Retrospective Evaluation of a New Neonatal Trigger Score

**OBJECTIVES:** To design and validate an objective clinical scoring system to identify unwell neonates, by using routinely collected bedside observations.

**METHODS:** A Neonatal Trigger Score (NTS) was designed by using local expert consensus and incorporated into a new observation chart. All neonates >35 weeks’ gestation admitted to the NICU over an 18-month period, and an age-matched “well” cohort, were retrospectively scored by using the newly constructed NTS and all established pediatric early warning system (PEWS) scores.

**RESULTS:** Scores were calculated for 485 neonates. The NTS score area under the receiver operating characteristic curve was 0.924 with a score of 2 or more predicting need for admission to the NICU with 77% sensitivity and 97% specificity. Neonates scoring ≥2 had increased odds of needing intensive care (odds ratio [OR] 48.7, 95% confidence interval [CI] 27.5–86.3), intravenous fluids (OR 48.1, 95% CI 23.9–96.9), and continuous positive airway pressure (OR 29.5, 95% CI 6.9–125.8). The NTS was more sensitive than currently established PEWS scores.

**CONCLUSIONS:** The NTS observation chart acts as an adjunct to clinical assessment, highlighting unwell neonates. Its simplicity allows successful and safe use by nonpediatric specialists. NTS out-performed PEWS, with significantly better sensitivity, particularly in neonates who deteriorated within the first 12 hours after birth (P < .001) or in neonates with sepsis or respiratory symptoms (P < .001). Neonates with a score of 1 should be reviewed and those scoring ≥2 should be considered for NICU admission for further management. *Pediatrics* 2013;131:e837–e842

**WHAT’S KNOWN ON THIS SUBJECT:** Trigger scores and early warning systems provide an objective measure of a patient’s condition, allowing earlier recognition of severe illness and adaptation of care. Such scores are established in adult and pediatric populations but remain unevaluated and rarely used in neonates.

**WHAT THIS STUDY ADDS:** This newly designed Neonatal Trigger Score provides an objective adjunct to multidisciplinary clinical assessment in detecting unwell neonates. It is more sensitive and specific than previously validated pediatric early warning system scores.
Trigger scores and early warning systems (EWSs) have been in use for many years. Providing an objective measure, they highlight unwell patients and those likely to deteriorate, enabling earlier recognition of severe illness and adaptation of care to focus resources on the most needy patients. Several pediatric early warning system (PEWS) scores exist and have been validated for use in both the emergency department and on the ward, but despite being used for infants on the pediatric ward, the neonatal population was largely overlooked during the evolution of the scores. Although validation studies included infants, none stratified analysis by age group or were sufficiently powered to assess the reliability of scoring in neonates; despite adopting age-specific cutoffs for heart rate and respiratory rate, the importance of each clinical indicator may vary with age. Therefore, a scoring system designed for a pediatric population cannot necessarily be extrapolated for use in neonates.

At present, there is no standardized clinical scoring system for neonates. Consequently, some hospitals developed their own neonatal observation charts (eg, Craigavon Area Hospital) (C.Flannigan, M. Hogan, personal communication, 2011), but none have been validated in providing an “unwell trigger score” in the same manner as other EWSs. The number of neonates suffering a catastrophic collapse after birth is small (0.05 per 1000 live term births), but larger numbers will require emergency admission for intensive care because of early-onset sepsis or respiratory compromise. Although there are limited data, term neonates have been documented to account for nearly half of intensive care admissions, even in tertiary units. Unwell neonates may initially demonstrate subtle clinical signs, but may rapidly decompensate if left unrecognized and untreated.

EWSs have been shown to identify trends independent of practitioner experience or clinical workload. Hence, we believe there is an important role for a system that provides midwives and clinicians with an objective clinical measure of changes in the physiologic status of neonates, and so designed a neonatal-specific trigger score by using routine nursing observations. This simple scoring system can be performed and interpreted by any member of the multidisciplinary team, making it easier to identify observation trends and allowing prompt intervention and institution of intensive care when needed.

**METHODS**

**Score Development**

The Neonatal Trigger Score (NTS) was developed by using expert group consensus and guidance from both the Neonatal Life Support 2010 guidelines and the National Institute for Clinical Excellence Postnatal care guideline. A neonatal scoring chart provided by Flannigan and Hogan (C. Flannigan, M. Hogan, personal communication, 2011) was also reviewed. The expert group included 5 consultant neonatologists, NICU nurses, and midwives. During the initial meeting, previously established early warning scores were reviewed and a total of 9 parameters proposed. During subsequent discussions, these were reduced to a total of 5 compulsory objective measures (temperature, heart rate, respiratory rate, respiratory distress, conscious level) and an additional objective measure if indicated by past history (pre-feed blood sugar level). This final selection was based on perceived clinical importance, ease of rapid assessment at the bedside, and interassessor reliability. Each parameter could score a minimum of 0, with maximum scores ranging from 1 to 3. The score from each separate parameter was then combined to generate a cumulative score (minimum 0, maximum 15) with a higher score reflecting greater deviation from “normal.”

**Participants**

All neonates born at >35 weeks’ gestation between February 1, 2010, and August 1, 2011, were eligible for inclusion. Those born in cardiorespiratory or respiratory arrest were excluded. Participants were divided into 2 groups:

Group 1: all neonates born in this period who required admission to the NICU from the labor or postnatal wards. Well neonates admitted to the NICU for social reasons (such as observation for development of neonatal abstinence syndrome) or for medical management of isolated jaundice or polycythaemia were excluded.

Group 2: a consecutive group of neonates born during the same period who remained well on the postnatal ward and did not require admission to the NICU.

**Clinical Data Collection**

Clinical data were extracted from the notes of all enrolled subjects. This included observations needed to complete the NTS, details of any intervention or treatments needed, and a discharge diagnosis for all neonates admitted to the NICU. In group 1, the “decision to admit” observations were recorded (ie, the observations documented at the time the infant was referred to the NICU for admission). If these had not been documented, the initial observations obtained immediately after transfer to the NICU were used as a substitute. In group 2, all recorded postnatal observations were recorded. Where data were not available for a given predictor, it was agreed in advance that it would be recorded as missing data, and no attempts would be made to derive data where data were not explicitly recorded. Moreover, infants in whom more than 20% of the final data set was missing were excluded, because it was agreed...
the score would be inaccurate and may introduce bias. The authors were split pairwise (L.W. and H.H., R.B. and M.K.) with each pair independently reviewing the case files and extracting data onto a standardized data-collection form, ensuring consistency and accuracy.

NTSs were calculated for all observation sets collected. Scoring was done by 2 independent investigators (L.W. and R.B.). Any score variations were adjudicated by a third investigator (H.H.). As admission for intensive care was being used as a surrogate marker for severe illness, group 1 was classified as the “unwell” cohort and group 2 formed the “well” cohort.

Score Evaluation
The data were analyzed by using SPSS version 17 (IBM SPSS Statistics, IBM Corporation, Armonk, NY). The sensitivity and specificity of the NTS were evaluated, including with receiver operating characteristic (ROC) curves, determining its ability to differentiate between well neonates and those requiring NICU admission, and intervention for different score cutoffs. Single-variable analysis was used to compare variables between groups 1 and 2. Unpaired $t$ tests were used to compare continuous variables and the $\chi^2$ test to compare categorical variables. Multivariable logistic regression was then undertaken to assess the independent associations between predictors and admission to NICU.

A comparison was also made to existing validated PEWS scores. Each scoring system used a varying number of items (ranging from 6 to 14), including some parameters, such as systolic blood pressure and capillary refill time, not included in NTS. All participants were re-scored by using these PEWS tools and the performance of the scores, in both detecting neonates needing admission to the NICU and also within specific diagnostic subgroups, was then compared by using the same methods as described previously.

Ethics
This study was approved by the hospital research and development committee and was deemed exempt from requiring ethical approval. Hospital data-management procedures were followed to ensure confidentiality of patient information.

RESULTS
A total of 270 neonates of gestation >35 weeks were admitted to the NICU over the 18-month period. Seventy-seven were excluded: 41 because of cardiopulmonary arrest at birth, 15 social admissions, 18 well neonates admitted for expectant treatment of jaundice or polycythemia, and 3 who did not meet the minimum data set requirement, resulting in 193 neonates in group 1. A total of 292 were recruited to group 2. The mean NTS was significantly higher for neonates in group 1 (2.8 compared with mean 0.35 in group 2, $P < .001$) (Table 1).

All single variables were significantly different between groups 1 and 2, which was upheld when multivariable analysis was performed on the continuous variables (Table 2). It was not possible to include the categorical variables in this analysis, as the zeros in the well cross-classification resulted in unstable estimates.

Score Cutoff
The area under the ROC curve was 0.924 (Fig 1). The sensitivity and specificity cutoffs of the NTS are shown in Table 3. From these, a cutoff score of 1 was decided as being a trigger score at which a neonatal doctor should be informed and a score of 2 reflecting the likelihood of admission to NICU (sensitivity 79.3%, specificity 93.5%).

A score of ≥2 was associated with increased need for intensive care (odds ratio [OR] 48.7, 95% confidence interval [CI] 27.5–86.3), intravenous fluid requirement (OR 48.1, 95% CI 23.9–96.9), or assisted ventilation with continuous positive airway pressure (OR 29.5, 95% CI 6.93–125.8). Conversely, a score of 0 was strongly predictive of infants being well enough to remain on the postnatal ward (OR 28.6, 95% CI 16.2–50.7).

Score Adjustment
Consideration was also given to which scoring parameters were the most predictive. Of the 19 well neonates who reached a score threshold of 2, 13 scored only because they were cold (lowest temperature 35.4°C), a common finding in neonates after birth. We therefore postulated that performance of the trigger score might be improved by excluding low temperature as a scoring parameter. An analysis of this adjusted score gave an area under the curve (AUC) of 0.936, with improved specificity. However, because of recent concerns over hypothermia being an unrecognized sign of sepsis, it was felt not appropriate to completely omit a low temperature. Therefore, a temperature <36°C will score 1, while a temperature between 36.0 and 36.4°C will score 2.

Comparison With PEWS
Using the published trigger threshold of each score and a score of 2 or more for NTS, our score was significantly more sensitive than all established PEWS tools in detecting unwell neonates (77% compared with 63%, 44%, 28%, 19%), with no loss to specificity (97% vs 98%).
and 100%). NTS performed especially well in neonates who clinically deteriorated within the first 12 hours after birth (P < .001), and in neonates with sepsis or respiratory symptoms (P < .001). Furthermore, the existing PEWS tools use a greater number of criteria, including several that are not routinely performed by midwives and the recording of which would substantially increase their workload.

**DISCUSSION**

Currently established early warning and trigger scores have not catered specifically to the neonatal population. The score described previously is the only evaluated neonatal-specific trigger score. By using specific measurable observations, this scoring system helps users to identify neonates on the labor or postnatal wards who are becoming unwell and are likely to require further intervention and intensive care. By using objective clinical parameters, the potential for variability when used by different individuals is limited. Although an obstacle common to the introduction of any EWS is the perceived extra workload on the nursing staff, alongside the implication they are not able to identify unwell patients, charting the observations on this new chart (Fig 2) adds little time to a routine bedside assessment with the color coding and “total score,” making it easier to identify trends and to know when to call for help.

This initial retrospective evaluation of the NTS revealed neonates scoring ≥2 were 48 times more likely to require NICU admission and subsequently need further intervention, such as assisted ventilation. Despite a score threshold of 3 having 100% specificity for unwell, the sensitivity is only 62.7%, so we believe using this cutoff would miss too many low-scoring but sick neonates. Reviewing all infants with a score of ≥2 would involve seeing only 1 extra well infant every 4 days in a hospital with 4000 deliveries per annum and so would not place a high burden on existing services.

By consecutively recruiting a large cohort of neonates, selection bias was...
reduced and the retrospective nature meant any clinical decisions could not have been influenced by the NTS. Because there is no single reliable gold standard definition of an unwell neonate, admission to NICU was used as a surrogate marker; however, this meant we did not identify neonates who were judged to be unwell enough to require an intervention on the postnatal ward, such as commencement of antibiotics or nasogastric feeds, but did not require NICU admission.

As seen from Table 3, not all unwell neonates requiring NICU admission reached the presumptive cutoff score of 2. When the case notes of these 40 neonates were reviewed, 15% \((n = 6)\) were admitted with bilious vomiting requiring surgical investigation and a further 13% \((n = 5)\) had isolated poor feeding, without any suspicion of sepsis. A symptom such as vomiting is difficult to incorporate into an objective scoring system but is obvious to all health care providers. During score conception, an additional scoring category was proposed: “concerned about patient,” which would have added 1 point to the total score. On further discussion, however, we elected to create a purely objective scoring system with the caveat that anyone concerned about an infant, regardless of their NTS, should seek appropriate help, which should be the case with any EWS.

Additionally, 6.5% \((n = 19)\) of the well infants reached the score threshold of 2. Most of these \((n = 13, 68\%)\) scored only because they were cold, a common finding in neonates after birth when there may have been delays in wrapping or dressing and the environmental temperature, especially in labor ward and theaters, is often low. Score adjustment by removing a temperature of 36.0°C to 36.4°C from the score marginally increased the AUC; however, significant hypothermia \((<36.0°C)\) was retained, as it may coexist with hypoglycemia and sepsis, highlighted in a recent National Patient Safety Signal. Infants with only mild temperature instability (temperature 36.0°C–36.4°C) and otherwise well, scoring 0 in all other categories, may be wrapped or put skin-to-skin. If after an hour the hypothermia persists, then even a seemingly well infant should be reviewed if the infant is not able to maintain a normal body temperature. Our score boasts improved sensitivity and comparable specificity over the currently established PEWS scores. Therefore, although neonates could be scored on PEWS charts, NTS is better at detecting unwell neonates. This was especially marked \((P < .001)\) in those with respiratory conditions requiring intensive care, which is particularly important, as respiratory illness accounts for a significant proportion of term infants admitted for intensive care (81% of neonates admitted to NICU in our study had respiratory distress or were tachypneic).

FIGURE 2
NTS observation and scoring chart.
There is also a significant difference between the aim of the NTS and the already established pediatric scores. Although pediatric EWSs are designed to detect children at risk for deterioration and allow earlier intervention on the ward with the aim of reducing ICU admissions,1,2,4–7 or assessment of children presenting to the emergency department,5 the NTS has a slightly different focus. As neonates can deteriorate rapidly and postnatal wards are not designed to look after unwell infants, any neonate detected at early stages of decompensation or illness is likely to need admission to the NICU for ongoing care. Therefore, we believe, the purpose of an NTS should be to detect these neonates to allow appropriate transfer.

CONCLUSIONS
This combined observation and trigger-scoring chart was designed to address an important challenge in neonatal clinical practice. This initial evaluation study leads us to recommend that if a neonate scores 1, the neonatal doctor should be informed so he or she can risk-stratify based on further information. If a neonate scores 2, he or she needs to be reviewed urgently, as the likelihood of needing intensive care admission is high. If any of the individual categories score a 3 (red column on chart) then the infant needs urgent attention and the midwife assessing the infant should decide whether a cardiac arrest emergency call is appropriate. Although performance of this score was good in this retrospective data set, the limitations of such a study must be acknowledged. Before widespread implementation, a further prospective study is required to fully evaluate this clinical risk scoring system, determining the impact on clinical decision-making and patient outcomes.

ACKNOWLEDGMENTS
We are grateful to Dr C. Flannigan and Dr M. Hogän, Craigavon Area Hospital, for sharing their version of a newborn early warning scoring system, to Dr Bob Phillips for his critical review of the manuscript and to the Neonatal Consultants at the Whittington Health for their support, with special thanks to Dr Wynne Leith.

REFERENCES
8. Edwards ED, Powell CV, Mason BW, Oliver A. Prospective cohort study to test the predictability of the Cardiff and Vale paediatric early warning system. Arch Dis Child. 2009;94(8):802–806
Retrospective Evaluation of a New Neonatal Trigger Score
Harriet Holme, Reena Bhatt, Marita Koumettou, Mark A.S. Griffin and Lucinda C. Winckworth

*Pediatrics* 2013;131;e837; originally published online February 18, 2013; DOI: 10.1542/peds.2012-0640

<table>
<thead>
<tr>
<th>Updated Information &amp; Services</th>
<th>including high resolution figures, can be found at: /content/131/3/e837.full.html</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 12 articles, 5 of which can be accessed free at: /content/131/3/e837.full.html#ref-list-1</td>
</tr>
<tr>
<td>Citations</td>
<td>This article has been cited by 1 HighWire-hosted articles: /content/131/3/e837.full.html#related-urls</td>
</tr>
<tr>
<td>Post-Publication Peer Reviews (P3Rs)</td>
<td>One P3R has been posted to this article: /cgi/eletters/131/3/e837</td>
</tr>
<tr>
<td>Permissions &amp; Licensing</td>
<td>Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: /site/misc/Permissions.xhtml</td>
</tr>
<tr>
<td>Reprints</td>
<td>Information about ordering reprints can be found online: /site/misc/reprints.xhtml</td>
</tr>
</tbody>
</table>
Retrospective Evaluation of a New Neonatal Trigger Score
Harriet Holme, Reena Bhatt, Marita Koumettou, Mark A.S. Griffin and Lucinda C. Winckworth

Pediatrics 2013;131;e837; originally published online February 18, 2013;
DOI: 10.1542/peds.2012-0640

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/131/3/e837.full.html