Duke | MARGOLIS CENTER for Health Policy

Advancing Endpoint Development for Preterm Neonates with Pulmonary Morbidities

JW Marriott • Washington, DC
October 2, 2018

Discussion Guide

Preterm neonates are often at risk of ongoing pulmonary morbidity, particularly due to developmental immaturity of the lungs at birth as well as damage caused by mechanical ventilation and long-term use of supplemental oxygen. One measure of pulmonary morbidity is captured by a diagnosis of bronchopulmonary dysplasia (BPD), which has been typically defined as the need for oxygen, with or without positive pressure respiratory support, at 36 weeks post-menstrual age (PMA) or "corrected" age. In clinical trials designed to improve pulmonary and respiratory outcomes, the most commonly used endpoint has been BPD. However, there is no consensus on whether BPD is a clinically meaningful endpoint to predict pulmonary and respiratory outcomes at infancy, childhood, and beyond.

Efforts are underway to support the development of validated efficacy endpoints, including Clinical Outcome Assessments (COAs), that incorporate longer-term improvement or changes in pulmonary outcomes and are data-driven. However, there are outstanding questions about how best to support the development of such endpoints.

To support further progress in this area, the Duke-Margolis Center for Health Policy is convening this expert workshop under a cooperative agreement with the U.S. Food and Drug Administration (FDA) to explore the work that has been done to date on efficacy endpoint development for preterm neonates with pulmonary morbidities. Workshop discussion will specifically highlight gaps and limitations in existing data and research in order to identify a path forward for developing validated endpoints and COAs that better represent short- and long-term outcomes associated with pulmonary morbidities of prematurity.

Defining Bronchopulmonary Dysplasia (BPD) as a Disease and Endpoint

BPD is one of the most common diseases of preterm birth, affecting 10,000-15,000 infants every year in the United States.[1] BPD refers to the need for supplemental oxygen, with or without additional respiratory support, due to complications from prematurity. Unlike other diseases related to severe prematurity, the rate of BPD in survivors of preterm birth has not declined over time. This may be due in part to advances in neonatal care that have increased survival for the smallest infants.[2] However, there have been few therapeutic advances in treatment over the past 25 years.[3] Many patients affected by BPD have recurrent hospitalizations and ongoing pulmonary morbidities, impacting their quality of life.

One of the challenges with using a diagnosis of BPD to correlate to short- and long-term outcomes associated with pulmonary morbidities of preterm birth is the lack of a universally accepted definition of the disease. The definition of BPD has evolved over time and there have been recent efforts to refine this definition further.[4] The most commonly used definition for BPD is the need for oxygen, with or without positive pressure respiratory support, at 36 weeks PMA or "corrected" age.[5] Another frequently

used categorization ranks neonates by severity for BPD including: none, mild, moderate, or severe.[6] These categorizations are based primarily on the level of supplemental oxygen or respiratory support that patients require at 36 weeks. However, neither of these definitions captures all infants who have pulmonary morbidities and neither reflects long-term pulmonary outcomes.

Another challenge with using a diagnosis of BPD to relate to long-term outcomes is the difficulty in describing the pathogenesis of the disease. There is a great deal of heterogeneity in patients who develop and have pulmonary morbidities after preterm birth. This can be due in part to different antenatal factors that increase the risk of developing these morbidities such as intrauterine growth restriction (IUGR), maternal hypertension, and maternal smoking.[7,8] Postnatal events that increase the risk of chronic pulmonary insufficiency include parenchymal lung disease, early pulmonary hypertension, and severe respiratory failure.[9] These various risk factors can lead to diverse experiences for these patients, particularly over the course of their childhoods. There are a range of interventions and techniques for managing the disease, many of which have not been systematically evaluated, resulting in variable outcomes.[10]

Because there are various definitions used to diagnose BPD, there has not been a consistent application of BPD as an endpoint in clinical trials. Moreover, there is no consensus whether the 36 week definition of BPD is an optimal endpoint in clinical trials, particularly to predict long-term outcomes in childhood. For example, some studies have failed to show significant differences in long-term pulmonary outcomes for infants with and without BPD at 36 weeks PMA, indicating that this endpoint may not be adequate for approval of therapeutic products.[11]

To mitigate some of the issues that arise using the 36 week definition, a 40 week PMA definition of BPD has been suggested as an alternative potential endpoint. This would remove certain characteristics that impact the predictive ability of BPD at 36 weeks PMA, such as immature respiratory control and feeding difficulties that are often resolved by 40 weeks PMA. Therefore, a 40 week PMA definition may be a better predictor of which patients will have long-term pulmonary issues; however, use of this time point requires further evaluation.[12]

Efficacy Endpoints in Medical Product Development

In order for a drug or biologic to be approved for marketing, it must be deemed safe and effective for its intended use – in other words, its potential benefits must outweigh its potential risks. Clinical trials are conducted to show that a treatment delivers this positive balance of benefits and risks. In clinical trials designed to measure the efficacy of a new medical product, endpoints are used to reflect the intended effects of the drug and demonstrate whether it provides clinical benefit. The specific endpoint chosen for a given study depends on the clinical trial design, the nature of the condition under study, and the expected effect of the treatment. Investigators typically use clinical outcomes, surrogate endpoints, biomarkers, and Clinical Outcome Assessments (COAs) to measure treatment benefit.

Clinical Outcomes

Clinical outcomes directly measure or reflect how an individual feels, functions, or survives.[13] These clinical outcomes measure whether there is an improvement of symptoms that is likely to outweigh any adverse events associated with a treatment. For example, in clinical trials for infants and children with

pulmonary morbidities, endpoints such as feeding capability, growth, development, and freedom from requiring medical/respiratory support or technology may measure how these patients are feeling and functioning. Functioning and feeling endpoints, however, may be more difficult to measure in neonates. Therefore, in neonatal trials, survival might be a reasonable endpoint for measuring improvement of early, severe symptoms.

Surrogate Endpoints

Surrogate endpoints may be used instead of clinical outcomes in some clinical trials. Surrogate endpoints are likely to *predict* clinical benefit in contrast to clinical outcomes which *reflect* clinical benefit. Surrogate endpoints may be used as a substitute in clinical trials when the use of a clinical outcome may not be practical, possible, or ethical. For example, surrogate endpoints may be useful when outcomes take a very long time to study or when the clinical benefit of improving the surrogate endpoint is well understood.[14]

Additional clinical trials must be designed to ultimately show that surrogate endpoints themselves are able to predict a specific clinical benefit. Surrogate endpoints that have undergone this process are called validated surrogate endpoints and may support medical product approval without the need for additional studies. For example, LDL cholesterol reduction is a validated surrogate endpoint for a reduction of CV events and has been used as the basis for approval of statins.[15] Validated surrogate endpoints generally enable clinical studies to be conducted in smaller numbers of patients over shorter periods of time, increasing the efficiency of the drug development process.

BPD, when defined as the need for oxygen, with or without positive pressure respiratory support, at 36 weeks PMA or “corrected” age, is a surrogate endpoint. However, as mentioned previously, this endpoint warrants further examination, as it is not a validated surrogate endpoint and may not be an optimal efficacy endpoint for regulatory submissions.

Biomarkers

A biomarker is a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.[16] Biomarkers can help diagnose a disease or predict future disease severity or outcomes.[17]

There has been difficulty identifying biomarkers with high predictive accuracy for either the risk of developing BPD or the risk of suffering certain outcomes for those diagnosed with BPD. Potential prognostic biomarkers that might identify this risk could use clinical variables (e.g. birth weight, gestational age), measurements of analytes in the fluid (e.g. blood, urine) and lung function measures, such as pulmonary function testing.[18]

Some have also suggested using enrichment biomarkers to identify subgroups of patients who have BPD. For example, subgroups could be identified by using specific risk or genetic factors. Once these subgroups are defined and validated, they then could be used to predict specific clinically important outcomes.[19] However, further research is needed to develop biomarkers for this particular use.

Clinical Outcome Assessments

Clinical outcome assessments (COAs) can also be used to evaluate the efficacy of a treatment. COAs measure a patient's symptoms, overall mental state, or the effects of a disease or condition on how the patient functions, and can be used to determine clinical benefit.[20] Most clinical studies submitted to FDA to support registration of a drug utilize at least one COA, either as part of the primary or secondary endpoint definition to evaluate the effect of a treatment, or as an exploratory endpoint that informs future research.[21]

COAs are generally divided into four broad categories, depending on how the assessment is conducted and reported: patient-reported outcomes (PROs), clinician-reported outcomes (ClinROs), observer-reported outcomes (ObsROs), and performance outcomes (PerfOs). A description of each of the four types is provided in **Table 1** below.

Table 1: Clinical Outcome Assessment Definitions[22]

COA Type	Definition
Patient-Reported Outcomes (PROs)	A measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else.
Clinician-Reported Outcomes (ClinROs)	A measurement based on a report that comes from a trained healthcare professional after observation of a patient's health condition. Most ClinRO measures involve a clinical judgment or interpretation of the observable signs, behaviors, or other manifestations related to a disease or condition. ClinRO measures cannot directly assess symptoms that are known only to the patient.
Observer-Reported Outcomes (ObsROs)	A measurement based on a report of observable signs, events, or behaviors related to a patient's health condition by someone other than the patient or a healthcare professional. Generally, ObsROs are reported by a parent, caregiver, or someone who observes the patient in daily life and are particularly useful for patients who cannot report for themselves (e.g., infants or individuals who are cognitively impaired). An ObsRO measure does not include medical judgment or interpretation.
Performance Outcomes (PerfOs)	A measurement based on a task(s) performed by a patient according to instructions that are administered by a health care professional. Performance outcomes require patient cooperation and motivation.

The selection of a COA for use in a trial begins with an understanding of the outcome(s) that is (are) most relevant and meaningful to patients with the specific condition. The type of COA selected will depend on the context of use and the concept of interest that will be measured, and more than one COA may be appropriate.

The degree to which any COA instrument is suitable and valid for a particular context of use depends in large part on its measurement properties. These properties include:

- *Validity* – Validity refers to both an instrument's *content validity* and *construct validity*.
    - Content validity denotes the evidence showing that a COA measures its concept of interest, is comprehensive in what it measures, and is the appropriate type of COA for the context in which it is used.
    - Construct validity signifies that relationships among items and concepts should be logically related. The strength of this relationship is established by assessing its *convergent* and *discriminant* validity as well as through *factor analysis.*
        - Convergent validity is intended to show that similar indicators designed to measure overlapping constructs are interrelated.[23]
        - Discriminant validity is indicated by showing that two items that are intended to be unrelated, are in fact unrelated.
        - Factor analysis is used to verify the latent dimensions of an instrument and show any underlying patterns.[24] The results of a factor analysis can provide confirmation of convergent and discriminant validity.
- *Reliability* – Reliability represents the ability of a COA instrument used in a clinical trial to yield consistent, reproducible estimates of treatment effect.[25] In ensuring that a COA instrument is reliable, there needs to be alignment across multiple reviewers administering the test, when applicable.
- *Ability to detect change* – The ability to detect change implies that an instrument can identify differences in scores over time to demonstrate treatment effect.

The usefulness and relevance of a COA endpoint depends on its ability to accurately capture the impact of an intervention and communicate those results to patients and other stakeholders. However, one of the key challenges in developing COA instruments and implementing them within clinical trials is determining what level or degree of change measured by the COA is considered clinically meaningful, rather than simply statistically significant. Showing that a change on a COA measurement is clinically meaningful and providing a treatment benefit is an important consideration in the endpoint development process.

The roadmap in **Appendix 1** illustrates the process of developing an outcome measurement for use in clinical trials.

Regulatory and Scientific Efforts to Support Development of New Endpoints for Preterm Neonates with Pulmonary Morbidities

Over the last several years, clinician-scientists, regulators, advocates, and industry partners have taken steps to facilitate the development of new efficacy endpoints for clinical trials in preterm neonates. In October 2014, FDA, in collaboration with the Critical Path Institute (C-Path) and the Burroughs Welcome Fund (BWF), held the first annual neonatal scientific workshop to accelerate the development of treatments for the neonatal population as well as to discuss the formation of a new international consortium dedicated to applying regulatory science to this population.[26]

Following the success of this workshop, the International Neonatal Consortium (INC) launched as part of C-Path in May 2015.[27] This consortium brings together stakeholders from 24 member countries with the goal of developing a predictable regulatory path for evaluating the safety and effectiveness of therapies for neonates. The INC has also formed working groups concentrating on certain disease areas, one of which is BPD. The BPD group has been particularly focused on developing new efficacy endpoints that better reflect long-term pulmonary outcomes.

These efforts have been complemented by a recent meeting sponsored by the National Institute of Child Health and Human Development (NIH) in October 2016.[28] The purpose of that workshop was to discuss the definitional and diagnosis challenges associated with BPD, review the current evidence for prevention, early intervention and treatment, and identify gaps in knowledge in current research. These efforts and those of the INC BPD working group will serve as the foundation for much of the day's discussion.[29]

Meeting Objectives

The purpose of this workshop is to advance the discussion on developing validated efficacy endpoints for preterm neonates with pulmonary morbidities and to identify an actionable research agenda that will move this field forward. Discussion will encompass what is clinically meaningful in endpoint development for preterm neonates with pulmonary morbidities, how to develop these endpoints - including key measurement concepts – and how to translate the existing data sources available to advance endpoint development.

Session I: Current State of Research and Challenges in Developing Endpoints for Preterm Neonates with Pulmonary Morbidities

Objective: Provide an overview of how today's workshop fits into broader efforts to advance the science on endpoints for preterm neonates with pulmonary morbidities, as well as the key issues and current limitations in this space

Session II: Identifying What is Clinically Meaningful to Stakeholders in Endpoint Development for Preterm Neonates with Pulmonary Morbidities

Objective: Explore what is clinically meaningful to stakeholders in endpoint development

Questions to address:

- What does the disease process look like and what are we trying to prevent/improve?
- What is important to you, your child, and your family?

Session III: Defining the Potential for Endpoint Development for Preterm Neonates with Pulmonary Morbidities

Objective: Outline the components that should be measured as well as the domains that may be included in potential clinically meaningful alternatives such as endpoints and COAs, and strategies to advance the development of these tools

Questions to address:

- What should the timing of endpoints be? 1 year, 2 years, or some other time frame?
- What type of endpoint(s) would best serve the needs of all the stakeholders? (These include COA, clinical endpoint, biomarker, among others)

Session IV: Exploring Endpoint and COA Development

Objective: Review how these instruments are developed, including a discussion of key concepts that should be considered such as validity and reliability, as well as other important attributes

Questions to address:

- What are the most important components in endpoint development and how are these incorporated when developing new endpoints?
- Do you foresee particular measurement areas that could hinder endpoint development in this space?
- Are there exemplar COAs that could be helpful as we begin to develop COAs for neonates with pulmonary insufficiency?
- What is the feasibility of follow-up programs and how do we address family, researcher, and sponsor concerns?

Session V: Characterizing Data Sources for Endpoint and COA Development Opportunities

Objective: Examine the current availability in data and how to translate that data to support endpoint and COA development for preterm neonates with pulmonary morbidities

Questions to address:

- What is possible with existing sources?
- What data are needed and how might they be obtainable?
- Is there an instrument/endpoint ready for testing? If not, what would be needed to make it ready for testing?

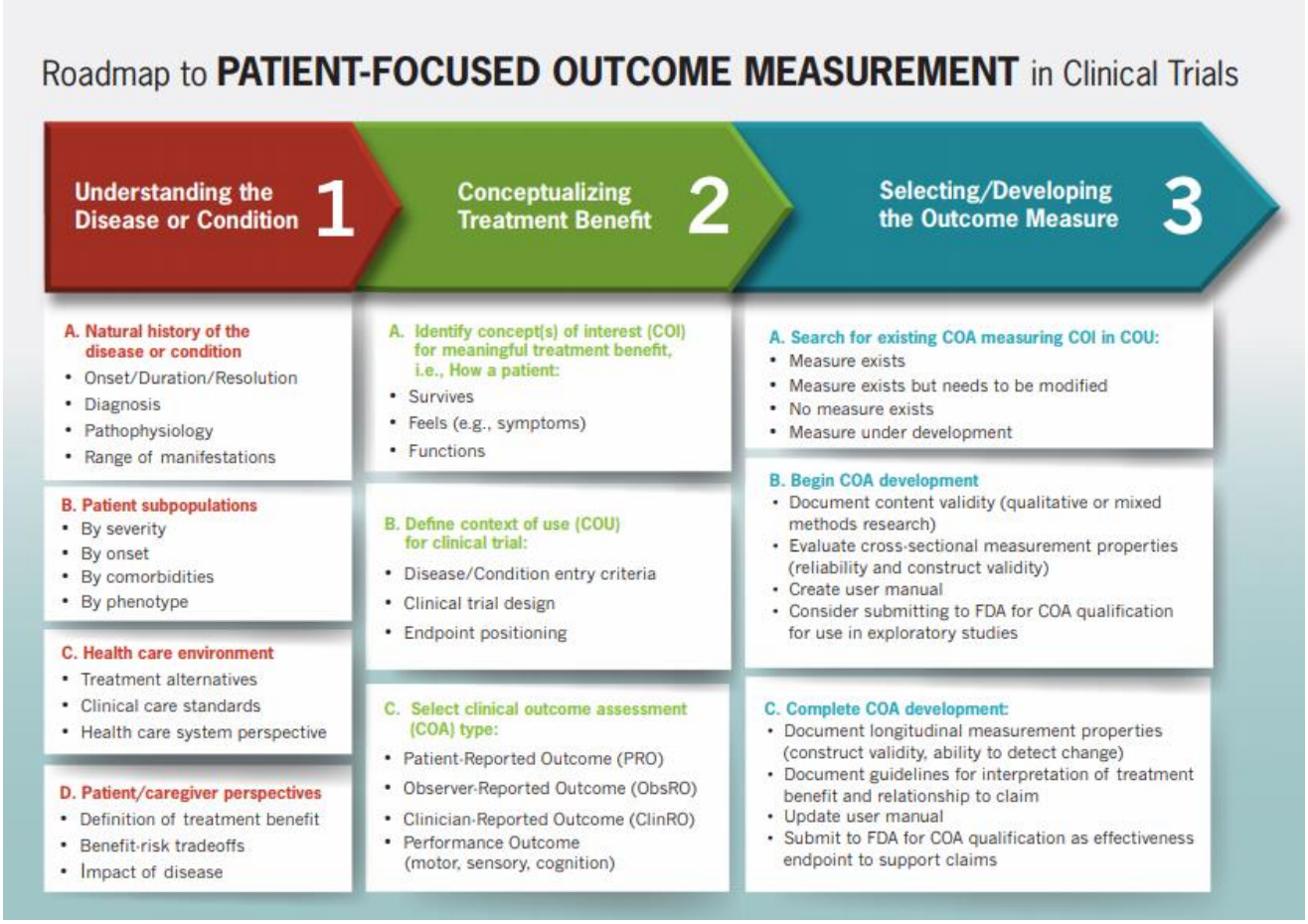
Session VI: Synthesis Discussion and Potential Paths Forward

Objective: Reflect on the day's discussion, identify the areas of consensus and potential paths forward, and highlight major gaps in knowledge and prioritize them for future research and discussion

Questions to address:

- What are actionable steps we can take or recommend to make real progress in the development of these endpoints?

Appendix 1: Roadmap to Patient-Focused Outcome Measurement in Clinical Trials



[1] Jensen E.A., & Schmidt B. (2014). Epidemiology of bronchopulmonary dysplasia. *Birth Defects Res A Clin Mol Teratol, 100*(3):145-57.
[2] Higgins, R.D., Jobe, A.H., Koso-Thomas, M., Bancalari, E., Viscardi, R.M., Hartert, T.V.,…Raju, T.N.K. (2018). Bronchopulmonary Dysplasia: Executive Summary of a Workshop. *The Journal of Pediatrics, 197*, 300-308.
[3] Iyengar A., & Davis J.M. (2015). Drug therapy for the prevention and treatment of bronchopulmonary dysplasia. *Front Pharmacol*, 6:12.
[4] Higgins, R.D., Jobe, A.H, Koso-Thomas, M., Bancalari, E., Viscardi, R.M., Hartert, T.V.,…Raju, T.N.K. (2018). Bronchopulmonary Dysplasia: Executive Summary of a Workshop. *The Journal of Pediatrics, 197*, 300-308.
[5] Higgins, R.D., Jobe, A.H, Koso-Thomas, M., Bancalari, E., Viscardi, R.M., Hartert, T.V.,…Raju, T.N.K. (2018). Bronchopulmonary Dysplasia: Executive Summary of a Workshop. *The Journal of Pediatrics, 197*, 300-308.
[6] Walsh M.C., Yao Q., Gettner P., Hale E., Collins M., Hensman A.,…McDavid, G. (2004). Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics*, *114*, 1305-11.
[7] Keller R.L., Feng R., DeMauro S.B., Ferkol T., Hardie W., Rogers E.E.,…Merrill J.D. (2017). Bronchopulmonary dysplasia and perinatal characteristics predict 1-year respiratory outcomes in newborns born at extremely low gestational age: a prospective cohort study. J Pediatr, *187*, 89-97
[8] Manuck T.A., Levy P.T., Gyamfi-Bannerman C., Jobe A.H., & Blaisdell C.J. (2016). Prenatal and perinatal determinants of lung health and disease in early life: a National Heart, Lung, and Blood Institute Workshop report. *JAMA Pediatr 170*(5).
[9] Steinhorn, R., Davis, J.M., Göpel, W., Jobe A.H., Abman, S., Laughon, M.,…Turner, M.A. (2017). Chronic Pulmonary Insufficiency of Prematurity: Developing Optimal Endpoints for Drug Development. *The Journal of Pediatrics*, *191*, 15-21.
[10] Higgins, R.D., Jobe, A.H., Koso-Thomas, M., Bancalari, E., Viscardi, R.M., Hartert, T.V.,…Raju, T.N.K. (2018). Bronchopulmonary Dysplasia: Executive Summary of a Workshop. *The Journal of Pediatrics, 197*, 300-308.
[11] Steinhorn, R., Davis, J.M., Göpel, W., Jobe A.H., Abman, S., Laughon, M.,…Turner, M.A. (2017). Chronic Pulmonary Insufficiency of Prematurity: Developing Optimal Endpoints for Drug Development. *The Journal of Pediatrics*, *191*, 15-21.
[12] Steinhorn, R., Davis, J.M., Göpel, W., Jobe A.H., Abman, S., Laughon, M.,…Turner, M.A. (2017). Chronic Pulmonary Insufficiency of Prematurity: Developing Optimal Endpoints for Drug Development. *The Journal of Pediatrics*, *191*, 15-21
[13] FDA-NIH Biomarker Working Group. (2016). Glossary. *BEST (Biomarkers, EndpointS, and other Tools) Resource.* Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK338448/
[14] https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/ucm606684.htm
[15] FDA-NIH Biomarker Working Group. (2016). Glossary. *BEST (Biomarkers, EndpointS, and other Tools) Resource.* Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK338448/
[16] FDA-NIH Biomarker Working Group. (2016). Glossary. *BEST (Biomarkers, EndpointS, and other Tools) Resource.* Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK338448/
[17] U.S. Food and Drug Administration. FDA Facts: Biomarkers and Surrogate Endpoints. Available at: https://www.fda.gov/aboutfda/innovation/ucm512503.htm
[18] Charitharth, V.L. & Namasivayam, A. (2015). Early Diagnosis, and Clinical Predictors of Bronchopulmonary Dysplasia. *Clinics in Perinatology, 42(4),* 739-754.
[19] Steinhorn, R., Davis, J.M., Göpel, W., Jobe A.H., Abman, S., Laughon, M.,…Turner, M.A. (2017). Chronic Pulmonary Insufficiency of Prematurity: Developing Optimal Endpoints for Drug Development. *The Journal of Pediatrics*, *191*, 15-21.
[20] U.S. Food and Drug Administration. Clinical Outcome Assessment Qualification Program. Available at: https://www.fda.gov/drugs/developmentapprovalprocess/drugdevelopmenttoolsqualificationprogram/ucm284077.htm
[21] Walton, MK, Powers, JH, Hobart, J, et. al. Clinical Outcome Assessments: Conceptual Foundation—Report of the ISPOR Clinical Outcomes Assessment – Emerging Good Practices for Outcomes Research Task Force. Value in Health. 2015; 18:741-752. doi.org/10.1016/j.jval.2015.08.006
[22] FDA-NIH Biomarker Working Group. (2016). Glossary. *BEST (Biomarkers, EndpointS, and other Tools) Resource.* Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK338448/

[23] Brown, T. (2006). Confirmatory factor analysis for applied research: Methodology in the social sciences. New York: The Guilford Press

[24] Brown, T. (2006). Confirmatory factor analysis for applied research: Methodology in the social sciences. New York: The Guilford Press

[25] US Food and Drug Administration. (2009). Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Retrieved from http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf

[26] U.S. Food and Administration. First Annual Neonatal Scientific Workshop – Roadmap for Applying Regulatory Science to Neonates. October 28-29, 2014. Available at: http://wayback.archive-it.org/7993/20161022055329/http://www.fda.gov/Drugs/NewsEvents/ucm410863.htm

[27] Critical Path Institute. Critical Path Institute Launches New Neonatal Consortium. May 19, 2015. https://c-path.org/critical-path-institute-launches-new-neonatal-consortium/

[28] Bronchopulmonary Dysplasia: Gaps in Knowledge and Opportunities for Research. October 17-18, 2016. Available at: https://www.nichd.nih.gov/about/meetings/2016/101716

[29] Steinhorn, R., Davis, J.M., Göpel, W., Jobe A.H., Abman, S., Laughon, M.,…Turner, M.A. (2017). Chronic Pulmonary Insufficiency of Prematurity: Developing Optimal Endpoints for Drug Development. *The Journal of Pediatrics*, *191*, 15-21