Simplification of the Finnegan Neonatal Abstinence Scoring System: retrospective study of two institutions in the USA

Enrique Gomez Pomar,1 Loretta P Finnegan,2 Lori Devlin,3 Henrietta Bada,1 Vanessa A Concina,1 Katrina T Ibonia,1 Philip M Westgate4

ABSTRACT

Objective To develop a simplified Finnegan Neonatal Abstinence Scoring System (sFNAS) that will highly correlate with scores ≥8 and ≥12 in infants being assessed with the FNAS.

Design, setting and participants This is a retrospective analysis involving 367 patients admitted to two level IV neonatal intensive care units with a total of 40,294 observations. Inclusion criteria included neonates with gestational age ≥37/0/7 weeks, who are being assessed for neonatal abstinence syndrome (NAS) using the FNAS. Infants with a gestational age <37 weeks were excluded.

Methods A linear regression model based on the original FNAS data from one institution was developed to determine optimal values for each item in the sFNAS. A backward elimination approach was used, removing the items that contributed least to the Pearson’s correlation. The sFNAS was then cross-validated with data from a second institution.

Results Pearson’s correlation between the proposed sFNAS and the FNAS was 0.914. The optimal treatment cut-off values for the sFNAS were 6 and 10 to predict FNAS scores ≥8 and ≥12, respectively. The sensitivity and specificity of these cut-off values to detect FNAS scores ≥8 and ≥12 were 0.888 and 0.883 for a cut-off of 6, and 0.637 and 0.992 for a cut-off of 10, respectively. The sFNAS cross-validation resulted in a Pearson’s correlation of 0.908, sensitivity and specificity of 0.860 and 0.873 for a cut-off of 6, and 0.525 and 0.986 for a cut-off of 10, respectively.

Conclusion The sFNAS has a high statistical correlation with the FNAS, and it is cross-validated for the assessment of infants with NAS. It has excellent specificity and negative predictive value for identifying infants with FNAS scores ≥8 and ≥12.

INTRODUCTION

The incidence of neonatal abstinence syndrome (NAS) has steadily increased since the 1970s and is now a significant public health problem.1–4 Tolia et al5 reported an almost fourfold increase in NAS neonatal intensive care unit (NICU) admissions from 7 cases/1000 admissions in 2004 to 27 cases/1000 in 2013. The median length of stay for infants with NAS increased from 13 to 19 days. The proportion of infants who received pharmacotherapy also increased, 74% in 2004–2005 to 87% in 2012–2013, resulting in a 35% increase in hospital costs.2

Several scoring systems have been proposed to evaluate infants with NAS.6 These included the Finnegan Neonatal Abstinence Scoring System (FNAS), the Lipsitz tool,7 the Neonatal Withdrawal Inventory,8 the Neonatal Narcotic Withdrawal Index9 and so on. However there is no consensus as to which tool to use, the cut-off points for treatment and the interval between assessments.10–11

The FNAS has been the most commonly used scale for the last 40 years.3 4 10–12 The FNAS was developed in 197511 and consisted of 21 items allowing a thorough assessment of infants with NAS. The tool was analysed for interuser reliability (mean inter-rater reliability coefficient of 0.82 (0.75–0.96)) and validated for the diagnosis of NAS. However, the lack of subsequent validation and inter-rater reliability is a major concern regarding the FNAS.6 After its original publication, the FNAS was slightly modified in its form but not in content.14 15

Strengths and limitations of this study

- A simplified scoring system (simplified Finnegan Neonatal Abstinence Scoring System (sFNAS)) is proposed developed from a large data set and cross-validated using a database from another institution.
- Cutoff values are provided for the sFNAS which predict the cutoff value.
- The retrospective nature of our study requires prospective validation of the inter-rater reliability of the sFNAS and also the clinical validity of this tool.
Finnegan et al have proposed using three consecutive scores of 8 or higher, or two consecutive scores of 12 or higher, to initiate pharmacological treatment. Contiguous scores less than 8 are often used as a measure of readiness for weaning pharmacological therapy. The validity of the cut-off point of the FNAS was evaluated in 2010 by Zimmermann et al in term newborns without opiate exposure. They found that 95% of the scores were less than 8, with some variability according to the day of life and time of day (lower at night); this led them to conclude that a value above 8 can be considered pathological. This study also supports the use of consecutive scores to identify infants who require pharmacological intervention.

The FNAS is a lengthy tool and given the continued increase in the number of infants diagnosed with NAS our goal was to create a shortened and simplified version of Finnegan Scoring System (sFNAS) that will highly correlate with the original FNAS for an efficient clinical assessment.

**METHODS**

This retrospective study was conducted using data from two institutions, both level IV NICUs and regional referral centres for the respective area. The study was approved by the Institutional Review Board of the University of Louisville and University of Kentucky. Data were collected from the electronic medical records at each institution. Inclusion criteria included neonates with gestational age ≥37 0/7 weeks, admission to the NICU for withdrawal after any in utero opioid exposure and assessment with the FNAS for signs of NAS. Infants with a gestational age less than 37 weeks and those with exposure to psychoactive substances without opioid exposure were excluded. At each institution, the nurses who performed the FNAS had training with experienced scorers (video, demonstration by trainer and reverse demonstration by trainee) and testing for reliability (trainer and trainee assess and score the same subject). Tests were two-sided at the 0.01 significance level.

In the development of the sFNAS, the original FNAS scores of ≥8 or ≥12 were considered as the pharmacological treatment cut-off values. Therefore, receiver operating characteristic (ROC) curves based on the accuracy of our shortened score to predict scores of ≥8 or ≥12 from the original FNAS were constructed. Optimal cut-off values for our sFNAS were then determined using the maximum proportion correctly classified (PCC). Sensitivities (Sn), specificities (Sp), positive predictive values (PPV) and negative predictive values (NPV) were also obtained, along with corresponding 95% CIs via the use of GEE.

To validate our sFNAS, we applied the scoring to the data from institution 2. Specifically, we computed for the corresponding 95% CIs via the use of GEE.

To validate our sFNAS, we applied the scoring to the data from institution 2. Specifically, we computed for the corresponding 95% CIs via the use of GEE.

**RESULTS**

The data comprised a whole year’s (2014) observations from 185 subjects of institution 1 contributing a total of 27,447 scores, and from 182 babies from institution 2 contributing a total of 12,847 scores. The number of observations per
that were not retained were observed in less than 10% of data from institution 1 (table 1). The majority of items in more than 10% of instances in both institutions. Gastro-respiratory rate >60/min and loose stools were observed. The signs namely tremors, excessive cry, sleep <3 hours, when an infant showed improvement in scores following initiation of pharmacotherapy; the goal was to minimise disturbing the infant to promote more sleep and allow ad lib feedings.

Some signs differed in percentages of occurrence between the two institutions. The sign most frequently noted for both institutions was the increased muscle tone. The signs namely tremors, excessive cry, sleep <3 hours, respiratory rate >60/min and loose stools were observed in more than 10% of instances in both institutions. Gastro-intestinal manifestations were more frequently observed in institution 2.

The backward elimination approach was applied to the data from institution 1 (table 1). The majority of items that were not retained were observed in less than 10% of instances; therefore, they did not contribute enough to the Pearson’s correlation in order to warrant their inclusion into the sFNAS. Motting was an item that was present in more than 10% of the observations; however, its elimination did not greatly affect the Pearson’s correlation (step 11). This was done to simplify the scoring even further, which is particularly useful when it is difficult for scorers to distinguish between different levels of an item. Optimal values for each remaining item were rounded to the nearest integer.

The backward elimination approach was applied to the data from institution 2 resulted in a Pearson’s correlation of 0.908 with the original FNAS. Mottling was an item that was present in more than 10% of the observations; however, its elimination did not greatly affect the Pearson’s correlation (table 1). The proposed sFNAS comprised 10 items as presented in table 4. The Pearson’s correlation of our sFNAS with the original FNAS is 0.914. The evaluation of the interrelatedness of the items that were removed and the items that were kept in the sFNAS is given in table 3. All items that were removed were significantly related to the items that were kept in the sFNAS, as demonstrated in table 3. Motting is the item that was most significantly related to the others, as it had significant associations with 8 out of the 10 items that were retained in the sFNAS. The sFNAS predicted the original FNAS scores ≥8 or ≥12 or higher with a respective area under the ROC curve of 0.952 and 0.982. Based on the PCC, the optimal treatment cut-offs for our shortened scale are 6 and 10, respectively. The PCC for institution 1 for the cut-off values of 6 and 10 was 0.884 and 0.980, respectively, and comparable values of 0.867 and 0.920, respectively, for institution 2. Values of Sn, Sp, PPV, NPV and PCC are given in table 4.

For the statistical validation of the sFNAS, the scoring system applied to the data from institution 2 resulted in a Pearson’s correlation of 0.908 with the original FNAS.
Table 2  Items and scores in the proposed simplified Finnegan Neonatal Abstinence Scoring System

| Item                      | Value | Scores |
|---------------------------|-------|--------|
| Cry                       | Yes (excessive or continuous) | 2 |
| Tremors                   | Tremors undisturbed (mild, moderate or severe) | 5 |
| Tremors                   | Tremors disturbed (mild, moderate or severe) | 1 |
| Increased muscle tone     | Yes   | 2 |
| Sleep                     | <1 hour | 3 |
| Nasal stuffiness          | Yes   | 1 |
| Respiratory rate          | >60/min | 1 |
| Excessive sucking         | Yes   | 1 |
| Poor feeding              | Yes   | 2 |
| Feed tolerance            | Regurgitation or projectile vomiting | 2 |
| Stools                    | Loose or watery stools | 2 |

Table 3  Interrelatedness of items expressed by ORs

| Item                           | OR       |
|--------------------------------|----------|
| Excoriation                    | 1.03     |
| Fever: any                     | 1.07     |
| Frequent yawn                  | 1.08     |
| Moro reflex                    | 1.14     |
| Motting                       | 1.15     |
| Myoclonic jerks                | 1.20     |
| Nasal flaring                  | 1.22     |
| Nasal stuffiness               | 1.25     |
| Respiratory rate: >60/min (retractions or none) | 1.26 |
| Excessive sucking              | 1.32     |
| Poor feeding                   | 1.36     |
| Feeding tolerance: regurgitation or projectile vomiting | 1.37 |
| Loose or watery stools         | 1.40     |

Observed values for Sn, Sp, PPV, NPV and PCC from the use of sFNAS are also presented in table 4.

When deriving a new scale, it is likely that some values that are near the cut-off get incorrectly classified. Therefore, we now look at the misclassified cases, that is, corresponding sFNAS scores that were close to the FNAS cut-offs of 8 and 12. There were 20,197 FNAS scores of 0–7 for the original data and 6,969 for cross-validation data. Of these observations, 9.8% and 9.6%, respectively, corresponded to sFNAS cut-off score of 6. The total FNAS scores from 0 to 11 for the original and cross-validation data sets, respectively, were 26,497 and 10,940. Only 0.6% and 0.9% were misclassified, respectively, corresponding to sFNAS score of 10, the higher cut-off for sFNAS. Of the FNAS scores ≥8 (original data set: 7,250 observations; cross-validation data: 5,784), 8.4% and 9.7%, respectively, were misclassified as sFNAS=5, below the cut-off of 6 for sFNAS. As to FNAS scores ≥12 (original observations: 950; cross-validation data: 1,813), 196 or 20.6% and 397 or 21.9%, respectively, were misclassified as sFNAS=9, lower than the higher sFNAS cut-off of 10.

DISCUSSION

We propose a simplified scoring system, the sFNAS, to evaluate infants with NAS. The sFNAS is a shortened FNAS derived by maximising its correlation with the original scale. Our scoring system consisting of 10 items provides a Pearson’s correlation of 0.914 with the original FNAS and a proposed cut-off of 6 and 10 (instead of 8 and 12). Using the proposed cut-off, the simplified scoring system provides...
excellent Sp and NPV. The cross-validation of sFNAS provided a Pearson’s correlation of 0.908 and detected adequate values of Sn, Sp, PPV and NPV. To our knowledge, this is the first shortened or simplified NAS scoring system that provided a cross-validation process.6

The 21-item original FNAS assesses the central nervous system, autonomic and gastrointestinal signs of infants with NAS. In its development, each of the categories was analysed to include the major clinical signs, resulting in a comprehensive scoring system that included clinically significant items that were often highly correlated, and thereby increasing the number of manifestations to be assessed. Although the sFNAS is shortened to 10 items, it maintains more than one item within each category of signs of NAS.

The assessment of the central nervous system signs in the sFNAS included four items (cry, tremors, tone and sleep) describing neurological excitability. These items were also described or included in other scoring systems describing neurological excitability. These reports indicate variation in the prevalence of the signs of some items between the two institutions in our study. Other autonomic manifestations noted in our infants were frequent yawning, fever, sneezing and sweating; these were rare and/or strongly associated with the manifestations included in the sFNAS.

Of the gastrointestinal manifestations, all four items in the original FNAS were retained in the sFNAS. Prolonged sucking with fewer pauses associated with increased spitting episodes has been reported in infants with NAS, but the significance of these signs remains controversial. Loose or watery stools alone contributed notably to the correlation as an independent and clinically significant sign in NAS.

There were differences in percentages of occurrences of some items between the two institutions in our study. Reports indicate variation in the prevalence of the signs of NAS across different centres. Since external factors and nurses variability have minimal influence on the FNAS, we believe that this variability seems to be more related to intrinsic neonatal factors. Regardless of the difference in percentages of occurrences of some items between our institutions, the Pearson’s correlation was high with adequate Sn and excellent Sp, PPV and NPV. High Sp and NPV are important for identifying infants that should not require pharmacotherapy.

There were other scales developed concomitantly with the FNAS. The Ostrea tool is a six-item scale that ranks multiple NAS signs but with no proposed guide for treatment. The Neonatal Narcotic Withdrawal Index was designed in 1981 with adequate validity, but it did not gain much popularity. The Narcotic Withdrawal Score by Lipsitz in 1975 included 11 items and evaluated infants twice daily. The tool proposed to initiate treatment when one score is greater than 4. However, studies have shown that a high score may not necessarily be obtained in subsequent assessments. Therefore, consecutive scores meeting cut-off would be pertinent to initiate pharmacological treatment.

Table 4 Values obtained after applying the new scoring on the original data set (internal validation) and the data set from institution 2 (cross-validation)

|                  | Sn (95% CI) | Sp (95% CI) | PPV (95% CI) | NPV (95% CI) | PCC (95% CI) | ROC |
|------------------|------------|------------|-------------|-------------|--------------|-----|
| Internal validation process | Cut-off of 6 0.888 (0.874 to 0.903) | 0.883 (0.870 to 0.895) | 0.731 (0.709 to 0.752) | 0.957 (0.950 to 0.963) | 0.884 (0.875 to 0.893) | 0.952 |
|                  | Cut-off of 10 0.637 (0.587 to 0.686) | 0.992 (0.990 to 0.994) | 0.746 (0.694 to 0.798) | 0.987 (0.984 to 0.990) | 0.980 (0.977 to 0.983) | 0.982 |
| Cross-validation process | Cut-off of 6 0.860 (0.845 to 0.876) | 0.873 (0.860 to 0.887) | 0.849 (0.833 to 0.866) | 0.883 (0.867 to 0.899) | 0.867 (0.859 to 0.876) | 0.967 |
|                  | Cut-off of 10 0.525 (0.491 to 0.559) | 0.986 (0.983 to 0.988) | 0.858 (0.833 to 0.882) | 0.926 (0.916 to 0.936) | 0.920 (0.911 to 0.930) | 0.927 |

NPV, negative predictive value; PCC, proportion correctly classified; PPV, positive predictive value; ROC, receiver operating characteristic; Sn, sensitivity; Sp, specificity.
There were previous attempts to modify the Finnegan scoring system. The Neonatal Withdrawal Inventory proposed an 8-point checklist that was derived from the FNAS with reported inter-rater reliability, sensitivity and specificity; however, the scoring system was not validated. Jansson et al. developed a new scoring system (MOTHER NAS scale), adding and removing items to the FNAS to create a 19-item scale, recommending treatment on scores ≥9. This score was used in the Maternal Opioid Treatment: Human Experimental Research (MOTHER) project. Subsequently, it was modified to develop a short screening tool, its validation awaits further studies. In 2015, institution 2 started assessing infants with NAS using the MOTHER NAS scale. From a data set of 17,150 observations in 276 infants over a period of 1 year, the sFNAS compared with the MOTHER NAS scale resulted in a Pearson’s correlation of 0.86. The sFNAS showed Sn=0.96, Sp=0.80, PPV=0.55, NPV=0.99 and PCC=0.83 to predict scores ≥9 on the MOTHER NAS scale. This correlation between the sFNAS and the MOTHER NAS scale is another support for the potential utility of the sFNAS in the assessment of NAS.

In 2013, Maguire et al. reported on an FNAS-short form. They performed factor analysis and their proposed scale contained seven items with a reported Pearson’s correlation of 0.917 with their original data set. Since this score is very similar to ours, we applied it to our data sets. The correlation between their seven-item score and the original FNAS score was 0.818 and 0.811 for institutions 1 and 2, respectively, which is lower than our correlations of 0.914 and 0.908 for institutions 1 and 2, respectively. This difference may be attributed to having more items included in the sFNAS.

As to limitations, the sFNAS was solely derived to shorten the original FNAS and not to improve on the psychometric properties of the original tool. Our study did not include a prospective clinical evaluation of inter-rater reliability, and therefore this will need to be addressed in future studies. Also, our study was not designed to determine the utility of the sFNAS in adjusting doses for pharmacological treatment; this is the next logical step after prospectively establishing its inter-rater reliability, sensitivity and specificity as a tool for initiation of treatment. Future studies should also include determination of the proportion of infants correctly classified as meeting threshold cut-offs from both the sFNAS and the original FNAS, as well as the percentage of those with discrepant scores. Indeed, the lack of uniformity in the assessment of infants with NAS can complicate the clinical care and increase the hospital costs for the affected infants.

CONCLUSION

The sFNAS provides a shortened and simplified assessment for infants with NAS. Developed with a rigorous statistical approach and with cross-validation, the sFNAS is not only abbreviated, but easily administered with minimal handling or interaction with the infant. The sFNAS having a high correlation with the original FNAS and with the MOTHER NAS scale makes it an attractive, efficient and simple alternative to the use of these lengthy tools. Further studies are needed to establish clinical utility, validity and reliability prior to a widespread application of the sFNAS.

Contributors EGP designed the study, was involved in data collection, results’ interpretation, drafted the initial manuscript, subsequent revisions and approved the final manuscript as submitted. LPF conceptualised the study, was involved in the analysis and interpretation of results, in review of draft and revisions, and approved the final manuscript as submitted. LD collaborated in the design of the study, involved in the data collection, analysis and interpretation of the results, critically reviewed the manuscript, and approved the final manuscript as submitted. VAC participated in the concept and design of the study, was involved in the data collection, critically reviewed the manuscript and approved the final manuscript as submitted. KTI participated in the concept and design of the study, designed the data collection instrument, involved in the data collection, interpretation of results, critically reviewed the manuscript and approved the final manuscript as submitted. HB co-conceptualised the study, collaborated with the design and development of the data collection instrument, was involved in the analysis and interpretation of the results, critically reviewed the manuscript draft and revisions, and approved the final manuscript as submitted. PMW carried out the extensive statistical analysis, was involved in the review and interpretation of results, critically reviewed the manuscript and approved the final manuscript as submitted. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Retroactive chart review of discharged patients.

Ethics approval Institutional Review Board of the University of Louisville and University of Kentucky.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All available data were used in the development of this study and can be obtained by contacting the corresponding author.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. Finnegan LP, Connaughton JF, Kron RE, et al. Neonatal abstinence syndrome: assessment and management. Addict Dis 1975;2:141–58.
2. Patrick SW, Schumacher RE, Benneyworth BD, et al. Neonatal abstinence syndrome and associated health care expenditures: United States, 2000–2009. JAMA 2012;307:1934–40.
3. 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.
4. Micard S, Brion F. Management of the opioid withdrawal in the neonates: French and European survey. Arch Pediatr 2003;10:199–203.
5. Tolia VN, Patrick SW, Bennett MM, et al. Increasing incidence of the Neonatal abstinence syndrome in U.S. Neonatal ICUs. N Engl J Med 2015;372:2118–26.
6. Bagley SM, Wachman EM, Holland E, et al. Review of the assessment and management of neonatal abstinence syndrome. Addict Sci Clin Pract 2014;9:19.
7. Lipsitz PJ. A proposed narcotic withdrawal score for use with newborn infants: A pragmatic evaluation of its efficacy. *Clin Pediatr* 1975;14:592–4.

8. Zahorodny W, Rom C, Whitney W, *et al*. The Neonatal withdrawal inventory: a simplified score of newborn withdrawal. *J Dev Behav Pediatr* 1998;19:89–93.

9. Green M, Suffet F. The Neonatal narcotic withdrawal index: a device for the improvement of care in the abstinence syndrome. *Am J Drug Alcohol Abuse* 1981;8:203–13.

10. Hudak ML, Tan RC. Committee on D. Neonatal drug withdrawal. *Pediatrics* 2012;129:e540–60.

11. Sarkar S, Donn SM. Management of neonatal abstinence syndrome in neonatal intensive care units: a national survey. *J Perinatol* 2006;26:15–17.

12. Mehta A, Forbes KD, Kuppala VS. Neonatal abstinence syndrome management from prenatal counseling to postdischarge follow-up care: results of a national survey. * Hosp Pediatr* 2013;3:317–23.

13. Finnegan LP, Kron RE, Connaughton JF, *et al*. Assessment and treatment of abstinence in the infant of the drug-dependent mother. *Int J Clin Pharmacol Biopharm* 1975;12:19–32.

14. Finnegan LP. Neonatal abstinence syndrome: assessment and pharmacotherapy. N N, ed. *Current therapy in neonatal-perinatal medicine*. 2nd ed. Ontario: BC Decker, 1990.

15. Finnegan LP, Kaltenbach K, Syndrome NA. In: Hoekelman R, Nelson NM, eds. *Primary pediatric care*. 2nd ed. St. Louis: Mosby, 1992:1367–78.

16. Orlando S. An overview of clinical tools used to assess neonatal abstinence syndrome. *J Perinat Neonatal Nurs* 2014;28:212–9.

17. Zimmermann-Baer U, Nötzli U, Rentsch K, *et al*. Finnegan neonatal abstinence scoring system: normal values for first 3 days and weeks 5–6 in non-addicted infants. *Addiction* 2010;105:524–8.

18. Maguire D, Cline GJ, Parnell L, *et al*. Validation of the Finnegan neonatal abstinence syndrome tool-short form. *Adv Neonatal Care* 2013;13:430–7.

19. LIANG K-YEE, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13–22.

20. Bada HS, Bauer CR, Shankaran S, *et al*. Central and autonomic system signs with in utero drug exposure. *Arch Dis Child Fetal Neonatal Ed* 2002;87:F106–12.