Published in final edited form as:

Acta Paediatr. 2018 March ; 107(3): 414–424. doi:10.1111/apa.14128.

Neonatal Eating Outcome Assessment: Tool Development and Inter-rater Reliability

Roberta Pineda, PhD, OTR/L^{1,2}, Rachel Harris, BA¹, Felicia Foci, MSOT¹, Jessica Roussin, BA¹, and Michael Wallendorf, PhD³

¹Program in Occupational Therapy, Washington University School of Medicine, St. Louis, MO

²Department of Pediatrics, Washington University School of Medicine, St. Louis, MO

³Division of Biostatistics, Washington University School of Medicine, St. Louis, MO

Abstract

Aim—To define the process of tool development and revision for the Neonatal Eating Outcome (NEO) Assessment and to report preliminary inter-rater reliability.

Methods—Tool development consisted of a review of the literature and observations of feeding performance among 178 preterm infants born < 32 weeks gestation. 11 neonatal therapy feeding experts provided structured feedback to establish content validity and define the scoring matrix. The tool was then used to evaluate feeding in 50 preterm infants born < 32 weeks gestation and 50 full term infants. Multiple revisions occurred at each stage of development. Finally, six neonatal occupational therapists participated in reliability testing by independently scoring five videos of oral feeding of preterm infants using version 4 of the tool.

Results—The intraclass correlation for the “pre-feeding” score was 0.71 (0.37-0.96), and the intraclass correlation for the “total” score was 0.83 (0.56-0.98).

Conclusion—The “total” score had good to excellent reliability. Fleiss' Kappa scores for all 18 scorable items ranged from slight agreement to moderate agreement. Items with lowest Kappa scores were revised, and additional feedback from therapists engaged in reliability testing was incorporated, resulting in final version 5.

Address correspondence to: Roberta Pineda, PhD, OTR/L, Washington University School of Medicine, Program in Occupational Therapy, 1 Children's Place, St. Louis MO, 63110, USA, [pineda_r@kids.wustl.edu], Phone: 314-286-1304, Fax: 314-286-1261.

Financial Disclosure: The NEO Assessment copyright is currently owned by the Office of Technology Management at Washington University. This tool can be accessed by the public through licensing, with potential profit to Washington University.

Conflict of Interest: Authors have indicated there are no conflicts of interest to report.

Contributor's Statements: Roberta Pineda: Dr. Pineda developed the Neonatal Eating Outcome assessment, assisted with conceptualization and design of the study, collected data, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Rachel Harris: Ms. Harris completed training material and table development, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Felicia Foci: Ms. Foci participated in study design and data collection, drafted the initial manuscript, revised the manuscript, and approved the final manuscript as submitted.

Jessica Roussin: Ms. Roussin participated in data collection, developed figures and tables, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Michael Wallendorf: Dr. Wallendorf chose appropriate statistical approaches and completed data analyses. He also approved the final manuscript as submitted.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Graphical abstract

Table of Contents: This manuscript outlines the process of development of the Neonatal Eating Outcome Assessment, a new developmental feeding assessment for high-risk infants. Inter-rater reliability is also reported.

Keywords

feeding; preterm; development; NICU; infant; outcome

Introduction

While survival rates for preterm infants have increased, the risk of morbidity remains high. Therapy services (physical therapy, occupational therapy, and speech-language pathology) in the neonatal intensive care unit (NICU) are on the rise, as more research identifies the developmental challenges that high-risk infants experience. Feeding is the primary occupation of infancy, and feeding problems can lead to negative feeding experiences, which can affect nutrition, developmental outcome, and parent interaction/social adaptation (1).

Feeding is a common focus of neonatal therapy, as feeding challenges affect one in five preterm infants (2). Evaluation and treatment of feeding relies on the availability and use of standardized, reliable, and valid feeding tools, which are important to aid in identification of feeding performance and to determine the efficacy of interventions. However, there are few tools designed to assess feeding in preterm infants in the NICU.

Some neonatal feeding assessments (the LATCH, Preterm Infant Breastfeeding Behavior Scale and the Systematic Assessment of the Infant at the Breast) were designed solely for breastfed infants. The Early Feeding Skills Assessment for Preterm Infants (3), the Support of Oral Feeding For Fragile Infants (SOFFI) (4), and the Infant Driven Feeding Assessment (5) were developed to guide the clinician as to when, along the pathway to term equivalent age, preterm infants should be introduced to oral feeding. Such assessments are often used as a larger model of care aimed at ensuring that infants drive when/if feeding occurs. An assessment to identify stages of nutritive sucking, which highlights factors related to compression, suction, respiration and rhythmicity that advance with postmenstrual age (PMA) has been developed (6). In addition, Lau, et al has also published a method to assess oral feeding skills (7) that relies on total volume of intake. However, many clinicians aim to better understand the functional components needed for successful oral feeding. One assessment, the Neonatal Oral Motor Assessment Scale (NOMAS), aims to discriminate dysfunctional feeding performance in neonates (8). In addition to reports of poor reliability (inter-rater reliability ranged from Kappa= -0.43-0.62) (9), the NOMAS is limited in that it is a static assessment of feeding, with no adapted scoring due to expected changes across PMA. None of the currently available feeding assessments aim to determine feeding performance, while considering the developmental changes that occur from early feeding attempts to those occurring around term equivalent age in preterm infants.

Oral feeding is one of the most complex tasks of infancy, and feeding is further complicated in preterm infants by attempts at oral feeding prior to term equivalent age. Full volume of intake can be contingent on a number of different factors, including arousal, reflex maturation, sensory and motor function, and coordination of suck-swallow-breathe synchrony (5, 10-13). Oral feeding prior to term equivalent age in the preterm infant is developmentally regulated (12, 14, 15). Reflexes necessary for feeding are typically coordinated and developed by 32-34 weeks PMA (16), but low levels of arousal can hinder the feeding process until after 36 weeks PMA (12). Most infants can orally feed by 40 weeks PMA, the time they would have been born if not born preterm. Cue-based feeding (4, 5), in which feeding attempts occur based on the infant's signs of readiness, has received recent attention and encourages feeding based on infant cues instead of on age or resolution of medical problems. However, early feeding prior to term equivalent age can result in lack of full volume of intake, inability to remain in an awake state, apnea, bradycardia, and/or oxygen desaturations (17-22). Such occurrences may not be abnormal, based on the PMA of the infant. An assessment that can differentiate normal from abnormal performance prior to term age is needed. A feeding tool that can accomplish this can potentially minimize unnecessary therapy services for feeding prior to developmental readiness, aid early identification to enable targeted interventions, and allow clinicians and researchers to better measure the efficacy of feeding interventions.

The Neonatal Eating Outcome (NEO) Assessment was designed to fill the need for such a tool. It was modeled after the scoring system of the Premie-Neuro, a neurobehavioral assessment that assigns scores for each item based on expected neurobehavioral performance across different PMAs (23). Such a model does not score infants based on the expected performance at term equivalent age, but instead acknowledges that appropriate performance may be different based on the progression of skills, which relies on maturity along the pathway from preterm birth to term equivalent age. By assigning scores based on the current PMA, an infant's emerging skills are not penalized if they have not reached the maturity expected of a full term infant.

This manuscript identifies the process of tool development, as well as identifies preliminary inter-rater reliability of the NEO Assessment.

Methods

Each part of this project that involved human subjects, which included infants as well as content experts, was approved by the Human Research Protection Office of Washington University School of Medicine in Saint Louis, and participants gave informed consent.

Overview of Procedures

This manuscript outlines the process of development of the NEO Assessment as well as the psychometric investigations that informed different stages of revision. First, the NEO Assessment was developed using research, clinical knowledge regarding the component skills needed for oral feeding, and observations of the feeding behaviors of 178 preterm infants at or near term equivalent age (version 1). Second, content validity was established by having 11 neonatal feeding experts review the NEO Assessment and provide structured

feedback that was then incorporated. Guided questions also enabled the content experts to define appropriate scoring of each item across PMA. Version 2 included all content expert feedback and the addition of the scoring matrix. Next, the NEO Assessment was used to assess 50 full term infants within the first few days of birth and 50 preterm infants across time. Revisions were made to improve clarity, to ensure appropriate options for a wide variety of feeding performance and to achieve agreement among the research team (versions 3 and 4). Finally, inter-rater reliability testing was done using six neonatal occupational therapists (using version 4). Version 5 was finalized after incorporating changes from the reliability testing, specifically targeting items with the lowest agreement and rater feedback. Version 5 is the final version that is available commercially.

The Neonatal Eating Outcome Assessment

The NEO Assessment was designed to be a comprehensive, standardized assessment of feeding for preterm infants in the NICU (24). The purpose of the NEO Assessment is to assess age-appropriate oral motor and feeding skills in preterm infants prior to and at term equivalent age. Scoring of the NEO Assessment is based on the infant's PMA at the time of assessment. The NEO can be used with breast or bottle fed infants.

The assessment is divided into four sections, I-Pre-Feeding Behaviors, II-Oral Feeding, III-Observations at the End of Feed, and IV-Non-Scored Items. Within these four sections the assessment evaluates feeding readiness, safety, and success with oral feeding in addition to complex feeding behaviors such as root and grasp, swallow, and suck-swallow-breathe coordination. The full assessment (Section I, II, III, and IV) is appropriate for infants who have initiated oral feeding (a minimum age of approximately 30-32 weeks PMA) through approximately 4-6 weeks corrected age (post-term). However, an abbreviated form (in which only Section I is scored) can be used for infants after 30 weeks PMA who are not yet orally feeding.

There are 19 items on the NEO Assessment that are scored. One of those items, Respiratory Support, is scored and used to define readiness to orally feed but is not included in the total score on the assessment. There are also 26 non-scored observations as part of Section IV that include potentially modifiable factors; these are defined so that scoring on the assessment can be done in the context of modifications, changes in feeding apparatus, or interventions done to impact feeding.

Refer to Table 1 for a list of the scored items on the NEO Assessment as well as a list of non-scored items.

Each item stem on the assessment has a range of performance choices, denoted by a letter score, which is chosen based on comprehensive criteria available in the manual. See Table 2 for an example of one of the item stems, I-6 Tongue, along with its associated performance options. A letter score is selected based on the infant's performance. This letter score is then converted to a numerical score based on the infant's PMA at the time of evaluation. See Table 3 for an example of how the numerical score based on the performance choice for item I-6 Tongue can differ based on the infant's PMA.

Each item is worth one, three, or five points (unless specified that the item is not scored). For each item, a score of 5 reflects a normal performance, a score of 3 reflects a questionable performance, and a score of 1 reflects a feeding challenge for that PMA. A questionable performance indicates that the skill may be emerging, may be impacted by concurrent medical complications, or may indicate abnormal performance. For infants who are orally feeding, all 18 scorable items are then tallied for a total score, with scores ranging from 18-95. For infants not yet orally feeding, an abbreviated form consisting only of the Pre-Feeding scores, the first 7 items of the NEO Assessment, can be tallied with scores ranging from 7-35. Therefore, there can be two scores generated by the NEO Assessment, a Pre-Feeding score (used when infants are not yet orally feeding and consists of the first 7 Pre-Feeding items) and the total score (used when infants are orally feeding and consists of scoring and tallying all 18 scorable items).

The NEO Assessment is available through the Washington University Office of Technology Management (24).

Tool Development

A review of the literature confirmed that there is a lack of available tools to assess neonatal feeding that have adequate psychometrics and adjust scoring based on developmental changes across PMA (25). The NEO Assessment was developed by outlining key components important for feeding success. Twenty-four factors related to oral motor skills and feeding performance were identified and used as item stems on the initial draft of the NEO Assessment, developed by author Roberta Pineda.

Then, a list of performance options for each item stem was defined, based on observations of different presentations of feeding and oral motor behaviors. To do this, one hundred and seventy eight preterm infants born 32 weeks gestation were enrolled as part of two overarching studies from 2007-2011 (26, 27). Infants were prospectively enrolled and were free of congenital anomalies. Infants had oral feedings videorecorded and evaluated between 35 to 41 weeks PMA. Videos consisted of a lateral view of the infant's face to enable visibility of the nipple in contact with the lips, as well as good visibility of jaw movement. Videos consisted of the first two minutes of oral feeding. These videos were used to define variance in feeding behaviors to ensure performance options for each item stem represented the significant variability in performance across multiple different infants. Four to six possibilities of performance for each item stem were defined. The item stem and performance choices went through multiple revisions with a goal of making them objective, consistent, and standardized, regardless of the infant's PMA. However, the score assigned to each item was intended to be based on expected performance related to PMA, and the development of the scoring matrix occurred later through the content validity process.

Content Validity

Fourteen experts were recruited via email for participation in the content validity portion of the study. These professionals were occupational therapists, speech-language pathologists, neonatal nurses, or neonatologists with a minimum of five years of experience in evaluating, treating, and/or managing feeding irregularities in the NICU. These participants were

considered to be knowledgeable experts in neonatal feeding, based on publication, presentation, or clinical expertise in the area of neonatal feeding.

After signing a non-disclosure agreement, content experts were sent packets via mail containing the NEO Assessment manuals and score sheets. Two participants withdrew, and one participant's packet was lost in the mail. A total of 11 experts provided feedback that was used in evaluating the content validity of the NEO Assessment.

Each expert reviewed the assessment and responded to the following question for each item: "Is the skill or knowledge measured by this item essential, useful but not essential, or not necessary to the performance of the construct?". Structured questions about the whole assessment included: 1) "Are all the relevant feeding behaviors included? If not, what is missing?", 2) "Are there any feeding behaviors that you might take out? If so, why?", and 3) "Are the scoring distinctions meaningful and clear? If not, please describe any problems you identify".

There were varied responses, and frequencies and percentages of answers were calculated for each assessment item. None of the items were deemed non-essential and removed from the assessment based on the expert's feedback. However, some of the items on the NEO Assessment (version 1) were combined, reducing the number of items from 24 to 19. Methods to determine feeding readiness and safety during oral feeding were added to the NEO Assessment in order to signal the clinician when the assessment may or may not be appropriate. Observations related to the breastfeeding process were better defined. Finally, the items related to swallow and rate necessitated clarification and rewording, prompting extensive revision.

Determining Scoring Based on PMA

Experts also provided feedback on each performance choice for each item stem and whether he/she perceived the behavior to be normal, abnormal, or questionable at each PMA (30, 31, 32, 33, 34, 35, 36, 37, 38, 39, and 40 weeks PMA). This feedback guided PMA-based item scoring (5=normal, 3=questionable, and 1=abnormal). PMA behavior categorizations were determined when greater than half of the experts had agreement about a specific item. Consensus was achieved often. However, occasionally agreement of more than half of the experts was not evident. In these cases the author (RP) incorporated expert feedback and carefully considered each behavior at the identified PMA to determine scoring. Expert feedback on clarity and meaningfulness of the assessment item and answer choices were considered. Most often, in these cases, the behavior was categorized as questionable, which best represented some disagreement among the experts in how the behavior should be defined. Significant modifications to the NEO Assessment were made based on the expert feedback and new scoring distinctions, resulting in NEO Assessment version 2.

Further Revision of the NEO Assessment and Preliminary Reliability with the Research Team

Fifty preterm infants who were born $\geq$ 32 weeks gestation and free of congenital anomalies were enrolled by day of life 7 from January to June 2015. In addition, 50 full term infants who were born >37 weeks gestation and free of congenital anomaly and any medical

interventions or complications at birth were enrolled by day of life 4 from March through June 2016. Preterm infants had videorecordings of feeding once oral feeding was initiated and then weekly until term equivalent age. Full term infants had an oral feeding videorecorded one time prior to day of life 4 and discharge from the mother-baby floor. Videos were done with a lateral view so the lips contacting the nipple were visible. The entire feeding or 20 minutes, whichever came first, was videorecorded. Parents and health care professionals were instructed to feed the infant as they would normally, and no instructions on how to feed were provided by the person recording. All of the preterm infant videos were of bottle feeding, but 24 of 50 (48%) of the full term infants were breastfed during videorecordings.

Seven members of the Washington University Neonatal and Developmental Research Team participated in NEO Assessments and engaged in group discussions. One was a neonatal occupational therapist, one was a psychologist, and five were occupational therapy graduate students at Washington University. Training occurred over a two hour period and was given by the occupational therapist (author, RP) who developed the tool. The research team observed selected videos, that had quality adequate for scoring, from the preterm and full term cohorts. Team members scored the NEO Assessment from the video independently, followed by engaging in open discussion among the group in an attempt to identify areas of the NEO Assessment that required revision and to reach agreement on appropriate scoring. Revisions were related to improving clarity in each item stem, making each performance choice mutually exclusive, depicting options for the range of feeding behaviors that were observed, and simplifying wording. The process of reviewing 3-5 videos, followed by scoring and discussing, continued for approximately 7 rounds with different feeding videos selected each round. Modifications were made to the tool, until consensus on the scoring and agreement about wording was achieved. Following agreement, team members independently scored 5 videos using the revised NEO Assessment. The Pre-Feeding intraclass correlation (ICC) was 0.73 (good), and the total score ICC was 0.92 (excellent).

Procedures: Inter-rater Reliability with Neonatal Therapists

Following development, testing, group discussion, and revision, reliability of the NEO Assessment version 4 was assessed with a convenience sample of six neonatal therapists. Neonatal therapists with at least 3 years of experience with feeding in the NICU were recruited for reliability testing. Two separate groups of therapists (one with four therapists and one with two therapists) were trained and conducted reliability testing. The training and reliability testing was conducted over Gotomeeting due to the therapists living in various locations throughout the United States. Trainings were designed to last one hour followed by reliability testing, however, the first training group needed to complete the training over a few sessions (over approximately three hours) due to significant dialogue generated among the group.

Five high quality videos of infants born 32 weeks estimated gestational age and currently orally feeding (with a range of age between 35-40 weeks PMA) were viewed on each rater's own computer screen, with videos presented in random order following the training. All infants in the videos selected were bottle fed. Videos were selected based on whether they

captured the infant's lips, cheek, and jaw line for the duration of the feeding and such that the quality was high enough to enable raters to visualize the infant. Raters were blinded to medical history, with the exception of knowing the current PMA, and raters scored the feedings using the NEO Assessment individually.

Letter assignments for each item were reported by each rater (due to criteria on each of the 19 items on the assessment). Each rater converted the letter score to a numerical score (1, 3, or 5), based on the infant's PMA (as previously described). Each rater also tallied a Pre-Feeding Score (adding all 6 scored Pre-Feeding items) and a total score (adding all 18 scored items) and converted them to categorical scores indicating feeding challenge, questionable or normal performance.

Statistical Analysis

Researchers utilized the Research Electronic Data Capture (REDCap) program by having each participant enter their assessment scores into a REDCap survey. Fleiss' Kappa statistics were used to determine reliability between raters on the letter score for each item, as well as to determine the reliability on the categorical score for both the "Pre-Feeding" and "total" score across raters. Inter-rater reliability of the "Pre-Feeding" and "total" continuous scores was also determined through intraclass correlations (28). A biostatistician, author MW, performed all statistical analyses using R 3.2.0.

Results

There were a total of six neonatal therapists who participated in reliability testing. Years of experience ranged from 4-39 years with a mean of 22 ± 15.1 years. All were neonatal occupational therapists: two had 4 years of experience in the NICU, one had 7 years of experience in the NICU, and two had greater than 30 years of experience in the NICU. Two of those who participated in the reliability testing also had participated in content validity.

Refer to Table 4 for the letter scores for each item, along with the calculated Fleiss' Kappa scores. Kappa scores <0 are considered to have poor agreement, 0.01-0.020 to have slight agreement, 0.21-0.40 to have fair agreement, 0.41-0.60 to have moderate agreement, 0.61-0.80 to have substantial agreement, and 0.81-1.00 to have almost perfect agreement (29).

The ICC of the continuous NEO Assessment Pre-Feeding score was 0.68 (confidence interval: 0.35-0.95), which is considered good. The ICC of the total continuous score was 0.83 (confidence interval: 0.56-0.98), which is considered excellent. ICC scores of less than 0.40 are considered poor, 0.40-0.59 are considered fair, 0.60-0.74 are considered good, and 0.75-1.00 are considered excellent (30).

Fleiss Kappa on the Pre-Feeding categorical score was 0.32 ($z=3.75$; $p<0.001$) and Fleiss Kappa on the total categorical score was 0.34 ($z=3.81$; $p<0.001$). These categorical scores were considered to have fair agreement.

Following the reliability testing, extensive conversation occurred with several of the neonatal therapists who participated. Revisions were made to the NEO Assessment to aid clarity,

optimize applicability to the clinical setting, and better capture breastfeeding and environmental factors in Section IV.

Discussion

Feeding challenges are widely recognized as a primary concern in preterm infants as they directly impact nutrition, developmental outcome and parent interactions/social adaptation. Neonatal feeding interventions are a vital component of care but are currently limited due to the lack of available feeding measures. The Neonatal Eating Outcome Assessment may fill the need for a developmental feeding assessment for preterm infants in the NICU. A rigorous process of tool development was undertaken, and preliminary results show that the NEO Assessment has content validity and good to excellent inter-rater reliability.

Feedback on the NEO Assessment from the neonatal feeding experts was positive, and perceptions about feeding performance prior to term equivalent age were used to refine the scoring matrix for each item, with scoring based on the infant's current PMA. Research team testing and discussions enabled revisions to scoring to aid clarity and ease of use. Reliability testing with neonatal therapists revealed good to excellent reliability on the total score of the NEO Assessment. However, Kappa statistics on each letter score (on the scored items of the NEO Assessment) provided important information to drive additional revisions.

Although the Kappa statistics of the categorical total score had fair agreement, there were some individual items that had slight agreement. Fleiss' Kappa scores for suck-swallow-breathe coordination, oral tone, feeding efficiency, suction, enjoyment of feeding, respiratory control during feeding, quality of sucking movements, and swallow informed the need for further revision of those items. The scoring criteria for these items were further revised based on feedback from raters and the research team (resulting in version 5).

While version 5, which was achieved from the process defined in this manuscript, is a final form that is now available to other clinicians and researchers, the tool will continue to undergo improvements, using statistical approaches to inform them. Using the final form, version 5, inter-rater reliability testing will be conducted again with neonatal therapists. RASCH analysis will be used to further describe reliability. In addition, other measures of neonatal feeding will be compared to the NEO Assessment to determine concurrent validity. Finally, predictive validity will be assessed by determining the ability of the NEO Assessment to define long term challenges with feeding. The need for a formal training and certification process will also be evaluated to ensure that professionals using the NEO Assessment have achieved adequate training to ensure good reliability.

Limitations

This study was limited in that participants used the NEO Assessment to evaluate videos of preterm infant feeding rather than using the assessment in a natural, face-to-face context.

Therefore, participants had to make assumptions regarding information not visible in the videos, such as infant arousal and non-nutritive sucking prior to feeding, as well as the level of support required for suck-swallow-breath coordination. These assumptions could have

hindered the participant's ability to score all criteria of the NEO Assessment. Additionally, the trainings and intra-rater reliability testing were held on an online virtual meeting which was susceptible to poor internet connection. The videos used for reliability testing were from infants between 35-40 weeks PMA, despite the tool being appropriate to use with infants from when oral feeding is initiated (as early as 30-32 weeks) to 40 weeks PMA. However, at the study site, very preterm infants were often not fed earlier than 35 weeks, resulting in a limited pool of videos to choose from for reliability testing. This study also did not assess reliability of infants who were not yet orally feeding using the abbreviated Pre-Feeding section only. Further assessment of the Pre-Feeding section of the NEO Assessment among infants who are early in PMA and not yet orally feeding is warranted. All videos used for reliability testing were of infants being bottle fed. As the tool is appropriate for breastfed infants, additional reliability testing of breastfed infants is also warranted. Finally, some letter scores of items, such as non-nutritive sucking and physiological stability, are challenging to interpret due to lack of variability in scoring. Carefully selecting infants who will have variable performance on specific items will allow better assessment of inter-rater reliability of those items in the future.

Conclusion

The purpose of this manuscript was define the process used to develop, achieve content validity, and assess reliability of the NEO Assessment. Significant progress in tool development with the NEO Assessment is reported here. The NEO Assessment is the first developmental feeding assessment to enable assessment of feeding performance across PMA. It holds promise for being a clinically useful and psychometrically sound tool, but more research is needed.

Acknowledgments

We would like to thank Robin Glass, Lynn Wolf, Erin Ross, Sandra Carroll, Peter Bingham, Gay Loyd Pinder, Tina Tan, Kara Waitzman, Elizabeth Clawson, Mary Stanford, Marjorie Meyer Palmer, Annie Aloysius, Janet Hamill, Sue Ludwig, Laura Madlinger-Lewis, Shannon Usher, Katie Bogan, Sarah Oberle, Jeanne Kloeckner, Jessica Roussin, Alan Jette, Wendy Coster, Roger Felding, Mary Slavin, Kelsey Dewey, Elizabeth Heiny, Katie Ross, Justin Ryckman, Sarah Wolf, Anna Annecca, Lisa Shabosky, Joy Bender, Kelsey Melchior, Laura Mazelis, Hayley Chrzastowski, Odochi Nwabara, Sonya Dunsirn, and Bailey Hall.

Funding Statement: This project was supported by NIH R24 (5R24HD065688-05) awarded to the Boston Rehabilitation Outcomes Center. Research reported in this publication was also supported by the National Institute of Health Comprehensive Opportunities for Rehabilitation Research Training (CORRT) Grant (K12 HD055931) and the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number U54 HD087011 to the Intellectual and Developmental Disabilities Research Center at Washington University.

References

1. Samara M, Johnson S, Lamberts K, Marlow N, Wolke D. Eating problems at age 6 years in a whole population sample of extremely preterm children. *Dev Med Child Neurol*. 2010; 52:e16–22. [PubMed: 19832883]
2. de Kieviet JF, Piek JP, Aarnoudse-Moens CS, Oosterlaan J. Motor development in very preterm and very low-birth-weight children from birth to adolescence: A meta-analysis. *JAMA*. 2009; 302:2235–42. [PubMed: 19934425]
3. Thoyre SM, Shaker CS, Pridham KF. The early feeding skills assessment for preterm infants. *Neonatal Netw*. 2005; 24:7–16.

4. Philbin MK, Ross ES. The SOFFI Reference Guide: text, algorithms, and appendices: a manualized method for quality bottle-feedings. *J Perinat Neonatal Nurs.* 2011; 25:360–80. [PubMed: 22071621]
5. Ludwig SM, Waitzman KA. Changing Feeding Documentation to Reflect Infant-Driven Feeding Practice. *Newborn Infant Nurs Rev.* 2007; 7:155–60.
6. Lau C, Alagugurusamy R, Schanler RJ, Smith EO, Shulman RJ. Characterization of the developmental stages of sucking in preterm infants during bottle feeding. *Acta Paediatr.* 2000; 89:846–52. [PubMed: 10943969]
7. Lau C, Smith EO. A novel approach to assess oral feeding skills of preterm infants. *Neonatology.* 2011; 100:64–70. [PubMed: 21212698]
8. Palmer MM, Crawley K, Blanco IA. Neonatal Oral-Motor Assessment scale: a reliability study. *J Perinatol.* 1993; 13:28–35. [PubMed: 8445444]
9. Zarem C, Kidokoro H, Neil J, Wallendorf M, Inder T, Pineda R. Psychometrics of the neonatal oral motor assessment scale. *Dev Med Child Neurol.* 2013; 55:1115–20. [PubMed: 23869958]
10. Rogers B, Arvedson J. Assessment of infant oral sensorimotor and swallowing function. *Ment Retard Dev Disabil Res Rev.* 2005; 11:74–82. [PubMed: 15856438]
11. Einarsson-Backes LM, Deitz J, Price R, Glass R, Hays R. The effect of oral support on sucking efficiency in preterm infants. *Am J Occup Ther.* 1994; 48:490–8. [PubMed: 8067370]
12. McGrath JM, Braescu AV. State of the science: feeding readiness in the preterm infant. *J Perinat Neonatal Nurs.* 2004; 18:353–68. quiz 69-70. [PubMed: 15646306]
13. Lau C. Oral feeding in the preterm infant. *Neoreviews.* 2006; 7:e19–e27.
14. Amaizu N, Shulman R, Schanler R, Lau C. Maturation of oral feeding skills in preterm infants. *Acta Paediatr.* 2008; 97:61–7.
15. Ross ES, Browne JV. Developmental progression of feeding skills: an approach to supporting feeding in preterm infants. *Semin Neonatol.* 2002; 7:469–75. [PubMed: 12614599]
16. Medoff-Cooper B, Shults J, Kaplan J. Sucking behavior of preterm neonates as a predictor of developmental outcomes. *J Dev Behav Pediatr.* 2009; 30:16–22. [PubMed: 19194323]
17. Lau C, Smith EO, Schanler RJ. Coordination of suck-swallow and swallow respiration in preterm infants. *Acta Paediatr.* 2003; 92:721–7. [PubMed: 12856985]
18. Griffith T, Rankin K, White-Traut R. The Relationship Between Behavioral States and Oral Feeding Efficiency in Preterm Infants. *Adv Neonatal Care.* 2017; 17:E12–E9.
19. Pickler RH, Best AM, Reyna BA, Wetzel PA, Gutcher GR. Prediction of Feeding Performance in Preterm Infants. *Newborn Infant Nurs Rev.* 2005; 5:116–23. [PubMed: 16467910]
20. Thoyre SM, Carlson J. Occurrence of oxygen desaturation events during preterm infant bottle feeding near discharge. *Early Hum Dev.* 2003; 72:25–36. [PubMed: 12706309]
21. Gewolb IH, Vice FL. Abnormalities in the coordination of respiration and swallow in preterm infants with bronchopulmonary dysplasia. *Dev Med Child Neurol.* 2006; 48:595–9. [PubMed: 16780630]
22. Slocum C, Arko M, Di Fiore J, Martin RJ, Hibbs AM. Apnea, bradycardia and desaturation in preterm infants before and after feeding. *J Perinatol.* 2009; 29:209–12. [PubMed: 19148108]
23. Daily DK, Ellison PH. The premie-neuro: a clinical neurologic examination of premature infants. *Neonatal Netw.* 2005; 24:15–22.
24. Pineda, RG. Neonatal Eating Outcome Assessment. 2016. [Available from: <https://otm.wustl.edu/technologies/neonatal-eating-outcome-neo-assessment/>]
25. Bickell M, Barton C, Dow K, Fucile S. A systematic review of clinical and psychometric properties of infant oral motor feeding assessments. *Dev Neurorehabil.* 2017:1–11.
26. Pineda RG, Tjoeng TH, Vavasseur C, Kidokoro H, Neil JJ, Inder T. Patterns of altered neurobehavior in preterm infants within the neonatal intensive care unit. *J Pediatr.* 2013; 162:470–6 e1. [PubMed: 23036482]
27. Madlinger-Lewis L, Reynolds L, Zarem C, Crapnell T, Inder T, Pineda R. The effects of alternative positioning on preterm infants in the neonatal intensive care unit: a randomized clinical trial. *Res Dev Disabil.* 2014; 35:490–7. [PubMed: 24374602]
28. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979; 86:420–8. [PubMed: 18839484]

29. Landis J, Koch G. The measurement of observer agreement for categorical data *Biometrics*. 1977; 33:159–74. [PubMed: 843571]
30. Cicchetti D. Guidelines, criteria, and rules of thumb evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994; 6:284–90.

Abbreviations

NICU	neonatal intensive care unit
NEO	Neonatal Eating Outcome
ICC	intra-class correlation
NOMAS	Neonatal Oral Motor Assessment Scale
EGA	estimated gestational age
PMA	postmenstrual age

Key Notes

- Evaluation and treatment of feeding relies on the availability of standardized, reliable, and valid feeding tools.
- The Neonatal Eating Outcome (NEO) Assessment is a new developmental feeding assessment for high-risk infants aimed at assessing oral motor and feeding skills in preterm infants prior to and at term equivalent age, with scoring based on postmenstrual age.
- This manuscript defines the process of tool development and revision and reports preliminary inter-rater reliability.

Table 1

Scored and non-scored items of the NEO Assessment.

<p>Pre-Feeding BehaviorsSection 1 Observation of non-nutritive sucking on a pacifier (or gloved finger) for 1-3 minutes. If the infant is orally feeding, this section is scored during the time prior to and at bottle nipple insertion or latch to the breast for oral feeding.</p>	I-1. Arousal/State Organization
	Level of ease of arousing the infant for the feeding within 15 minutes of the scheduled feeding/care time. Includes observation of the initial state, and if no spontaneous waking is evident, re-assessment following a diaper change, gently changing the infant's position, and/or assessing vital signs.
	I-2. Physiological Stability
	Stability of the infant's heart rate, oxygen saturation level, and respiratory rate.
	I-3. Respiratory Support
	Amount of respiratory support being delivered to the infant as their standard of care.
	I-4. Rooting and Grasp
	Infant's demonstration of the rooting reflex and grasp in preparation for feeding. It is also an assessment of perioral/facial sensitivity. The rooting and grasp response should be assessed prior to oral feeding by gently stroking each side of the mouth with a finger, pacifier, or breast and observing the infant's response. If a response cannot be seen on the sides of the mouth, stroke the upper and lower lip surfaces.
	I-5. Initiation of Sucking
	Infant's response to the bottle nipple being placed in the mouth or the response to latching onto the breast. Scored based on the overall responsiveness to the nipple/breast in the mouth, not a single incident or one observation. Sucking initiation is defined as demonstrating a minimum of 3 sucks in a 5-second period.
<p>Oral FeedingSection II Observation of first 10 minutes of oral feeding. Items are scored based on the criteria that best represents the infant's performance.</p>	I-6. Tongue
	Infant's tongue position and response in relation to finger, nipple, breast or pacifier. Scored based on the predominant tongue behavior that appears to persist, affect performance and/or catch attention.
	I-7. Non-Nutritive Sucking
	Number of sucks and how many there are in relation to the rest breaks. Observe if there is a predictable pattern and how frequent each suck is.
	II-1. Suck-Swallow-Breathe Coordination
	Infant's ability to coordinate sucking, swallowing, and breathing for successful oral feeding, with particular focus on the first two minutes of oral feeding when the infant adjusts to the flow of milk and then achieves a pattern of suck-swallow-breathe.
	II-2. Sucking Burst Length
	Infant's feeding efficiency after the initial adjustment to oral feeding in the first minute(s) but before fatigue. Scored after the first minute of oral feeding after the infant has gotten into a pattern, but within the first 5 minutes of the oral feeding.
	II-3. Suction
	Infant's ability to form a vacuum and enable movement of liquid from the bottle or breast.
	II-4. Oral Tone
	Infant's oral tone, including the tongue, cheeks and lips.
	II-5. Quality of Sucking Movements
	Measures the extent to which coordinated jaw and tongue movements support rhythmic sucking movements in the first 10 minutes of oral feeding.
	II-6. Behavioral Response to Feeding
	Observations of positive or negative engagement feeding behaviors (i.e. social interaction, calm behavior, stress signs, aversion, etc.). Scored by the most prevalent

	characterization but observe for signs related to discomfort in the first 10 minutes of oral feeding.	
	II-7. Fluid Loss	
	Measures how much fluid is lost in expressing liquid from the breast/bottle in order to assess the efficiency of the infant's feeding skills.	
	II-8. Swallow	
	Infant's ability to safely swallow with no clinical evidence of aspiration. Scored with careful observation of the quality and pattern of swallows during the feeding.	
	II-9. Respiratory Control During Feeding	
	Infant's respiration coordination in response to feeding. Scored with observation for indications of increased work of breathing and/or airway noises such as stridor.	
Observations at the End of FeedingSection III Scored after feeding has been stopped by the caregiver or the infant or after 20 minutes of oral feeding.	III-1. Feeding Completion	
	Evaluation of why the feeding is discontinued and/or to define behaviors that are present after a 20-minute attempt at oral feeding.	
	III-2. State Maintenance	
	Assessment of what state an infant is able to maintain for 20 minutes of oral feeding. Observation of how long the infant was able to maintain an awake state prior to stimulation (the NEO Assessment manual includes detailed explanation of infant states).	
	III-3. Volume Consumed	
	Evaluation of difference between the starting and ending volume in the bottle (or pre/post weights for breastfeeding infants) to determine how much liquid was expressed or consumed, minus any spillage.	
Items Not ScoredSection IV Captures environmental, positioning, and adaptive strategies used during the assessment process	Section IV Items Not Scored	
	IV-1. Feeder IV-2. Feeder Experience IV-3. Infant Feeding Experience IV-4. Infant's Current Feeding Regime IV-5. Nipple Type IV-6. Milk Type IV-7. Interruptions	IV-8. External Supports Provided IV-9. Position During Feed IV-10. Heart Rate IV-11. Respiratory Rate IV-12. Oxygen Saturation IV-13. Time IV-14. Infant Anatomic and Functional Variations
	Section IV Items Not Scored—Breastfeeding Only	
	IV-15. Mother's Milk Supply IV-16. Pumping Schedule IV-17. Maternal Comfort Level IV-18. Breastfeeding Contraindications IV-19. Mother's Flow Rate IV-20. Breast (s) IV-21. Maternal Anatomy	IV-22. Maternal Complications IV-23. Managing Breastfeeding IV-24. Maternal Motor Factors IV-25. Maternal Environmental Factors IV-26. Mother-Infant Dynamics

Table 2

Example of performance choices (labeled with a letter) associated with item stem. **I-6. Tongue** - This item describes what the tongue does in relation to the finger, nipple, breast, or pacifier. This should be **scored based on the predominant position and response of the tongue that appears to persist, affect performance, and/or catch attention**. For infants not yet orally feeding, it may be ideal to use a gloved finger for non-nutritive sucking to appropriately feel and score this item. Observations of the tongue before and after nipple insertion or latch at the breast can also be used.

A.	<u>Flaccid or non-responsive</u>	The tongue lies flat in the mouth with absent central groove. The infant may demonstrate up and down jaw movements but the tongue does not participate in the sucking movements. A score of 'A' should also be given if the tongue fails to move with the presence of the nipple, breast, finger or pacifier. Infants with a flaccid tongue may also demonstrate low tone in jaw and lips (see Section II-4).
B.	<u>Flat with some tongue cupping</u>	The tongue is predominantly flat with lack of a central groove, but some or intermittent cupping may be observed. Tongue cupping may result in flat and thin tongue positioning with a wide trough down middle of tongue. Score this item if the tongue appears wide. The lateral aspects of the tongue can intermittently be visualized during sucking. Overall tone may be low.
C.	<u>Elevated and retracted</u>	Prior to or after insertion of the nipple, breast, finger or pacifier, the tongue is either elevated and/or retracted in the mouth. The nipple, breast, finger or pacifier may be inserted under the tongue. The feeder may need to put the bottle in the infant's mouth on the side or use strategies to enable the tongue to drop down. Jaw clenching may be prevalent before or after nipple, breast, pacifier, or finger insertion. Overall facial tone may be high. Score 'C' if it is challenging to open the infant's mouth to place the nipple, pacifier or finger. For breastfeeding, score 'C' if the infant has difficulty dropping the tongue down or if there is increased tightness of the tongue that causes pain. As the infant opens the mouth, the tongue may be observed to be elevated and retracted in the mouth. Strategies may be needed to bring the tongue down. Jaw clenching also may be observed. The mother's nipple may be lipstick shaped or flattened with a compression stripe when removed from the baby's mouth.
D.	<u>Tongue cupping with central groove</u>	The tongue cups around the nipple, pacifier, or finger and molds to it with a prevalent central groove. During bottle-feeding, the tongue is not seen during active sucking, because it remains cupped around the nipple, pacifier, or finger within the labial borders. For breastfeeding, score 'D' if infant achieves wide-open mouth posture and a U-shaped configuration of the tongue is observed as the baby is going toward the breast. The tongue is down and cupped around the breast, with much of the areola in the mouth. The tongue may be visualized during breastfeeding due to the wide-open mouth posture.
E.	<u>Deviated to side, tongue thrusting, or tongue bunching</u>	<i>1) The tongue is not held symmetrically in the mouth. OR 2) The infant thrusts the tongue out of the mouth before or after nipple, pacifier or finger insertion. OR 3) The infant bunches the tongue in the middle and although the tongue does not extend out of the mouth, the nipple, pacifier or finger is pushed out by the shape and position of the tongue.</i>

Table 3

Scoring based on PMA.

			Postmenstrual Age (weeks)											Score	
Item			Choices	30	31	32	33	34	35	36	37	38	39	40+	
Section IPre-Feeding Behaviors			A. Flaccid or non-responsive B. Flat with some tongue cupping C. Elevated and retracted D. Tongue cupping with central groove E. Deviated to side, tongue thrusting, or tongue bunching		E=1 A, B, C=3 D=5	A, E=1 B, C=3 D=5	A, B, E=1 C=3 D=5								A, B, C, E=1 D=5

Table 4

Letter scores assigned by each of 6 raters and 5 infants and associated Fleiss' Kappa statistics on each scorable item of the NEO Assessment.

Item I-1: Arousal/State Organization							Item I-2: Physiological Stability						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	D	D	D	D	D	D	Baby 1	D	D	D	D	D	D
Baby 2	D	D	D	D	D	D	Baby 2	D	D	C	C	D	D
Baby 3	B	A	B	B	A	A	Baby 3	D	D	D	D	D	D
Baby 4	D	D	D	D	C	D	Baby 4	D	D	D	D	D	D
Baby 5	C	C	C	B	B	D	Baby 5	D	D	D	D	D	D
Kappa=0.43 (Moderate Agreement) z=5.81 p<0.001							Kappa=0.143 (Slight Agreement) z=1.24 p=0.22						
Item I-3: Respiratory Support							Item I-4: Rooting and Grasp						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	D	D	D	D	D	D	Baby 1	B	C	C	C	B	E
Baby 2	D	D	D	D	D	D	Baby 2	D	C	D	D	C	D
Baby 3	D	D	D	D	D	D	Baby 3	A	B	A	B	A	A
Baby 4	D	D	D	D	D	D	Baby 4	D	D	D	D	D	D
Baby 5	D	D	D	D	D	D	Baby 5	B	C	B	B	B	C
*Statistics could not be calculated due to consistency of scoring across all infants							Kappa=0.37 (Fair Agreement) z=5.69 p<0.001						
Item I-5: Initiation of Sucking							Item I-6: Tongue						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	B	C		B	E	E	Baby 1	E	E	E	E	E	E
Baby 2	D	D	D	D	C	D	Baby 2	B	C	D	D	C	D
Baby 3	B	C	C	B	B	A	Baby 3	C	D	D	D	B	A
Baby 4	D	D	D	D	D	D	Baby 4	D	D	D	D	D	D
Baby 5	D	C		B	A	B	Baby 5	C	D	C	D	C	C
Kappa=0.24 (Fair Agreement) z=3.58 p<0.001							Kappa=0.39 (Fair Agreement) z=5.59 p<0.001						
Item I-7: Non-Nutritive Sucking							Item II-1: Suck-Swallow-Breathe Coordination						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	D	D	D	D	D	D	Baby 1	C	C	D	C	C	C
Baby 2	D	D	D	D	D	D	Baby 2	C	B	B	B	B	B

Item I-1: Arousal/State Organization							Item I-2: Physiological Stability						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 3	D	C	D	D	D	D	Baby 3	C	D	C	C	C	B
Baby 4	D	D	D	D	D	D	Baby 4	C	C	C	D	C	C
Baby 5	D	D	D	D	D	D	Baby 5	C	C		D	C	C
Kappa=-0.04 (Slight Agreement) z=-0.30 p=0.77							Kappa=0.18 (Slight Agreement) z=1.75 p=0.08						
Item II-2: Feeding Efficiency							Item II-3: Suction						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	B	C	B	B	B	C	Baby 1	B	B	C	B	B	C
Baby 2	C	C	C	C	C	C	Baby 2	C	C	C	C	C	C
Baby 3	C	C	B	B	B	B	Baby 3	C	C	C	C	C	C
Baby 4	B	C	B	C	B	C	Baby 4	C	C	B	C	B	C
Baby 5	D	C		D	C	C	Baby 5	D	C	C	C	B	C
Kappa=0.17 (Slight Agreement) z=1.84 p=0.07							Kappa=0.18 (Slight Agreement) z=1.75 p=0.08						
Item II-4: Oral Tone							Item II-5: Quality of Sucking Movements						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	B	D	C	B	C	C	Baby 1	B	C	B	C	B	D
Baby 2	B	B	C	C	C	C	Baby 2	C	C	C	D	B	D
Baby 3	D	B	C	C	B	C	Baby 3	C	C	C	C	B	D
Baby 4	C	C	C	C	C	C	Baby 4	C	D	B	D	C	D
Baby 5	D	C	D	C		C	Baby 5	A	C	C	C	C	D
Kappa=-0.01 (Slight Agreement) z=-0.15 p=0.89							Kappa=-0.07 (Slight Agreement) z=-0.84 p=0.40						
Item II-6: Enjoyment of Feeding							Item II-7: Fluid Loss						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	C	B	D	C	A	C	Baby 1	A	C	C	D	C	C
Baby 2	C	C	C	C	B	D	Baby 2	D	D	D	D	D	D
Baby 3	B	C	C	C	B	B	Baby 3	C	C	C	C	B	C
Baby 4	C	D	B	C	B	C	Baby 4	C	D	D	D	D	C
Baby 5	B	D	D	D	B	D	Baby 5	B	C	C	C	C	C
Kappa=0.02 (Slight Agreement) z=0.26 p=0.80							Kappa=0.38 (Fair Agreement) z=4.15 p<0.001						

Item I-1: Arousal/State Organization							Item I-2: Physiological Stability						
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Item II-8: Swallow													
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	D	C	D	D	C	D	Baby 1	D	D	D	B	D	D
Baby 2	C	D	C	D	D	D	Baby 2	D	C	C	B	C	D
Baby 3	C	D	D	D	B	D	Baby 3	D	D	D	D	C	D
Baby 4	D	D	D	D	C	D	Baby 4	B	C	C	C	D	B
Baby 5	C	D	D	D	D	D	Baby 5	D	D	D	C	B	D
Kappa=-0.15 (Slight Agreement) z=-1.43 p=0.15						Kappa=0.20 (Slight Agreement) z=2.28 p=0.02							
Item III-1: Feeding Completion													
Rater:	1	2	3	4	5	6	Rater:	1	2	3	4	5	6
Baby 1	B	B	B	B	B	B	Baby 1	B	B	B	B	B	C
Baby 2	B	D	B	B	B	B	Baby 2	C	E	E	C	B	E
Baby 3	D	D	D	D	B	D	Baby 3	D	D	D	D	A	A
Baby 4	B	D	B	B	B	A	Baby 4	E	E	E	C	E	E
Baby 5	D	D	D	D	D	D	Baby 5	D	D	D	D	B	E
Kappa=0.52 (Moderate Agreement) z=4.92 p<0.001						Kappa=0.34 (Fair Agreement) z=5.34 p<0.001							
Item III-3: Volume Consumed													
Rater:	1	2	3	4	5	6							
Baby 1	C	D	C	C	C	D							
Baby 2	D	E	E	E	F	E							
Baby 3	F	F	F	F	F	F							
Baby 4	E	E	F	E	E	D							
Baby 5	F	F	F	F	F	F							
Kappa=0.49 (Moderate Agreement) z=6.73 p<0.001													