The American Journal on Addictions, 25: 370–373, 2016

ISSN: 1055-0496 print / 1521-0391 online
DOI: 10.1111/ajad.12388

Brief Report: Psychometric Assessment of the Neonatal Abstinence Scoring System and the MOTHER NAS Scale

Hendrée E. Jones, PhD,[1,2] Carl Seashore, MD,[3] Elisabeth Johnson, FNP, PhD,[1] Evette Horton, PhD,[1] Kevin E. O'Grady, PhD,[4] Kim Andringa, PhD,[1] Matthew R. Grossman, MD,[5] Bonny Whalen, MD,[6] Alison Volpe Holmes, MD, MPH[6]

[1]UNC Horizons and Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
[2]Departments of Psychiatry and Obstetrics and Gynecology, School of Medicine, Johns Hopkins University, Baltimore, Maryland
[3]Department of Pediatrics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
[4]Department of Psychology, University of Maryland, College Park, College Park, Maryland
[5]Department of Pediatrics, Yale School of Medicine, New Haven, Connecticut
[6]Department of Pediatrics, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire

Background and Objectives: The present study examined the psychometric characteristics of the Neonatal Abstinence Scoring System (NASS; "Finnegan Scale") and the MOTHER NAS Scale (MNS).
Methods: Secondary analysis of data from 131 neonates from the Maternal Opioid Treatment: Human Experimental Research (MOTHER) study, a randomized trial in opioid-dependent pregnant women administered buprenorphine or methadone.
Results: Both the NASS and MNS demonstrated poor psychometric properties, with internal consistency (Cronbach's αs) failing to exceed .62 at first administration, peak NAS score, and NAS treatment initiation.
Conclusions: Findings support the need for development of a NAS measure based on sound psychometric principles.
Scientific Significance: This study found that two frequently used measures of neonatal abstinence syndrome suffer inadequacies in regard to their basic measurement characteristics. (Am J Addict 2016;25:370–373)

INTRODUCTION

Neonatal abstinence syndrome (NAS) is characterized by signs and symptoms indicating central nervous system hyperirritability, and dysfunction of the autonomic nervous system, gastrointestinal tract, and respiratory system,[1] resulting from prenatal exposure to opioids.[2] Untreated NAS can be a source of significant morbidity.[1] Treatment often requires extended hospitalization, pharmacotherapy, and monitoring of the neonate.

Diagnosis and treatment for NAS have been based on any one of numerous instruments, the most notable being the Neonatal Abstinence Scoring System (NASS), often termed the "Finnegan Scale."[3] The scoring systems devised for these instruments has nonetheless been criticized as subjective.

The NASS appears to be the most frequently used instrument to measure NAS, with estimates of its use ranging between 52% and 65% of hospital settings where NAS is treated.[4,5] However, the issue of how best to measure NAS is complicated by the wide variety of modifications of the NASS that are in use—local variants of the measure are used more frequently in both research and clinical practice than is the original NASS.[4,5] The MOTHER NAS Scale (MNS)[2], a new instrument developed to measure NAS and based in part on the NASS, represents a potential improvement to the NASS.

The present study answers two specific questions: What are the psychometric properties of the NASS and the MNS?; and is one measure superior in terms of its psychometric properties?

METHODS

MOTHER Study

The MOTHER (Maternal Opioid Treatment: Human Experimental Research) study[2] was a double-blind, double-dummy, flexible-dosing, two-group trial in opioid-dependent pregnant women randomly assigned to either methadone or buprenorphine pharmacotherapy in the context of comprehensive prenatal care and substance use treatment.

MOTHER had exclusionary criteria that required participants to have neither current DSM-IV abuse of nor dependence on either alcohol or benzodiazepines. Moreover,

Received February 24, 2016; revised May 4, 2016; accepted May 14, 2016.
Address correspondence to Jones, UNC Horizons, 127 Kingston Drive, Chapel Hill, NC 27514.
E-mail: Hendree_Jones@med.unc.edu

MOTHER utilized a contingency management approach to minimize use of illicit substances, and participants were rewarded with monetary vouchers for providing urine samples negative for opioids and other illicit substances excluding their respective study medication for which testing was not conducted. This contingency management program yielded a very low level of illicit substance use during the study: 73% of the urine samples were opioid-negative and 81% were cocaine-negative during prenatal testing, and only 13% of the sample tested positive for opioids at delivery.

The information that follows summarizes aspects of MOTHER central to the present paper.

MOTHER Sites

Seven university hospitals provided outcome data. The number of participants at some US sites was quite small, so the seven sites were combined into three categories (US Urban vs. US Rural vs. European) in order to create a factor that would control for possible site effects.

MOTHER Participants

Recruitment targeted opioid-dependent pregnant women with a singleton fetus between 6 and 30 weeks gestation who met eligibility criteria.[2] One hundred seventy-five women participated, of whom 131 delivered an infant while enrolled in the study. The secondary analyses reported in the present paper use the NAS data from these neonates.

Measurement of Neonatal Abstinence Syndrome

Neonatal Abstinence Scoring System (NASS)

The NASS[3] consists of 31 signs of central nervous, gastrointestinal, or respiratory/vasomotor disturbances that are scored either zero (absence of) or, variously, one through five, indicating presence and/or degree of severity, with a score range of 0–46, inclusive.

MOTHER NAS Scale (MNS)

The MNS[2] includes all NASS items, but uses a revised scoring procedure in an attempt to incorporate a number of modifications and improvements. It does not score 14 NASS items due to their overlap with other items (eg, nasal flaring) or because they are unresponsive to opioid treatment (eg, myoclonic jerking), and adds two items: failure to thrive (to assess neonates who experience excessive weight loss due to their hypertonicity, excessive movements, and/or difficulties feeding) and irritability (to assess neonates who are irritable but do not cry). Thus, MNS total scores are based on 19 weighted item scores, and can range from 0 to 42, inclusive. Jones et al.[2] discuss the degree of inter-rater agreement.

Raters were trained in administration and scoring of the NASS (and hence 17/19 of the MNS items) as well as the two additional MNS items.[2]

The MNS was administered every 3–4 hours during neonatal inpatient hospitalization for a minimum of 10 days by trained staff. NAS treatment utilized a protocol based on MNS scores.[2]

Statistical Analyses

Scores on all NASS and MNS items for each participant were recorded in the study database. Thus, obtaining NASS scores after completion of the MOTHER study[2] was possible.

Because the primary focus was on the psychometric properties of the two measures, corrected item-total correlations and internal consistency reliabilities (Cronbach's α)[6,7] were calculated for each measure, controlling for pharmacotherapy condition, site, and neonatal gender. Cronbach's α represents how closely related a set of items are as a group, and is often viewed as an index of "accuracy" of measurement of a measure. The difference between two Cronbach's αs was evaluated with a test of dependent Cronbach's αs.[8]

A critical issue in determining the psychometric properties of the NASS and MNS was the time at which NAS was assessed. To address this issue systematically, corrected item-total correlations and Cronbach's α were estimated at three administration time points: (1) at first MNS administration; (2) at peak MNS score; and (3) at time of medication treatment initiation in the neonates treated for NAS. Estimates at first administration of the MNS and at peak MNS score might represent underestimates, given possible floor and ceiling effects, respectively. To maximize the range and, hence, the variability of NASS and MNS scores, neonates not diagnosed with NAS were randomly matched within medication condition by the day and time of assessment with neonates diagnosed with NAS at treatment initiation, thus maximizing the possible range of NASS and MNS scores in the total sample.

RESULTS

Table 1 contains the corrected item-total correlations and the internal consistency reliability (Cronbach's αs) for the NASS and MNS at the three administration time points. At first administration, the NASS had 6/31 items with no variance, 5/31 items significantly positively, and 1/31 significantly negatively correlated ($ps < .05$) with the total NASS score, while the MNS had 6/19 items with no variance, and 6/19 items significantly positively correlated with the total MNS score. At peak NAS score, the NASS had 1/31 items with no variance, 10/31 items were significantly positively correlated, and 4/31 items were significantly negatively correlated with the total NASS score, while the MNS had 11/19 items significantly positively and 1/19 items significantly negatively correlated with the total MNS score. At treatment initiation, the NASS had 2/31 items with no variance, 9/31 items that were significantly positively correlated, and 1/31 items significantly negatively correlated with the total NASS score, while the MNS had 1/19 items with no variance and 11/19 items significantly positively correlated with the total MNS score.

The Cronbach's αs were not significantly different between the NASS and the MNS at all three administration time points ($ps > .1$).

TABLE 1. Corrected item total correlations and internal consistency αs for the Neonatal Abstinence Scoring System (NASS) and MOTHER NAS Scale (MNS) at first assessment, Peak NAS score, and treatment initiation ($N = 131$)

	First assessment		Peak NAS score		Treatment initiation	
	NASS	MNS	NASS	MNS	NASS	MNS
Excessive high pitched crying	.17		−.08		**.19**	
Continuous high pitched crying	0		.13		.16	
Crying		**.21**		**.30**		**.42**
Sleeps <1 hour after feeding	−.03		.13		.12	
Sleeps <2 hour after feeding	−.09		**−.28**		.03	
Sleeps <3 hour after feeding	−.01		**−.23**		−.11	
Sleeping		−.11		**.24**		**.29**
Hyperactive moro reflex	**.23**		−.15		.15	
Markedly hyperactive moro reflex	.12		**−.29**		.11	
Moro reflex		**.30**		**.19**		**.25**
Mild tremors: disturbed	**−.29**		**−.25**		**−.29**	
Mod-severe tremors: disturbed	**.35**		**.27**		**.38**	
Disturbed tremors		**.34**		**.27**		**.34**
Mild tremors: undisturbed	.06		.07		−.01	
Mod-severe tremors: undisturbed	**.28**		.17		**.20**	
Undisturbed tremors		**.46**		**.34**		**.32**
Increased muscle tone	**.31**	**.40**	.13	**.25**	**.25**	**.36**
Excoriation	.08	0	.12	.17	.06	.14
Myoclonic jerk	−.12		.12		−.03	
Generalized convulsions	0	0	−.16	**−.20**	0	0
Sweating	0	0	**.29**	**.35**	.18	.16
Fever 37.2–38.3 °C	.04		**.30**		.13	
Fever ≥38.4 °C	**.21**		0		−.01	
Fever ≥37.2 °C		.07		**.36**		.14
Frequent yawning (>3)	0	0	−.13	−.10	.02	.07
Mottling	.08		.08		.04	
Nasal stuffiness	−.05	.06	**.33**	**.29**	**.19**	**.25**
Sneezing (>3)	.09	.04	.14	.05	.09	.03
Nasal flaring	.15		**.27**		0	
Respiratory rate (>60/min)	.12	.02	.20	**.20**	**.23**	**.22**
Respiratory rate (>60/min with retractions)	.15		.13		−.03	
Excessive sucking	.04		**.32**		**.45**	
Poor feeding	.14	.02	.09	.08	**.24**	**.28**
Regurgitation	−.02		.08		.02	
Projectile vomiting	−.04		**.36**		.05	
Vomiting		−.07		.02		.01
Loose stools	0	0	**.27**	.07	**.23**	.17
Watery stools	0		**.19**		.18	
Failure to thrive		0		−.01		**.21**
Excessive irritability		**.31**		**.50**		**.47**
Internal consistency αs	.38	.43	.36	.48	.45	.62

NAS, Neonatal Abstinence Syndrome; NASS, Neonatal Abstinence Scoring System; MNS, MOTHER NAS Scale; α, Cronbach's α measure of internal consistency.

Corrected item-total correlations are corrected for the fact that the total would also contain the item itself. Corrected item-total correlations are estimated controlling for site, pharmacotherapy condition, and neonatal gender (ie, site, pharmacotherapy condition, and neonatal gender have been partialled out of the correlation between each item and its total score). Corrected item-total correlations that are zero are due to the fact that the item in question had no variance. Significant ($p < .05$) corrected item-total correlations are in bold.

DISCUSSION

Findings from the present study clearly indicate that neither NAS instrument can be said to have acceptable psychometric properties. These findings are hardly surprising, for three reasons. First, both instruments required considerable clinical interpretation of neonatal behavior, even with extensive training of raters. Second, both instruments were developed using what has been termed a "rational approach" to scale construction—that is, items are developed based on a fund of knowledge of the researcher(s), rather than following an empirical procedure for scale development. Third, item scoring is likewise based on a rational weighting scheme that was not empirically tested. As a result, some subset(s) of items suffer from any one or more drawbacks, including ambiguity, floor or ceiling effects, or lack of discrimination. As such, rationally derived measures often suffer in terms of reliability—and hence, validity. Finally, for both measures, the problem with ambiguity is magnified by a number of items that prove difficult for raters to make discriminations between degree or amount of symptom expression (eg, undisturbed vs. disturbed tremors).

Critical review of both measures indicates that they assess many signs that do not necessarily relate to the clinical course or severity of NAS in newborns (eg, sneezing, yawning, tremors) yet neglect or underemphasize more clinically relevant symptoms such as excessive weight loss or poor weight gain, feeding difficulty, difficulty sleeping, and inability to be consoled.[9,10]

The present study is limited by its use of secondary analyses of data collected as part of a study whose primary aims did not focus on a psychometric examination of either the NASS or the MNS. A study whose focus was examination of the psychometric characteristics of the two measures could have reached different conclusions. Moreover, use of the MNS rather than the NASS to determine NAS treatment might have biased the findings regarding reliability of either measure in some unknown way.

CONCLUSIONS

Current measures in use do not accurately reflect the clinical condition of the newborn with NAS with an acceptable degree of reliability. Findings support the need for the development of a measure of NAS based on sound psychometric principles.

All MOTHER grants are from the National Institute on Drug Abuse (NIDA) unless noted otherwise: Brown University, R01 DA015778; Johns Hopkins University, R01 DA015764; Medical University of Vienna, R01 DA018417; Thomas Jefferson University, R01 DA015738; University of Toronto, R01 DA015741; University of Vermont, R01 DA 018410 and M01 RR109; Vanderbilt University, R01 DA 017513 and M01 RR00095, and Wayne State University, R01 DA15832. Reckitt Benckiser Healthcare, Hull, UK supplied buprenorphine tablets (and the associated placebo) via the National Institute on Drug Abuse (NIDA).

Neither NIDA nor Reckitt Benckiser had any role in the design and conduct of the study; data acquisition, management, analysis, and interpretation; or preparation, review, and approval of the manuscript.

Declaration of Interest

HEJ received reimbursement for her time and travel from Reckitt Benckiser during the conduct of the MOTHER study. KEO'G received reimbursement for his time from Reckitt Benckiser in the period following the conduct of the MOTHER study. Neither has received reimbursement since 2011. All other authors declare no conflicts of interest. The authors alone are responsible for the content and writing of this paper.

REFERENCES

1. Finnegan LP, Kaltenbach K. Neonatal abstinence syndrome. In: Hoekelman RA, Friedman SB, Nelson NM, Seidel HM, eds. *Primary Pediatric Care*, 2nd edn. St. Louis: Mosby; 1992. pp. 1367–1378.
2. Jones HE, Kaltenbach K, Heil SH, et al. Neonatal abstinence syndrome after methadone or buprenorphine exposure. *N Engl J Med*. 2010; 363: 2320–2331.
3. Finnegan LP, Connaughton JF Jr., Kron RE, et al. Neonatal abstinence syndrome: Assessment and management. *Addict Dis*. 1975;2:141–158.
4. Sarkar S, Donn SM. Management of neonatal abstinence syndrome in neonatal intensive care units: A national survey. *J Perinatol*. 2006; 26:15–17.
5. O'Grady MJ, Hopewell J, White MJ. Management of neonatal abstinence syndrome: A national survey and review of practice. *Arch Dis Child Fetal Neonatal Ed*. 2009;94:F249–F252.
6. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334.
7. Develles RF. *Scale development*. Thousand Oaks, CA: Sage Publications; 1991.
8. Alsawalmeh YM, Feldt LS. A modification of Feldt's test of the equality of two dependent alpha coefficients. *Psychometrika*. 1994; 59:49–57.
9. Whalen BL, Holmes AV. Neonatal abstinence syndrome and the pediatric hospitalist. *Hosp Pediatr*. 2013;3:324–325.
10. Holmes AV, Atwood EC, Whalen B, et al. Rooming-in to treat neonatal abstinence syndrome: Improved family-centered care at lower cost. *Pediatrics*. 2016. pii: e20152929.