Published in final edited form as:

Dev Med Child Neurol. 2013 December ; 55(12): . doi:10.1111/dmcn.12202.

Psychometrics of the neonatal oral motor assessment scale

Cori S Zarem¹, Hiroyuki Kidokoro², Jeffrey Neil^{3,2,4}, Michael Wallendorf⁵, Terrie Inder^{2,3,4}, and Roberta Pineda^{1,2}

¹Program in Occupational Therapy, Washington University School of Medicine, St Louis, Missouri, USA

²Department of Pediatrics, Washington University School of Medicine, St Louis, Missouri, USA

³Department of Neurology, Washington University School of Medicine, St Louis, Missouri, USA

⁴Department of Radiology, Washington University School of Medicine, St Louis, Missouri, USA

⁵Division of Biostatistics, Washington University School of Medicine, St Louis, Missouri, USA

Abstract

AIM—To establish the psychometrics of the Neonatal Oral Motor Assessment Scale (NOMAS).

METHOD—In this prospective cohort study of 75 preterm infants (39 females, 36 males) born at 30 weeks' or less gestation (mean gestational age 26.56wk, SD 1.90, range 23–30wk; mean birthweight 967.33g, SD 288.54, range 480–2240), oral feeding was videotaped before discharge from the neonatal intensive care unit (NICU) discharge. The NOMAS was used to classify feeding as normal, disorganized, or dysfunctional. Neurobehavior was assessed at term equivalent, and infants underwent magnetic resonance imaging. Children returned for developmental testing at 2 years corrected age. Associations between NOMAS scores and (1) neurobehavior, (2) cerebral injury and metrics, and (3) developmental outcome were investigated using ²-analyses, *t*-tests, and linear regression. For reliability, six certified NOMAS evaluators rated five randomly selected NOMAS recordings and re-scored them in a second randomized order. Reliability was calculated with Cohen's kappa coefficient.

RESULTS—Dysfunctional NOMAS scores were associated with lower Dubowitz scores [*t*=−2.14; mean difference −2.32 (95% confidence interval [CI] −0.157 to −4.49); *p*=0.036], higher stress on the NICU Network Neurobehavioral Scale (*t*=2.61; mean difference 0.073 [95% CI 0.017 to 0.129]; *p*=0.0110, and decreased transcerebellar diameter (*t*=−2.22; mean difference −2.04 [CI=−3.89 to −0.203]; *p*=0.03). No significant associations were found between NOMAS scores and 2 year outcome.

INTERPRETATION—Some concurrent validity was established with associations between NOMAS scores and measures of infant behavior and cerebral structure. The NOMAS did not show predictive validity in this study of preterm infants at high risk of developmental delay. Reliability was variable and suboptimal.

Full attainment of oral feeding is one of the most complex tasks of infancy^{1,2} and one of three criteria required for discharge from the neonatal intensive care unit (NICU).³ Infants in the NICU are often delayed in oral feeding and are at increased risk of oral-motor dysfunction, non-organic failure to thrive, and dysphagia.^{4–6} Additionally, poor feeding in neonates has been associated with subsequent developmental delay and feeding difficulties

in childhood.⁴ Children with feeding complications have higher incidences of brain injury, and often experience long-term motor, behavioral, and cognitive dysfunction.^{5–9} Early identification of these problems is important to enable interventions to optimize function,¹⁰ and relies on the availability of valid and reliable assessment tools.

The Neonatal Oral Motor Assessment Scale (NOMAS), developed by Marjorie Meyer Palmer in 1985, is a widely used neonatal feeding evaluation.^{2,7,11–13} The tool is not commercially available. Raters are certified through a 3-day course, conducted most often by Palmer. As a requirement set by Palmer, participants are certified at the end of the course if they correctly classify five videotaped feedings as ‘normal’, ‘disorganized’, or ‘dysfunctional’. The NOMAS is the only available neonatal feeding evaluation that can be used with term or preterm infants and with infants who are breast or bottle-fed.¹

The NOMAS is a 28-item observational checklist of tongue and jaw movement. Following observation of non-nutritive sucking, the first two minutes of oral feeding are evaluated. Feeding is classified as normal, disorganized, or dysfunctional.¹¹ Using the NOMAS, normal defines an infant demonstrating coordination of suck–swallow–breathe responses and feeding efficiency.¹³ Infants who demonstrate difficulty coordinating suck–swallow–breathe are classified as disorganized. Any abnormal movement interrupting the feeding process leads to a classification of dysfunctional.^{11,12} Dysfunctional characteristics include excessively wide jaw excursions interrupting the seal on the nipple, lateral jaw deviation, a flaccid/retracted tongue, or total absence of movement.¹³

Though the NOMAS was developed over 20 years ago, there is limited research investigating its psychometric properties.¹ One study reported that at age 2 years, 100% of infants classified as normal on the NOMAS will not have developmental delay, and 100% of infants classified as dysfunctional will have developmental delay.¹⁴ The ability of the NOMAS to predict developmental delay in infants categorized as normal or dysfunctional is cited during the certification course. However, this study assessed 17 infants and was retrospective, with many confounding factors, weakening its results and implications. Another study found the NOMAS to have modest internal consistency within the normal and disorganized categories (Cronbach $\alpha > 0.70$) and acceptable convergent validity for infants 32 to 35 weeks' postmenstrual age (Spearman's $r = 0.51–0.69$).² The study did not include calculations for the dysfunctional category, and called for further investigation of the NOMAS. Another study found the interrater reliability to be ‘moderate’ to ‘substantial,’ with kappa (κ) = 0.40 to 0.65 and intrarater reliability to be variable at $\kappa = 0.33$ to 0.95. This was considered unacceptable for a diagnostic tool.¹² The NOMAS has been associated with 12-month outcome; infants categorized as normal performed better on developmental tests than infants persistently categorized as disorganized.⁹ However, this study excluded infants with cerebral injury, which is common in preterm infants and other high-risk populations.

The ability of assessments like the NOMAS to identify those in need of early intervention could enhance the short and long-term care of preterm infants, and help minimize the effects of feeding related morbidities.¹⁰ It is therefore necessary to further investigate the psychometric properties of the NOMAS to provide clinicians with accurate information, and to work toward developing standardized and reliable tools to assess feeding in high-risk populations. The aim of this study was to evaluate the concurrent and predictive validity of the NOMAS, as well as its interrater and intrarater reliability.

METHOD

This research was part of a longitudinal study investigating the developmental trajectory of preterm infants. It was approved by the human subjects committee at the study site. Infants

were recruited from the 75-bed NICU of St Louis Children's Hospital, MO, USA. Parents provided informed consent for serial neurobehavioral evaluation, feeding assessment, and magnetic resonance imaging (MRI) during hospitalization, as well as developmental assessment at age the corrected age of 2 years.

Consecutive inborn infants admitted to the NICU were recruited from 2008 to 2010. Participants were born 30 weeks' gestation or less and recruited within 72 hours of birth. Infants with known or suspected congenital anomalies were excluded from this study. Infants were not excluded based on any research procedures/testing.

With a primary outcome of the NOMAS predicting the presence of developmental delay, a power analysis was conducted. A sample size of 50 would provide 80% power when the true difference in proportions is 0.31, using a Pearson's χ^2 test with an alpha of 0.05. A large difference in proportions was chosen because of previous reports of 100% predictive power of the NOMAS, with 100% of infants with feeding categorized as dysfunctional having developmental delay at 2 years corrected age.¹⁴

Independent variable: NOMAS

At term equivalent (37–42 weeks' gestational age), an oral feeding session was recorded using video, and scored using the NOMAS. Video recordings included a close-up lateral view of the neck, jaw, and mouth. The videos included a clear view of the lips in contact with the nipple. Videotaping commenced prior to placement of the nipple in the infant's mouth and was stopped after two minutes of oral feeding. The feeding was conducted by parents/family when present, therapists, or nurses. A single, certified evaluator determined the NOMAS category, based on bedside clinical observations and video analysis. This evaluator understood the study purpose to be investigating the longitudinal development of preterm infants.

Concurrent validity

Information collected from the medical records included birthweight, gestational age at birth, Critical Risk Index for Babies (CRIB) scores, and days of ventilation, continuous positive airway pressure, and total parenteral nutrition. The CRIB score is a measure of medical severity calculated in the first 12 hours of life.¹⁹ All medical factors were analyzed for associations with NOMAS scores.

At term equivalent, infants were assessed using the Dubowitz neurological examination (Dubowitz),¹⁵ NICU Network Neurobehavioral Scale (NNNS)¹⁶ and MRI testing. Infants also underwent routine cranial ultrasound to detect intraventricular hemorrhage at one week and one month of life.

The Dubowitz is a 34-item neurological assessment yielding an optimality score, which has been shown to decrease with increased severity of cerebral injury.^{15,17} The Dubowitz was used to analyze concurrent validity as a measure of neurological function.

The NNNS is a comprehensive assessment of neurobehavioral function yielding 13 summary scores including the following: habituation, orientation, handling, quality of movement, self-regulation, suboptimal reflexes, stress, arousal, hypertonia, hypotonia, asymmetry, excitability and lethargy.¹⁶ Summary scores were used to analyze concurrent validity as a measure of term neurobehavior.

Infants underwent sedative-free MRI to obtain anatomical images with an axial magnetization-prepared rapid gradient echo T1-weighted sequence (TR/TE 1500/3ms, voxel

size $1 \times 0.7 \times 1 \text{ mm}^3$) and a turbo spin echo T2-weighted sequence (TR/TE 8600/160ms, voxel size $1 \times 1 \times 1 \text{ mm}^3$, echo train length 17). Images were interpreted by a trained neuroradiologist. MRI and ultrasound results determined the presence of cerebral injury, defined as the presence of grade III–IV intraventricular hemorrhage, cystic periventricular leukomalacia, and/or cerebellar hemorrhage.

Brain metrics data were obtained using reference points on specific MRI slices, and measurements taken by computer analysis. These included bifrontal diameter, biparietal diameter (BPD), bone–biparietal diameter ratio (BBPD), interhemispheric distance (IHD), and transcerebellar diameter.¹⁸ To compare brain and ventricle size, another measure was calculated by dividing the ventricular diameter by the BBPD. This incorporates the IHD and BPD to correlate cerebral mass with skull size.

Diffusion tensor imaging (DTI) measures were obtained from MRI bilaterally in the following regions: anterior limb of the internal capsule, posterior limb of the internal capsule, optic radiation, frontal lobe, cingulum bundle, centrum semiovale, and the corpus callosum. DTI was interpreted by a trained evaluator.

To assess concurrent validity, associations between NOMAS scores and Dubowitz scores, NNNS summary scores, the presence of cerebral injury, brain metrics, and DTI measures were analyzed.

Predictive validity

Infants returned for testing at 2 years' corrected age, using the Bayley Scales of Infant Development, third edition (BSID-III).²⁰ The BSID-III is a norm-referenced assessment, considered to be the criterion standard in developmental evaluation.²⁰ Composite subscores for cognitive, language and motor skills were investigated. In addition, a categorical variable of developmental delay was created for analysis. Developmental delay was defined as having any BSID-III composite score less than 70; two standard deviations below the mean (100).^{20,21}

Recent evidence suggests that the BSID-III overestimates developmental scores.²¹ Adjustments are suggested to BSID-III scores to account for this difference.²¹ Analyses with both original and adjusted BSID-III scores were conducted.

Interrater and intrarater reliability

Six certified NOMAS raters completed reliability testing. Five NOMAS recordings were randomly selected, presented in a randomized order, and scored. Raters were blinded to medical history. After a minimum 2-week period, raters re-scored the same recordings, presented in another randomized order. Scores were collected and processed to determine interrater and intrarater reliability.

Statistical analysis

Independent samples *t*-tests and Mann–Whitney *U* tests were used to determine associations between the NOMAS and the NNNS and Dubowitz optimality scores. MRI/DTI outcomes were analyzed using linear regression, while controlling for postmenstrual age at ultrasound. Predictive validity was investigated using χ^2 analysis to determine the presence of developmental delay related to NOMAS scores, and independent sample *t*-tests to determine differences in BSID-III scores across NOMAS categories. A 95% confidence interval (CI) was calculated to represent the true value of dysfunctional infants with developmental delay at age 2 years. Additionally, a comparison between the proportion of dysfunctional infants in

this cohort and the proportion of dysfunctional infants in the 1999 Palmer cohort (100%)¹³ was made using Fisher's exact test. Cohen's kappa statistics were used to calculate reliability. Analyses were performed using IBM SPSS 20 software (IBM Corporation, NY, USA) SAS 9.3 software (SAS Institute, Cary, NC, USA), $\alpha=0.05$.

RESULTS

Ninety-eight infants were enrolled. Of these, five withdrew and 14 died, leaving 79 infants in the cohort who were discharged from the NICU. NOMAS scores were unavailable for four infants: one early discharge, one transfer, one infant not orally feeding at discharge, and one missing data point. There were no differences in medical factors among those who withdrew and those who remained in the study, among those who had and did not have NOMAS scores, and among those who did and did not return for developmental testing. Data from 75 infants (39 females, 36 males; mean gestational age 26.56wk [SD 1.90], range 23–30wk; mean birthweight 967.33g, [SD 288.54g], range 480–2240g; $n=73$ bottle-fed, $n=2$ breastfed during the NOMAS evaluation) were included in analysis. Of these, 44 (59%) were categorized as disorganized feeders and 31 as dysfunctional feeders, with none categorized as normal feeders. Infants categorized as dysfunctional had longer periods of mechanical ventilation [$t=2.60$, mean difference 5.11 (CI=2.11 to 22.67); $p=0.02$]. No other medical variables were significantly associated with NOMAS scores (see Table I).

Term evaluation

Dysfunctional NOMAS scores were associated with increased signs of stress on the NNNS [$t=2.61$, mean difference 0.073 (95% CI 0.017 to 0.129); $p=0.011$] and decreased neurological function on the Dubowitz [$t=-2.14$; mean difference -2.32 , (95% CI -0.157 to -4.49); $p=0.036$]. No other associations were found between NOMAS category and neurobehavior at term (see) Table II.

Cerebral injury was present in 19.4% ($n=14$) of the cohort that survived until NICU discharge and had NOMAS scores. No significant associations between presence of cerebral injury and NOMAS scores were observed. See Table III for brain metrics and NOMAS categorization. Dysfunctional NOMAS scores were associated with decreased transcerebellar diameter [$t=-2.22$, mean difference -2.04 , (CI= -3.89 to -0.203); $p=0.03$]. No other MRI/DTI measures were associated with NOMAS scores.

Predictive validity

Following discharge, an additional two infants died and one withdrew from the study, leaving 72 infants in the cohort. Fifty-nine (82%) returned for developmental testing at 2 years (see Table III for mean BSID-III scores by NOMAS category). Seventeen percent ($n=10$) of the cohort had developmental delay at age 2 years. There were no associations between NOMAS scores and BSID-III cognitive, language and motor outcome. There was no association between NOMAS scores and developmental delay, using adjusted and unadjusted scores on the BSID-III.

Seventeen percent ($n=4$) of infants categorized as dysfunctional had developmental delay at age 2 years. The confidence interval of this proportion is 0.01–0.33. The proportions between the two cohorts (100% from the 1999 cohort of $n=6$ ¹³ and 17% in the current cohort) were significantly different (Fisher's exact test $p=0.017$).

Reliability

Interrater reliability ranged between $r = -0.43$ and 0.62 . Intrarater reliability ranged from $r = 0.33$ to 1.00 (see Table IV).

DISCUSSION

Key findings include that predictive validity of the NOMAS was not supported, contradicting previous reports of 100% predictive validity for infants with dysfunctional feeding.¹⁴ However, the NOMAS does have some concurrent validity through associations with measures of neurobehavior and brain structure at term. Interrater and intrarater reliability were variable and suboptimal.

Associations between NOMAS scores and days of mechanical ventilation were found. Adverse oral stimuli, such as prolonged mechanical ventilation are suspected to relate to long-term feeding outcomes.^{8,22} This was supported in the current study, manifesting in poorer NOMAS scores. However, no other perinatal exposures were associated with NOMAS outcome.

Some concurrent validity of the NOMAS was established through associations with measures of early neurobehavior. Infants categorized as dysfunctional demonstrated lower Dubowitz scores, which incorporates motor-based items, including assessment of suck and tone.¹⁵ As efficient oral feeding requires efficient oral-motor control and skills,^{23,24} associations with the NOMAS provide some rationale that they are testing similar domains of function.

Infants categorized as dysfunctional also demonstrated more signs of stress on the NNNS. Stress signs may include an inability to achieve an awake state, startles/tremors, cyanosis, and other disruptive behaviors. Stress signs may interrupt the feeding process and manifest in poor performance on the NOMAS. No other study to date has investigated concurrent associations of the NOMAS with neurobehavior; therefore, these associations are new contributions to the literature and merit further investigation.

The NOMAS was associated with one variable in cerebral structural analysis. Feeders in the dysfunctional category demonstrated decreased transcerebellar diameter. The cerebellum is associated with fine and gross motor coordination, and in infancy plays a vital role in the coordination of reflexive feeding movements.^{23–25} Studies have established that the cerebellum is particularly vulnerable in preterm infants.²⁵ Cerebellar volume in preterm infants correlates with cerebral injury and volume loss, later head circumference measurements, and overall growth.²⁵ This may be related to later feeding difficulties, which have been linked with neonatal cerebral injury and long-term neurological conditions.⁵ No other study has investigated associations between NOMAS scores and concurrent brain development. The results of this study demonstrate that although NOMAS scores are not associated with many types of cerebral injury, the cerebellum can have important associations with feeding performance.

Predictive validity of the NOMAS was not supported because of lack of associations with measures of functional outcome at age 2 years. Less than 20% of dysfunctional feeders demonstrated developmental delay. This contradicts earlier research establishing 100% predictive validity of the dysfunctional category.¹⁴ Feeding problems have been associated with functional impairment;^{5,8,24} however, we have demonstrated low predictive validity of this tool in a prospective study with high-risk preterm infants.

Reliability values were variable and substandard. NOMAS-certified raters in this study failed to demonstrate adequate interrater reliability. This contradicts research describing moderate to substantial interrater reliability of the NOMAS.¹³ Some values for interrater reliability were negative. These are interpreted as equivalent to zero, with decreased likelihood that the results are due to chance. Failure to achieve adequate interrater and intrarater reliability jeopardizes the clinical utility of the tool. The rater who determined NOMAS category for the entire cohort achieved an intrarater κ of 1.00, and had 18 years of experience in the NICU. The other rater with the highest agreement ($\kappa = 0.62$) was another therapist with significant feeding experience in the NICU. Those with less experience in neonatal feeding, including students and clinicians not primarily working with neonates, are represented in the lower range of κ values. This may indicate that the course itself is not sufficient to prepare raters to use the NOMAS reliably. Clinical experience, in conjunction with NOMAS training, may enhance the utility of the tool and the ability of clinicians to assess infants appropriately. This merits further investigation to fully discern factors influencing reliability of the NOMAS.

This study investigated a diverse sample, i.e. those influenced post-discharge by various factors, including young maternal age, family structure, socio-economic status, and therapeutic interventions. All of these factors may influence outcome measured at follow-up. The NOMAS cannot account for these factors, which may impede its ability to predict later function. This is the first study to investigate psychometrics of the NOMAS prospectively with this population. Previous studies included small sample sizes ($n \leq 35$), were retrospective and/or confounded.^{1,13,14} This study included a larger sample size and compared the NOMAS to medical factors in the NICU, neurobehavior, brain structure, and developmental outcomes at corrected age 2 years. Comparing the NOMAS to these factors is both a novel addition to the literature and critical to understanding the utility of the NOMAS.

The NOMAS is considered one of the best available tools for the evaluation of neonatal feeding. Some changes that may improve the utility of the NOMAS include establishing objective scoring criteria in a manual, which is not currently available, and changes to the certification, emphasizing clinical experience beyond the course. The NOMAS has not been recently updated, and modifications may improve its clinical utility.

There were some limitations to this study. Though this study was conducted using a larger and more diverse sample than previous investigations, it remains lower than ideal for a psychometric investigation. However, adequate power was achieved to state that the NOMAS dysfunctional category failed to predict developmental delay 100% of the time, as previously reported.¹⁴ Interrater and intrarater reliability were determined using a convenience sample of raters. Although certified to score the NOMAS, three raters were occupational therapy students and had limited NICU clinical experience. However, the upper range of interrater and intrarater reliability is represented in the statistics and remains suboptimal. Finally, none of the infants in the cohort were categorized as normal on the NOMAS. All participants were preterm infants born at 30 weeks' or less gestation with a high risk of feeding impairment and developmental problems, which lends credibility to the test as it discerned abnormalities in all infants, as expected. As many factors were examined at term, an increased likelihood existed of multiple findings.

Future directions for research include further and more detailed evaluations of NOMAS items (versus categories) and investigation into cerebral structural differences that were found in this study. Further investigation into developmental differences between NOMAS categories observed in this cohort is warranted. The BSID-III is the best available tool for measuring developmental function. It was expected that any existing differences between

NOMAS groups would be detected at follow-up. More research is needed to analyze why differences were not found.

The NOMAS appears to have some clinical utility, but modifications to the tool and/or certification may better prepare evaluators to administer and score the assessment appropriately. Currently, the NOMAS lacks adequate evidence to support the claim of predicting future function. With very few structured assessments available to use in the NICU, an enhanced tool may improve early identification of feeding problems to enable services to optimize outcome.

Acknowledgments

We wish to acknowledge the contributions of Ryan Colvin and JiaJing Chen for assistance with biostatistics. We also thank Tara Crapnell, Lisa Tiltges, Laura Madlinger, Chrysty Sturdivant and Cheryl Klohr for assisting with the study. Finally, we thank the infants and their families whose participation makes this research possible. This project was supported by the National Institute of Health (ROI HD 057098), the Doris Duke Charitable Foundation, and the Washington University Intellectual and Developmental Disabilities Research Center (NIH/NICHD P30 HD062171).

Abbreviations

BBPD	Bone-biparietal diameter
BPD	Biparietal diameter
BSID-III	Bayley Scales of Infant Development, third edition
DTI	Diffusion tensor imaging
IHD	Interhemispheric distance
NICU	Neonatal Intensive Care Unit
NNNS	NICU Network Neurobehavioral Scale
NOMAS	The Neonatal Oral Motor Assessment Scale

REFERENCES

1. Howe TH, Lin KC, Fu CP, Su CT, Hsieh CL. A review of psychometric properties of feeding assessment tools used in neonates. *J Obstet Gynecol Neonatal Nurs*. 2008; 37:338–349.
2. Howe TH, Sheu CF, Hsieh YW, Hsieh CL. Psychometric characteristics of the Neonatal Oral-Motor Assessment Scale in healthy preterm infants. *Dev Med Child Neurol*. 2007; 49:915–919. [PubMed: 18039238]
3. Committee on Fetus and Newborn. Hospital discharge of the high-risk neonate. *Pediatrics*. 2008; 122:1119–1126. [PubMed: 18977994]
4. Hawdon JM, Beauregard N, Slattery J, Kennedy G. Identification of neonates at risk of developing feeding problems in infancy. *Dev Med Child Neurol*. 2000; 42:235–239. [PubMed: 10795561]
5. Lefton-Greif MA. Pediatric dysphagia. *Phys Med Rehabil Clin N Am*. 2008; 19:837–851. ix. [PubMed: 18940644]
6. 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]
7. da Costa SP, van den Engel-Hoek L, Bos AF. Sucking and swallowing in infants and diagnostic tools. *J Perinatol*. 2008; 28:247–257. [PubMed: 18200022]
8. 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–e22. [PubMed: 19832883]

9. Tsai SW, Chen CH, Lin MC. Prediction for developmental delay on Neonatal Oral Motor Assessment Scale in preterm infants without brain lesion. *Pediatr Int.* 2010; 52:65–68. [PubMed: 19400913]
10. Fucile S, Gisel EG, McFarland DH, Lau C. Oral and non-oral sensorimotor interventions enhance oral feeding performance in preterm infants. *Dev Med Child Neurol.* 2011; 53:829–835. [PubMed: 21707601]
11. Braun MA, Palmer MM. A pilot study of oral-motor dysfunction in "at-risk" infants. *Phys Occup Ther Pediatr.* 1985; 5:13–26.
12. da Costa SP, van der Schans CP. The reliability of the Neonatal Oral-Motor Assessment Scale. *Acta Paediatr.* 2008; 97:21–26. [PubMed: 18201309]
13. Palmer MM, Crawley K, Blanco IA. Neonatal Oral-Motor Assessment scale: a reliability study. *J Perinatol.* 1993; 13:28–35. [PubMed: 8445444]
14. Palmer MM, Heyman MB. Developmental outcome for neonates with dysfunctional and disorganized sucking patterns: Preliminary findings. *Infant–Toddler Intervention: The Transdisciplinary Journal.* 1999; 9:299–308.
15. Dubowitz, LMS.; Dubowitz, V.; Mercuri, E. *The Neurological Assessment of the Preterm and Full Term Newborn Infant.* London: Mac Keith Press; 1999.
16. Lester, BM.; Tronick, EZ. *The NICU network neurobehavioral scale manual.* Baltimore, MD: Paul H. Brookes Publishing Co.; 2004.
17. Woodward LJ, Mogridge N, Wells SW, Inder TE. Can neurobehavioral examination predict the presence of cerebral injury in the very low birth weight infant? *J Dev Behav Pediatr.* 2004; 25:326–334. [PubMed: 15502549]
18. Mathur AM, Neil JJ, Inder TE. Understanding brain injury and neurodevelopmental disabilities in the preterm infant: the evolving role of advanced magnetic resonance imaging. *Semin Perinatol.* 2010; 34:57–66. [PubMed: 20109973]
19. Network TIN. The CRIB (clinical risk index for babies) score: A tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. *Lancet.* 1993; 342:193–198. [PubMed: 8100927]
20. Bayley, N. *Bayley scales of infant and toddler development.* San Antonio, TX: Harcourt Assessment; 2006.
21. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW. Underestimation of developmental delay by the new Bayley-III Scale. *Arch Pediatr Adolesc Med.* 2010; 164:352–356. [PubMed: 20368488]
22. Dodrill P, McMahon S, Ward E, Weir K, Donovan T, Riddle B. Long-term oral sensitivity and feeding skills of low-risk pre-term infants. *Early Hum Dev.* 2004; 76:23–37. [PubMed: 14729160]
23. Amaizu N, Shulman R, Schanler R, Lau C. Maturation of oral feeding skills in preterm infants. *Acta Paediatr.* 2008; 97:61–67. [PubMed: 18052999]
24. Ross ES, Browne JV. Developmental progression of feeding skills: an approach to supporting feeding in preterm infants. *Semin Neonatol.* 2002; 7:469–775. [PubMed: 12614599]
25. Volpe JJ. Cerebellum of the premature infant: rapidly developing, vulnerable, clinically important. *J Child Neurol.* 2009; 24:1085–1104. [PubMed: 19745085]

What this paper adds

- Dysfunctional NOMAS scores were associated with lower Dubowitz scores, higher stress scores on the NNNS, and decreased trans cerebellar diameter.
- NOMAS predictive validity was not supported, contradicting previous reports.
- Interrater and intrarater reliability of the NOMAS is variable and suboptimal.

Table I

Description of the sample

Characteristic	<i>NOMAS Category</i>		Mean Difference	<i>p</i> ^a
	Disorganized (<i>n</i> =44) mean (SD) or <i>n</i> (%)	Dysfunctional (<i>n</i> =31) mean (SD) or <i>n</i> (%)		
Gestational age at birth (wk)	27.14 (1.76)	26.06 (2.70)	1.07	0.06
Birthweight, g	997.61 (240.46)	923.87 (345.24)	73.74	0.28
Females/Males	24/20	15/16	–	0.60
Total parenteral nutrition (d)	20.73 (15.56)	31.35 (28.04)	–10.63	0.06
CPAP (d)	6.57 (8.70)	7.42 (11.27)	–0.85	0.71
Days on mechanical ventilation	8.55 (16.41)	20.94 (24.91)	–12.39	0.02 ^c
Critical Risk Index for Babies	3.32 (3.32)	4.47 (3.87)	–1.15	0.18
Cerebral injury ^b	8 (18.2%)	6 (19.4%)	–	0.92

^a Investigation of differences across infants with disorganized and dysfunctional feeding on the Neonatal Oral Motor Assessment Scale (NOMAS) using independent samples *t*-tests, Mann–Whitney *U* tests, and χ^2 analyses.

^b Cerebral injury was defined as presence of grade III–IV intraventricular hemorrhage, cystic periventricular leukomalacia, and/or cerebellar hemorrhage. CPAP, continuous positive airway pressure. Bold type indicates significance.

Table II

Associations between NOMAS, neurobehavior, and brain structure

Variable	NOMAS Disorganized (<i>n</i> =44)	NOMAS Dysfunctional (<i>n</i> =31) mean (SD)	<i>p</i> ^{<i>b</i>}
Dubowitz Optimality Score	19.77 (4.48)	17.45 (4.84)	0.04
NNNS			
Habituation ^{<i>a</i>}	8.17 (6.50, 9.00)	7.50 (7.00, 9.00)	0.12
Orientation	3.32 (1.15)	3.2 (1.43)	0.73
Handling ^{<i>a</i>}	0.63 (0.63, 0.75)	0.75 (0.63, 0.81)	0.11
Quality of Movement	3.60 (0.72)	3.39 (0.84)	0.23
Self Regulation	4.51 (0.82)	4.20 (0.69)	0.09
Suboptimal Reflexes	6.48 (2.09)	7.42 (2.46)	0.08
Stress	0.33 (0.12)	0.41 (0.13)	0.01
Arousal	4.13 (0.86)	4.26 (0.95)	0.55
Hypertonia ^{<i>a</i>}	1.00 (1.00, 2.00)	2.00 (1.00, 3.00)	0.25
Hypotonia ^{<i>a</i>}	1.00 (0.00, 1.00)	1.00 (0.00, 1.00)	0.42
Asymmetry ^{<i>a</i>}	2.00 (0.00, 4.00)	3.00 (1.00, 4.00)	0.21
Excitability	5.05 (2.57)	6.10 (2.51)	0.08
Lethargy	6.77 (2.67)	7.16 (3.15)	0.57
MRI			
Bifrontal diameter	60.00 (6.15)	58.26 (6.26)	0.26
Biparietal diameter	70.43 (4.90)	69.08 (6.00)	0.31
Interhemispheric distance	3.15 (1.57)	3.39 (1.43)	0.51
Transcerebellar diameter	48.59 (3.37)	46.55 (4.27)	0.03
Right ventricle: Brain size ratio	8.53 (1.99)	7.88 (1.33)	0.20
Left ventricle: Brain size ratio	8.71 (1.96)	7.84 (1.61)	0.09

^{*a*}Reported as median (25%, 75%) due to lack of normality.^{*b*}Investigation of differences across infants with disorganized and dysfunctional feeding on the Neonatal Oral Motor Assessment Scale (NOMAS) using independent samples *t*-tests, Mann-Whitney *U* tests, and linear regression. Post-menstrual age at time of MRI was controlled for when investigating differences in brain measures; bold type indicates significance. NNNs, NICU Network Neurobehavioral Scale.

Table III

Predictive validity: associations between NOMAS and 2-year outcome

<i>NOMAS categories</i>				
Domain	Disorganized (n=35) mean (SD) or n (%)	Dysfunctional (n=24) mean (SD) or n (%)	Mean difference or odds ratio	<i>p</i>^a
BSID-III				
Cognitive	84.69 (10.19)	87.71 (10.32)	-3.02	0.27
Language	91.62 (12.16)	86.96 (13.25)	4.66	0.17
Motor	86.29 (11.28)	84.42 (10.48)	1.87	0.52
DD	6 (17.1%)	4 (16.7%)	OR=0.93 ^b	0.62

^aInvestigation of differences across infants with disorganized and dysfunctional feeding on the Neonatal Oral Motor Assessment Scale (NOMAS) using independent samples *t*-tests and ² analysis.

^bReference group is dysfunctional group. BSID-III, Bayley Scales of Infant Development, third edition; DD, developmental delay (counted as yes if <70 on any composite score on the BSID-III).

Table IV

Kappa scores for interrater and intrarater reliability

	Rater 1 ^a	Rater 2	Rater 3	Rater 4	Rater 5 ^a	Rater 6 ^a
Rater 1 ^a		-0.11	-0.43	0.29	0.17	0.62
Rater 2			0.38	0.06	-0.39	-0.18
Rater 3				-0.25	-0.23	-0.36
Rater 4					0.29	-0.36
Rater 5 ^a						-0.15
Intrarater	1.00	0.38	1.00	N/A ^b	0.55	0.33
Year certified	2008	2011	2011	2011	2011	1999
Years of infant experience	18	0	0	0	10	30

^aRater is a licensed and registered occupational therapist (OTR/L). Raters 2, 3, and 4 are occupational therapy students, but were certified to use the Neonatal Oral Motor Assessment Scale (NOMAS).

^bKappa not calculable because of lack of variation during one visit. This rater demonstrated self-agreement in all except one video. Reliability was calculated using five video recordings of NOMAS evaluation.