

All attributes were identified to be of at least some importance (i.e., score >0) by ≥25% of the clinician sample. Therefore, no predefined factors were removed in the subsequent round of the survey, except for PRN use of pulmonary vasodilators, which was considered clinically infeasible and removed following clarification by the clinical expert consultants. Likewise, when grouping the free-text responses, no responses were endorsed by more than 25% of the sample. Following an additional review of the free-text responses by the clinical expert consultants, it was decided that no responses warranted the inclusion of additional attributes in the Round 2 survey, and the free-text option was removed in subsequent rounds of the survey.

3.2.2 | Round 2

In the second-round survey, the most important factors in terms of mean absolute importance were home mechanical ventilation (8.60), supplemental oxygen at ≥2 L/min (8.47), rehospitalizations (7.76), and supplemental oxygen at <2 L/min (7.63) (Figure 1). The same four factors were ranked most important relative to the others.

When assessing the IQR of the absolute importance ratings for each factor, most factors had an IQR of 1.00, with a maximum of 2.00, indicating that there was fairly good consensus in the absolute importance ratings across the sample (both overall and within clinician groups).

3.2.3 | Round 3

In Round 3, the factors ranked most important in terms of mean absolute importance were home mechanical ventilation (8.89),

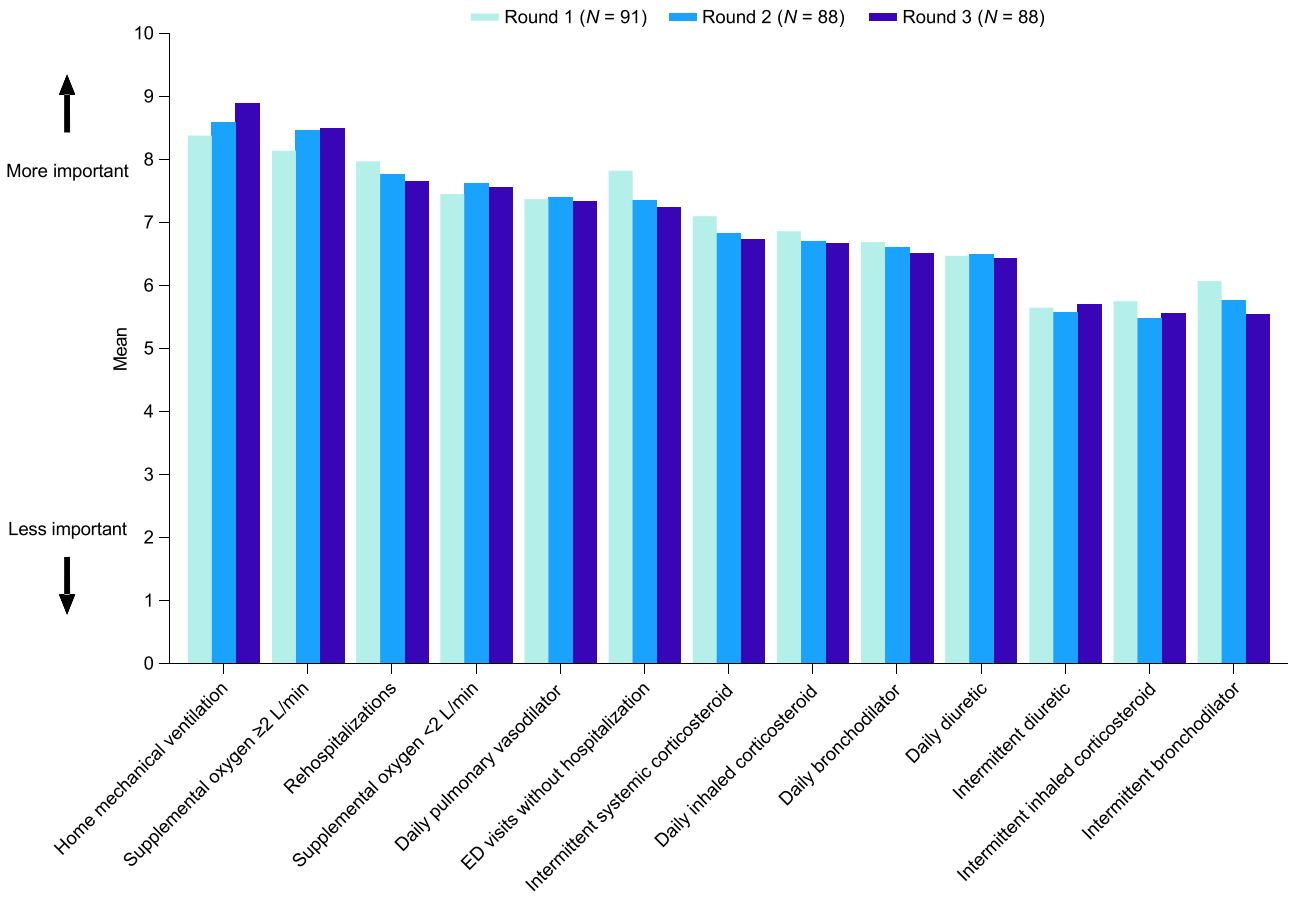

FIGURE 1 Mean absolute importance of factors in evaluating CLDP severity. Each factor was rated individually on a scale of 0 (*not at all important*) to 10 (*very important)* and is presented in order of the Round 3 results (most to least important). CLDP, chronic lung disease of prematurity; ED, emergency department [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 ANOVA results comparing Rounds 2 and 3 absolute importance ratings among clinicians ($N = 88$)

Attribute	Absolute importance rating[a], mean (*SD*)		*F* value	*p* value
	Round 2	Round 3		
Home mechanical ventilation	8.60 (1.62)	8.89 (1.22)	4.42	.0385*
Supplemental oxygen ≥ 2 L/min	8.47 (1.47)	8.49 (1.49)	0.03	.8630
Supplemental oxygen <2 L/min	7.63 (1.41)	7.56 (1.13)	0.23	.6337
Respiratory-related rehospitalizations	7.76 (1.31)	7.65 (1.04)	0.70	.4045
Daily administration of pulmonary vasodilator	7.40 (1.26)	7.34 (1.20)	0.17	.6783
ED visits without hospitalization	7.36 (1.32)	7.24 (1.17)	0.80	.3737
Daily administration of diuretics	6.50 (1.40)	6.43 (1.10)	0.26	.6123
Intermittent administration of systemic corticosteroid	6.83 (1.34)	6.73 (1.08)	0.47	.4969
Daily administration of inhaled corticosteroid	6.70 (1.20)	6.67 (1.24)	0.05	.8190
Daily administration of a bronchodilator	6.60 (1.22)	6.51 (1.12)	0.45	.5051
Intermittent administration of diuretics	5.58 (1.43)	5.70 (1.07)	0.67	.4141
Intermittent administration of a bronchodilator	5.76 (1.11)	5.55 (1.32)	3.11	.0815
Intermittent administration of inhaled corticosteroid	5.48 (1.61)	5.56 (1.36)	0.20	.6573

Abbreviations: ANOVA, analysis of variance; ED, emergency department.

[a]As rated on a scale from 0 (*not at all important*) to 10 (*very important*).

*Statistically significant at $\alpha = .05$ level.

supplemental oxygen ≥2 L/min (8.49), rehospitalizations (7.65), and supplemental oxygen <2 L/min (7.56); the same four factors were also ranked most important in terms of relative importance to the others (Figures 1 and S1). When assessing the IQR of the absolute importance ratings for each factor, all factors had an IQR of 1.00, with the exception of home mechanical ventilation, which had an IQR of 2.00, indicating good consensus in the absolute importance ratings across the sample. ANOVA results comparing the Rounds 2 and 3 absolute importance ratings showed no statistically significant differences in ratings between rounds, except for mechanical ventilation ($F = 4.42$; $p = .0385$) (Table 3). Combined IQR and ANOVA results indicated that good consensus was achieved after the Round 3 survey.

3.3 | Discrete choice experiment findings

The exploratory DCE results from Round 2 showed that supplemental oxygen had the largest influence on profile classification. If a profile described the need for "mechanical ventilation" rather than the reference level of "no supplemental oxygen," the predicted probability that it would be classified as asymptomatic/minimal lung disease decreased from 0.30 to 0.06, and the predicted probability that it would be classified as severe lung disease increased from 0 to 0.28 (Figure 2). If a profile described the need for daily pulmonary vasodilator use rather than the reference level of "none," the predicted probability that it would be classified as asymptomatic/minimal lung disease decreased from 0.30 to 0.17, and the predicted probability that it would be classified as moderate lung disease increased from 0.15 to 0.24. Overall, the results of the exploratory DCE supported the Delphi findings on the importance of factors.

3.4 | CLDPSS instrument scoring

Based on the Round 3 mean overall importance values for each factor (Figure 3), the factor weights were calculated. For each variable, the *worst* level was selected (e.g., mechanical ventilation for the supplemental oxygen variable), and the sum of these important values was calculated and then rescaled to 100; the individual final importance values were then rescaled down to the final weights for each factor (Table S3). These were home mechanical ventilation, 23.4; at least one respiratory-related rehospitalization, 15.1; daily pulmonary vasodilator use, 13.6; at least one ED visit without hospitalization, 12.2; daily use of

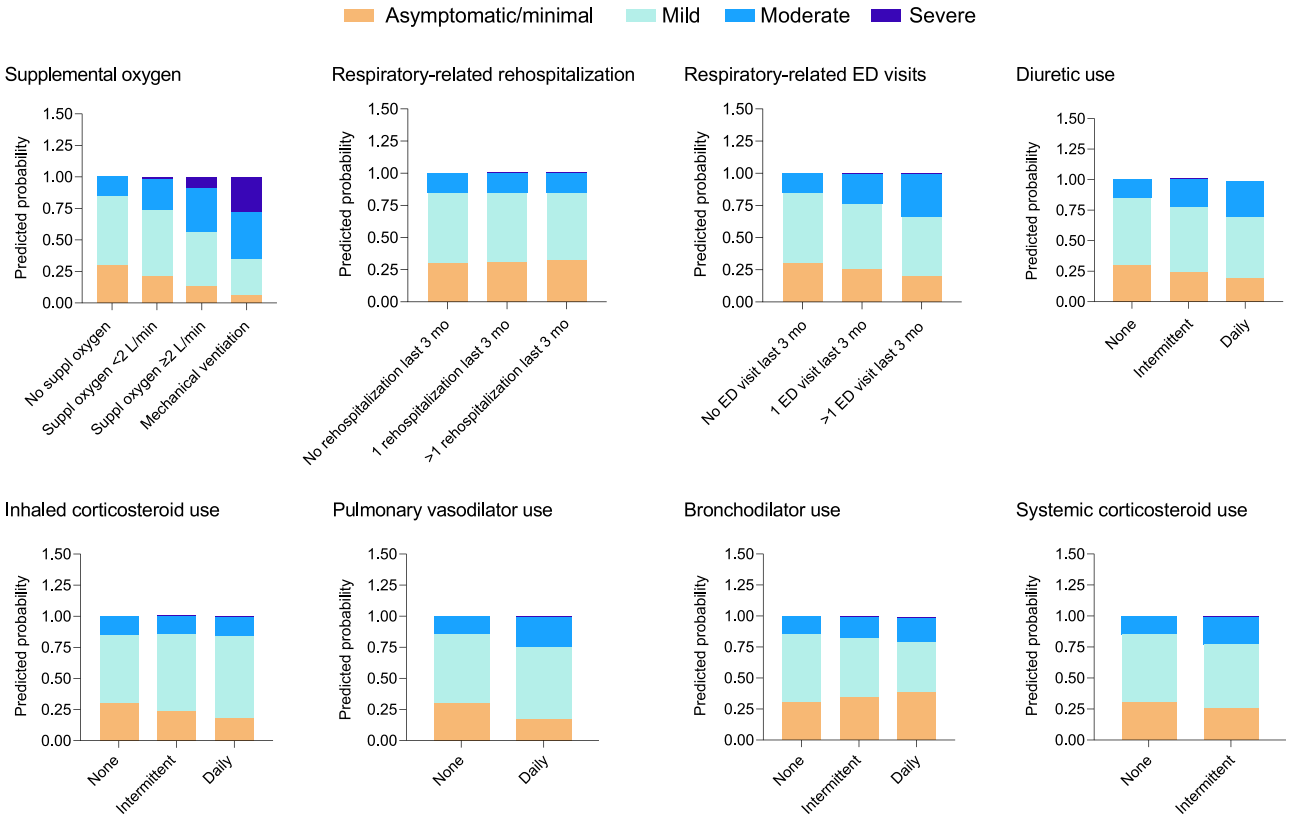

FIGURE 2 Predicted probability of choosing asymptomatic/minimal versus mild, moderate, or severe lung disease for each attribute level, holding all other attributes at their lowest level. BiPAP, bilevel positive airway pressure; CPAP, continuous positive airway pressure; ED, emergency department; NIPPV, nasal intermittent positive pressure ventilation; suppl, supplemental [Color figure can be viewed at wileyonlinelibrary.com]

diuretics, 9.3; intermittent use of systemic corticosteroids, 9.2; daily use of inhaled corticosteroids, 8.8; and daily use of a bronchodilator, 8.4.

4 | DISCUSSION

We solicited the opinions of a diverse group of physicians to explore their real-world impressions as to what defines different severities of CLDP at 12 months CA, based on a set of prespecified respiratory-related factors. Findings from this Delphi consensus-building study reveal that the most important factors for clinicians in assessing CLDP severity from NICU discharge through 12 months CA included home mechanical ventilation, supplemental oxygen at ≥2 L/min, respiratory-related rehospitalization, and supplemental oxygen at <2 L/min. However, clinicians did not characterize as unimportant any of the 13 prespecified factors discussed in the Delphi survey.

The current study built on work previously conducted by Gage et al.[15] and includes additional questions on systemic corticosteroid use and pulmonary vasodilator therapy. Further, while Gage et al. confined their Delphi-based survey to pediatric pulmonologists at North American academic health science centers, we sought greater generalizability in the current study by including pediatricians and neonatologists, in addition to pediatric pulmonologists, and by surveying a global sample of physicians from 11 countries across North America, Europe, Asia, and South America. This approach acknowledged the wide variations in neonatology clinical practice. For example, in countries other than the United States, some infants might be treated by a pediatrician, as opposed to a subspecialist (i.e., pediatric pulmonologist).

The results of this study informed the weighting and scoring of factors in the CLDPSS, a novel instrument to measure the severity of CLDP after discharge from the NICU through 12 months CA. The CLDPSS is a continuous outcome measure that is more discriminating and informative than any dichotomous measure, and it is easily measured with clinical data that are typically available until 1 year CA. This is important, given that the Prematurity and Respiratory Outcomes Program study found that many infants with severe BPD at 36 weeks PMA reported no respiratory morbidity at 1 year CA, while other infants who did not have a BPD diagnosis at 36 weeks PMA did have morbidity at 1 year CA.[12]

CLDPSS data can be collected directly from caregivers for assessment of CLDP in clinical trials of premature infants with

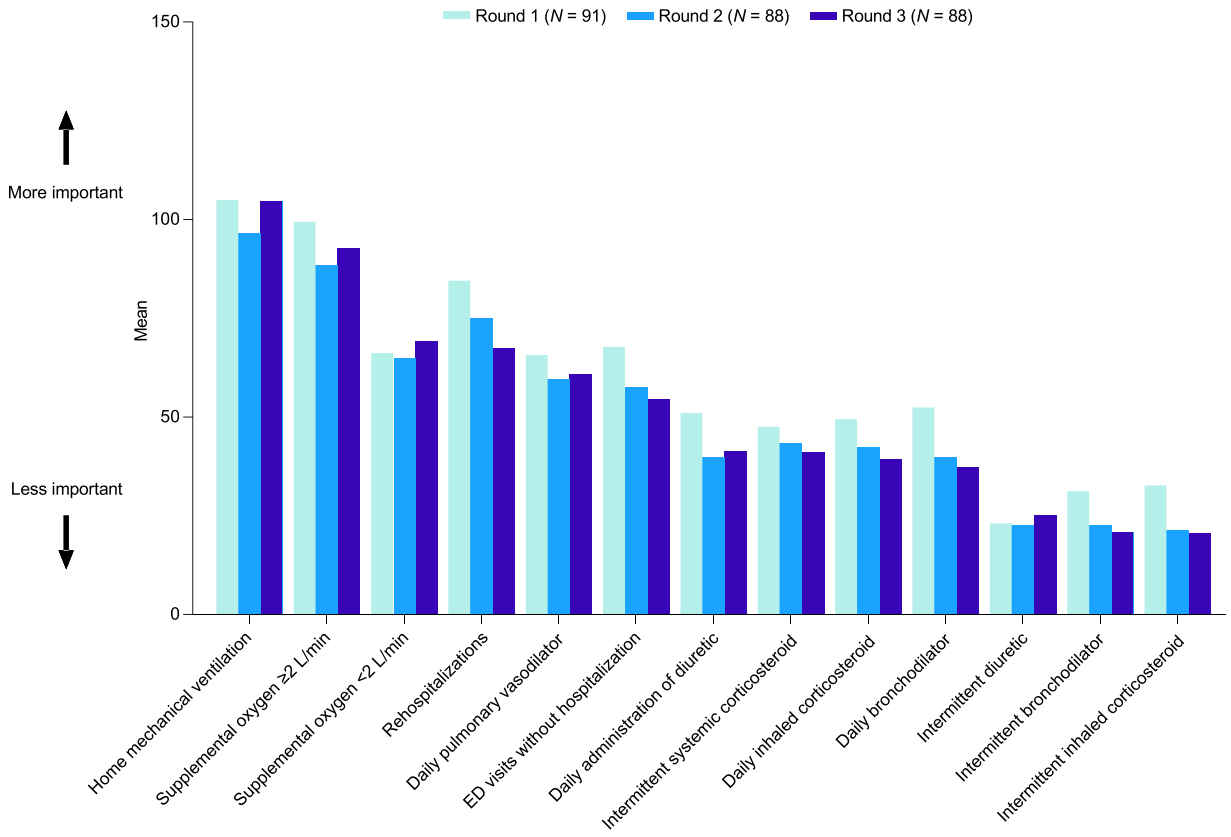

FIGURE 3 Mean overall importance of factors in evaluating CLDP severity. Calculated as the product of absolute importance rating and adjusted relative importance rank (i.e., ranking values adjusted such that "1 = least important"); therefore, higher values indicate greater relative importance; presented in order of the Round 3 results (most to least important). CLDP, chronic lung disease of prematurity; ED, emergency department [Color figure can be viewed at wileyonlinelibrary.com]

respiratory issues. Further, the scale is brief, resulting in less respondent (caregiver) or clinician burden, and can aid in routine monitoring of preterm infants with respiratory issues. The CLDPSS will provide standardized assessments of long-term pulmonary outcomes that may be useful in interventional neonatal studies. Data may be comparable across real-world clinical practices and may aid in benchmarking of outcomes in this population. If these data can determine factors during the NICU stay that correlate with long-term pulmonary morbidity better than a diagnosis of BPD at 36 or 40 weeks PMA, the CLDPSS will be an important addition to long-term care.

The CLDPSS development phase (clinician feedback) is complete. The reliability and validity of the CLDPSS scale will be evaluated via a prospective study assessing clinimetric properties, using patient data in routine clinical care settings.

The Delphi approach allowed the differentiation of the importance of specific factors among a diverse group of clinicians in multiple countries. The finding that the daily use of pulmonary vasodilators was only moderately important (i.e., not among the four most important factors) to the surveyed physicians was surprising. Previous studies have shown that when BPD-associated pulmonary hypertension is severe enough to require daily pulmonary vasodilators, it is associated with severe BPD with increased morbidity and mortality through 1 year CA.[18,19] Further support for this association comes from a 2020 retrospective study of very preterm infants with severe BPD, where investigators found that a diagnosis of pulmonary hypertension was a primary predictor of mortality.[20] The ranking of vasodilators outside the four most important factors here could be due, at least in part, to the three granular levels of supplemental oxygen use (mechanical ventilation, ≥2 L/min, and <2 L/min). If these factors were grouped as one supplemental oxygen variable, pulmonary vasodilator use would be the third most important variable (behind supplemental oxygen and rehospitalizations). However, we believe that grouping the three levels of supplemental oxygen had a minor effect only on the ranking of pulmonary vasodilators, and the finding is more likely explained by the diversity of surveyed clinicians in the current study.

We do not have an understanding of the factors that drove the clinicians' importance ratings/rankings in the current study, but it is possible that the frequency with which treatment is used in clinical practice may have had an impact. This could explain why the final weights for the pulmonary vasodilator variable were not as high as expected when compared with the other variables. It is possible that clinical practice changes resulting

from the emerging literature on the management of infants with BPD/pulmonary hypertension is variable across centers and geographical locations. Additional differences among respondents that could account for our findings include years of experience, the degree of patient severity they usually manage, and their formal training or clinical specialty (e.g., pulmonology, pediatrics, neonatology).

Our study has some limitations. We provided clinical scenarios at one point in time (i.e., as in a cross-sectional study), so we cannot be certain how our approach will perform in the same child over time. We did not provide physiologic measurements of respiratory (e.g., infant pulmonary function testing, hypoxic oxygen challenge, lung clearance index) or cardiovascular (e.g., pulmonary artery pressures) function, so we do not know how sensitive the predictors will be to changes in the patient's physiologic status over time.

Furthermore, clinicians were recruited via online research panels, which could potentially introduce sample bias. Clinicians had different specializations, were from different countries, and were not necessarily academically active in this area of research; therefore, they might have assigned different importance to factors due to differences in clinical practice—for example, approaches to ventilatory support, whether it be supplemental oxygen, positive airway pressure, and so forth. Lastly, we speculate that pediatric pulmonologists see more severe cases than pediatricians do, while neonatologists are typically only involved in early NICU care and not when the child is 1 year CA.

In conclusion, we identified factors important for clinicians in assessing CLDP severity after NICU discharge through 12 months CA following extremely preterm birth to develop a weighted CLDPSS. The CLDPSS adds granularity to previous instruments in terms of supplemental oxygen; respiratory-related ED visits; and administration of vasodilators, systemic corticosteroids, and intermittent diuretics. The CLDPSS will enable the standardized assessment of long-term pulmonary morbidity in neonatal trials and clinical practice.

ACKNOWLEDGMENTS

Under direction of the authors, Rosalind Bonomally, MSc, employee of Excel Medical Affairs, provided writing assistance for this manuscript. Editorial assistance in formatting, proof reading, and copy editing was also provided by Excel Medical Affairs. Shire, a Takeda company, provided funding to Excel Medical Affairs for support in editing this manuscript. The interpretation of the data was made by the authors independently. This study was funded by Takeda.

CONFLICT OF INTERESTS

Hugh M. O'Brodovich, Robin Steinhorn, Robert M. Ward, and Mikko Hallman were paid consultants to the Shire, a Takeda company, in connection with this study. Ethan J. Schwartz and Magdalena Vanya are employees of ICON and performed contracted research for the Shire, a Takeda company, in connection with this study. Ellen M. Janssen was an employee of ICON at the time the study was carried out. Linda Han is an employee of and owns stock/stock options in Takeda. Alexandra Mangili and Sujata P. Sarda were employees of Takeda at the time the study was carried out.

ORCID

Sujata P. Sarda https://orcid.org/0000-0003-3729-9539

REFERENCES

1. Siffel C, Kistler KD, Lewis JFM, Sarda SP. Global incidence of bronchopulmonary dysplasia among extremely preterm infants: a systematic literature review. *J Matern Fetal Neonatal Med*. 2019: 1-11.
2. Lapcharoensap W, Gage SC, Kan P, et al. Hospital variation and risk factors for bronchopulmonary dysplasia in a population-based cohort. *JAMA Pediatr*. 2015;169(2):e143676.
3. Stoll BJ, Hansen NI, Bell EF, et al. Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. *Pediatrics*. 2010;126(3):443-456.
4. Mowitz ME, Ayyagari R, Gao W, Zhao J, Mangili A, Sarda SP. Health care burden of bronchopulmonary dysplasia among extremely preterm infants. *Front Pediatr*. 2019;7:510.
5. Smith VC, Zupancic JAF, McCormick MC, et al. Rehospitalization in the first year of life among infants with bronchopulmonary dysplasia. *J Pediatr*. 2004;144(6):799-803.
6. Baraldi E, Filippone M. Chronic lung disease after premature birth. *N Engl J Med*. 2007;357(19):1946-1955.
7. Mandell EW, Kratimenos P, Abman SH, Steinhorn RH. Drugs for the prevention and treatment of bronchopulmonary dysplasia. *Clin Perinatol*. 2019;46(2):291-310.
8. Ley D, Hallberg B, Hansen-Pupp I, et al. rhIGF-1/rhIGFBP-3 in preterm infants: a phase 2 randomized controlled trial. *J Pediatr*. 2019;206:56-65.
9. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2001;163(7):1723-1729.
10. 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.
11. Higgins RD, Jobe AH, Koso-Thomas M, et al. Bronchopulmonary dysplasia: executive summary of a workshop. *J Pediatr*. 2018; 197:300-308.
12. Keller RL, Feng R, DeMauro SB, et al. Bronchopulmonary dysplasia and perinatal characteristics predict 1-year respiratory outcomes in newborns born at extremely low gestational age: a prospective cohort study. *J Pediatr*. 2017;187:89-97.
13. US Department of Health and Human Services. *Patient-Focused Drug Development: Collecting Comprehensive and Representative Input*. 2020; https://www.fda.gov/media/139088/download. Accessed September 24, 2020.
14. Corwin BK, Trembath AN, Hibbs AM. Bronchopulmonary dysplasia appropriateness as a surrogate marker for long-term pulmonary outcomes: a systematic review. *J Neonatal Perinatal Med*. 2018; 11(2):121-130.
15. Gage S, Kan P, Oehlert J, et al. Determinants of chronic lung disease severity in the first year of life: a population based study. *Pediatr Pulmonol*. 2015;50(9):878-888.
16. von der Gracht HA. Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technol Forecast Soc Change*. 2012;79(8):1525-1536.
17. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR

Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health*. 2013;16(1):3-13.
18. Lagatta JM, Hysinger EB, Zaniletti I, et al. The impact of pulmonary hypertension in preterm infants with severe bronchopulmonary dysplasia through 1 year. *J Pediatr*. 2018;203:218-224.
19. Krishnan U, Feinstein JA, Adatia I, et al. Evaluation and management of pulmonary hypertension in children with bronchopulmonary dysplasia. *J Pediatr*. 2017;188:24-34.
20. Wu KY, Jensen EA, White AM, et al. Characterization of disease phenotype in very preterm infants with severe bronchopulmonary dysplasia. *Am J Respir Crit Care Med*. 2020;201(11):1398-1406.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: O'Brodovich HM, Steinhorn R, Ward RM, et al. Development of a severity scale to assess chronic lung disease after extremely preterm birth. *Pediatric Pulmonology*. 2021;56:1583–1592.
https://doi.org/10.1002/ppul.25279