Development and initial validation of the EDIN scale, a new tool for assessing prolonged pain in preterm infants

T Debillon, V Zupan, N Ravault, J-F Magny, M Dehan

Abstract
Objective—To develop and validate a scale suitable for use in clinical practice as a tool for assessing prolonged pain in premature infants.

Methods—Pain indicators identified by observation of preterm infants and selected by a panel of experts were used to develop the EDIN scale (Échelle Douleur Inconfort Nouveau-Né, neonatal pain and discomfort scale). A cohort of preterm infants was studied prospectively to determine construct validity, inter-rater reliability, and internal consistency of the scale.

Results—The EDIN scale uses five behavioural indicators of prolonged pain: facial activity, body movements, quality of sleep, quality of contact with nurses, and consolability. The validation study included 76 preterm infants with a mean gestational age of 31.5 weeks. Inter-rater reliability was acceptable, with a κ coefficient range of 0.59–0.74. Internal consistency was high: Cronbach’s α coefficients calculated after deleting each item ranged from 0.86 to 0.94. To establish construct validity, EDIN scores in two extreme situations (pain and no pain) were compared, and a significant difference was observed.

Conclusions—The validation data suggest that the EDIN is appropriate for assessing prolonged pain in preterm infants. Further studies are warranted to obtain further evidence of construct validity by comparing scores in less extreme situations.

Keywords: prolonged pain; pain scale; assessment; validation; premature

Pain assessment in neonates is among the most difficult challenges faced by health professionals and clinical researchers. All existing tools for neonates were developed for acute or post-operative pain.4, 5 They are helpful for identifying an acute painful event or comparing analgesic treatments used to combat procedural pain. A large number of responses to pain were described in the course of the development of these tools, providing valuable semiological data.

However, several lines of evidence indicate that preterm infants are subjected to not only acute pain but also prolonged pain. Firstly, neonatal intensive care involves many painful procedures. For instance, Barker & Rutter7 found that the number of procedures performed during the neonatal intensive care unit (NICU) stay of an infant weighing 560 g at birth was 488. Secondly, hormonal or metabolic responses indicative of stress have been reported in neonates with hyaline membrane disease or recent surgery, and abated with analgesic treatment.8–10 Thirdly, preterm babies often spend a considerable time in hospital: in a recent study, for instance, the median duration of assisted ventilation in a cohort of 182 very low birthweight infants was 36 days, and the median hospital stay was 105 days.9 Intensive care provided over such a long period constitutes a prolonged painful experience.

No tools are available for assessing prolonged pain—that is, lasting several hours or days—in preterm babies. Clearly, there is a clinical need for such a tool.10 In a previous study, we began to describe prolonged pain semiology.11 Now, we seek to meet this need by developing and validating a new pain scale suitable for clinical practice.

Methods
Development of the EDIN (Échelle Douleur Inconfort Nouveau-Né, neonatal pain and discomfort scale) was a two phase process. In the first phase, useful indicators of pain were identified, the scale was drafted, and its content validity was evaluated. The data were gathered by observation over one year of preterm infants born at 25–36 weeks gestational age and admitted to an NICU or a conventional neonatal unit (CNU). Severity of illness was the main difference between the patients in these two units: critically ill neonates were admitted to the NICU and neonates requiring simple care to the CNU. Newborns with cerebral hypoxia-ischaemia were excluded because of the possibility of brain damage interfering with the evaluation of behavioural indicators. In the second phase of the study, the construct validity, inter-rater reliability, and internal consistency of the EDIN were assessed on the basis of a study of 76 preterm infants.

PHASE 1: IDENTIFICATION OF USEFUL PAIN INDICATORS AND EVALUATION OF CONTENT VALIDITY
Data collection was by video recording using a camera providing a full view of the infant’s face and body. Of interest were the following situations: (a) assisted ventilation for hyaline membrane disease; (b) prolonged assisted ventilation for bronchopulmonary dysplasia; (c) first days of necrotising enterocolitis; (d) period
after ductus arteriosus closure. All indicators of potential use for evaluating pain or wellbeing were discussed by a panel composed of neonatologists, nurses, psychologists, and physiotherapists. On the basis of a content validity evaluation of each indicator, the panel selected the indicators found to be highly relevant for assessing pain or wellbeing in preterm infants, and subsequently eliminated the indicators likely to be difficult to observe in clinical practice.

PHASE 2: EVALUATION OF CONSTRUCT VALIDITY, INTER-RATER RELIABILITY, AND INTERNAL CONSISTENCY

In this part of the study, the task of determining the EDIN scores was given to the nurses, who spent more time with the infants than the other health care providers. After an eight hour period of infant observation, the score was calculated. The nurses had no specific training in pain assessment and determined the EDIN scores while performing their usual work. Five different hospitals participated in the study, and the EDIN scale was tested by several groups of nurses. Gestational age, birth weight, place of residence, and diagnosis at admission were also recorded for each preterm infant.

Two comparisons were performed to begin to establish construct validity. Firstly, EDIN scores in NICU preterm infants were determined before and eight hours after intravenous fentanyl at a dose of 1 µg/kg/h after an initial bolus of 1 µg/kg. All the infants in this group had an EDIN score $\geq 7/15$ before fentanyl (for this scale, a higher score represents more pain). Secondly, EDIN scores were compared on the day of NICU admission (a day with procedure related pain) and on the day before discharge (a day without pain).

Inter-rater reliability was assessed by asking two nurses to independently determine EDIN scores in the same preterm infants admitted to a CNU. These assessments were also used to evaluate the homogeneity of the EDIN items, which reflects internal consistency.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Description</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Facial activity</strong></td>
<td>0. Relaxed facial activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Transient grimaces with frowning, lip purse and chin quiver or tautness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Frequent grimaces, lasting grimaces</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Permanent grimaces resembling crying or blank face</td>
<td></td>
</tr>
<tr>
<td><strong>Body movements</strong></td>
<td>0. Relaxed body movements</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Transient agitation, often quiet</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Frequent agitation but can be calmed down</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Permanent agitation with contraction of fingers and toes and hypertonia of limbs or infrequent, slow movements and prostration</td>
<td></td>
</tr>
<tr>
<td><strong>Quality of sleep</strong></td>
<td>0. Falls asleep easily</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Falls asleep with difficulty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Frequent, spontaneous arousals, independent of nursing, restless sleep</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Sleepless</td>
<td></td>
</tr>
<tr>
<td><strong>Quality of contact with nurses</strong></td>
<td>0. Smiles, attentive to voice</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Transient apprehension during interactions with nurses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Difficulty communicating with nurses. Cries in response to minor stimulation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Refuses to communicate with nurses. No interpersonal rapport. Moans without stimulation</td>
<td></td>
</tr>
<tr>
<td><strong>Consolability</strong></td>
<td>0. Quiet, total relaxation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Calms down quickly in response to stroking or voice, or with sucking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. Calms down with difficulty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Disconsolate. Sucks desperately</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1** EDIN scale (Échelle Douleur Inconfort Nouveau-Né, neonatal pain and discomfort scale). Scoring method: nurses observe the infant for several hours during and between caring and feeding, and test the efficacy of consoling. They then score each EDIN item and calculate the total EDIN score as the sum of the five items.
VALIDITY INDICATORS AND EVALUATION OF CONTENT (n = 12).

Informations (n = 3), or nasogastric tube nutrition (n = 5), intestinal obstruction (n = 4), malformations (n = 3), or nasogastric tube nutrition (n = 12).

The reason for admission was respiratory distress syndrome (n = 45), maternal-fetal...

Mean birth weight was 1667 g (range 750–4000). The reason for admission was respiratory distress syndrome and mean birth weight was 1667 g (range 750–4000).

Phase 2 was conducted in 76 preterm infants, mean gestational age was 31.5 weeks (range 26–36) and mean birth weight was 1667 g (range 750–2980). The reason for admission was respiratory distress syndrome (n = 45), maternal-fetal infection (n = 12), necrotising enterocolitis (n = 5), intestinal obstruction (n = 4), malformations (n = 3), or nasogastric tube nutrition (n = 12).

STATISTICAL ANALYSIS

All data are reported as mean (SD) or mean (SEM). Between group differences were evaluated using paired t tests and p values < 0.05 were considered statistically significant.

In the assessment of inter-rater reliability, the level of agreement between the two nurses was determined for each item based on the weighted κ coefficient, which expresses the level of agreement beyond the level that can be attributed to chance alone.12 κ values ≥ 0.75 indicate excellent agreement, κ values < 0.4 indicate poor agreement, and intermediate values indicate acceptable agreement. Internal consistency was evaluated on the basis of Cronbach’s α coefficient; α values ≥ 0.8 indicate good internal consistency.

Results

Phase 1: Identification of useful pain indicators and evaluation of content validity

Indicators of pain and wellbeing recorded during the observation period were essentially behavioural and physiological markers. Among these indicators, those selected by the panel fell into the following five groups: (a) facial expression, (b) body movements, (c) quality of sleep, (d) quality of contact with nurses or sociability, and (e) consolability. Furthermore, the panel individualised two syndromes denoting prolonged pain: agitation and immobility. Agitation syndrome was characterised by facial contraction, often with cry characteristics; excessive spontaneous body movements (continual flexion-extension of the limbs with appreciable hypertonia); an inability to sleep; failure to respond to consolation by voice or stroking; and exacerbation of the above symptoms during interactions initiated by the nurse. Immobility syndrome was defined as facial contraction with a blank expression; a paucity of spontaneous movements; frequent arousals in response to mild stimulation; and an increase in facial contraction with appreciable hypertonia of the limbs during interactions with the nurses. The panel excluded all physiological indicators based on lack of specificity and difficulty of observation over several hours. Thus, the EDIN score was the sum of each of the scores for the five behavioural items described above, with each item being scored on a four point scale where 0 indicated wellbeing and 3 severe prolonged pain (fig 1).

Phase 2: Evaluation of construct validity, inter-rater reliability, and internal consistency

Phase 2 was conducted in 76 preterm infants, 40 in an NICU and 36 in a CNU. Mean gestational age was 31.5 weeks (range 26–36) and mean birth weight was 1667 g (range 750–2980). The reason for admission was respiratory distress syndrome (n = 45), maternal-fetal infection (n = 12), necrotising enterocolitis (n = 5), intestinal obstruction (n = 4), malformations (n = 3), or nasogastric tube nutrition (n = 12).

Construct validity

EDIN scores before and eight hours after analgesia were compared in the 40 preterm infants admitted to an NICU for respiratory distress syndrome. The mean total EDIN score was significantly lower after than before analgesia (4.7 (2.1) v 9.2 (1.7); mean (SD); p < 0.0001, paired t test) (fig 2). The mean (SEM) difference between before and after fentanyl was 4.4 (0.4) (95% confidence interval 3.6 to 5.2).

In the 36 CNU infants, EDIN scores on the admission day and the day before discharge were compared. Mean (SD) postconceptional ages at these two time points were 33.9 (2.5) and 37.5 (1.7) weeks. The mean (SD) total EDIN score was significantly lower before discharge than at admission (1.5 (1.5) v 4.5 (3.7); p = 0.0009, paired t test) (fig 3). The mean (SEM) difference between the two scores was 3.0 (0.5) (95% confidence interval 2.0 to 3.9).

These differences suggest that the EDIN discriminated between pain and no pain situations in our populations of preterm infants.

Inter-rater reliability

EDIN scores were determined independently by two nurses in each of the 36 CNU infants. The κ coefficients were 0.66 for facial expression, 0.72 for body movements, 0.74 for quality of sleep, 0.59 for sociability, 0.65 for consolability, and 0.69 for the total score.

Internal consistency

Internal consistency was evaluated in the 36 CNU infants. Cronbach’s α coefficients calculated after deletion of each individual item were
A new tool for assessing prolonged pain in preterm infants

0.94 for facial expression, 0.89 for consolability, and 0.86 for quality of sleep and body movements. The standardised item \( \alpha \) for the five items was 0.92.

**Discussion**

Pain is a subjective phenomenon that is difficult to assess in infants who are too young to communicate their distress by words. Scales have been developed to allow objective evaluation of pain in infants. The EDIN scale uses five behavioural items to identify and quantify prolonged pain in preterm infants. The EDIN scale score at a given point in time reflects the presence of wellbeing or pain during the preceding hours. The main findings from our study were as follows: (a) content validity was established on the basis of an evaluation by a panel of experts; (b) inter-reliability was acceptable as shown by the \( \kappa \) coefficient values of 0.4–0.75 for each item and for the total score; (c) internal consistency was high, with Cronbach’s \( \alpha \) coefficients ranging from 0.86 to 0.94 for each item and equal to 0.92 for the five items; (d) construct validity was confirmed by the significant difference between scores before and after analgesic treatment and during and after a period of pain.

**ITEM SELECTION**

All the items of the EDIN scale are behavioural indicators—that is, the EDIN scale is unidimensional. Pain, however, is a complex multidimensional phenomenon, and several tools for evaluating acute pain use a combination of physiological, behavioural, and contextual indicators. Physiological indicators such as heart rate or oxygen saturation have been found to be useful for identifying responses of neonates to acute noxious stimuli. However, physiological indicators are not well suited to the assessment of prolonged pain: whereas in acute pain, changes in physiological indicators are concurrent with the painful event, making the link between the two easy to grasp, in prolonged pain, physiological indicators are often modified not only by the pain but also by disorders such as respiratory distress syndrome or apnoea with bradycardia. In other words, physiological indicators are not specific for prolonged pain. Consequently, the panel excluded all physiological indicators from the EDIN scale.

Behavioural state, illness severity, gestational age at birth, and pain experience have been shown to influence the response of infants to acute pain. These factors were not included in the EDIN scale, and we did not investigate their potential effects in the setting of prolonged pain. Further studies in a typical situation of prolonged pain, such as necrotising enterocolitis, would be useful to investigate the potential impact of these factors. Factors found to have a significant impact could then be incorporated into the EDIN scale to make this tool more consistent with the multidimensional nature of pain. In this first study, we preferred the development of a global score which could be used easily in clinical practice.

Some indicators, such as quality of contact or consolability, may seem too subjective to allow accurate assessment. However, inter-rater reliability was good for these items in our study. This result could be explained by the prolonged relation between nurses and preterm infants: over several hours, the nurses observed the newborns in different situations and determined their responses to stimulations. Consolability and sociability have been used in scales developed for the assessment of postoperative pain, and an adverse effect of prolonged pain on interactions with others has been shown in older infants. Evaluation of consolability and sociability requires extensive experience with neonates. For this reason, neonatal unit nurses may well be the key persons for identifying prolonged pain in their patients.

In the EDIN scale, changes in facial activity and body movements are described using common language words such as “frequent” or “transient”, whereas terms allowing more accurate quantification are used in the Neonatal Facial Coding System and in the Infant Body Coding System. These two scales require a trained observer or a video recording and are consequently not easy to use in clinical practice. Because suitability for clinical practice was a priority for the EDIN scale, simple terms were used to describe each item. Furthermore, we found that inter-rater reliability was good for both facial activity and body movements.

**VALIDATION OF THE EDIN SCALE**

Establishing construct validity was both the most important and the most difficult step of our validation study, because no yardstick was available. Comparison of scores in situations with and without pain has been used by other authors to establish construct validity. To determine whether the EDIN scale, some discriminated satisfactorily between situations with and without pain, we compared the same infants before and after analgesic treatment, and at admission and on the day before discharge. In both comparisons, the EDIN scale score was significantly lower in the no pain situation. However, pain was not the only factor that differed between the two situations. Furthermore, in the admission versus discharge comparison, postconceptional age was higher at the time of the second evaluation; however, at the time of the first evaluation, it is unlikely that the lower postconceptional age contributed to the higher EDIN scale score. Indeed, it has been reported that behavioural responses to acute pain became stronger as postconceptional age increased. Our findings suggest construct validity, but further studies are needed to investigate the discriminating efficacy of the EDIN scale by comparing scores in less extreme situations.

The high internal consistency shown in our study is probably ascribable to the unidimensional nature of the EDIN scale, because it suggests that all five items measure the same component of prolonged pain and have similar importance in the assessment of pain. High internal consistency is more difficult to achieve
with multidimensional tools. Moderate correlations between the physiological and behavioural indicators of the Premature Infant Pain Profile (PIPP) have been reported. However, results obtained using the PIPP are consistent with the gate control theory, which holds that multiple indicators reflecting multiple dimensions provide more information than a single indicator.24

WHY A SCALE FOR PROLONGED PAIN ASSESSMENT? Many tools for assessing acute pain in neonates have been developed. Although these tools are useful for evaluating procedure related pain and treatment efficacy, they are not suitable for identifying pain that lasts for several hours or days. We are not aware of any tools for evaluating prolonged pain in neonates. Yet many abnormalities common in ill neonates undoubtedly cause prolonged pain; examples include abdominal distention during enteral nutrition and nasal lesions during nasal continuous positive pressure ventilation. If not routinely looked for, prolonged pain in neonates can go unrecognised and therefore untreated. In particular, a blank facial expression and/or a paucity of body movements often fail to suggest pain; in our study, this pattern of behaviour was common in patients with necrotising enterocolitis. Use of the EDIN scale by the nurses once or twice a day allows the identification of prolonged pain in neonates.

The EDIN scale was developed not only to allow the identification of prolonged pain, but also to improve its management. Accurate evaluation of pain is essential to gauge patients’ needs in terms of analgesic treatment and to evaluate the efficacy of analgesics. A consensus has been reached that appropriate pain management is a key priority in neonatal care. Firstly, pain can induce short lived alterations in homeostasis. In one study, for example, laboratory test abnormalities were shown during the postoperative period and improved under analgesic treatment. Secondly, there is preliminary evidence that pain during the neonatal period may have long lasting effects on the future behaviour of the patient. Although data on these effects remain limited, findings from studies in full term patients suggest that acute pain during the neonatal period may be followed by increased reactivity to subsequent pain.25 On the other hand, Johnston and Stevens21 reported that, in preterm infants, behavioural reactivity decreased after a painful procedure during the neonatal period. Thirdly, pain should be controlled as effectively as possible for ethical reasons.

In summary, we have developed the EDIN scale for assessing prolonged pain in preterm infants in clinical practice. The findings from our validation study suggest that the EDIN scale is a reliable tool. However, further work is needed to assess the influence on EDIN scale scores of factors such as postconceptional age and experience with pain and to further evaluate construct validity, which is the key validation criterion for clinical scales.

We would like to thank Dr Anne Gauvain-Piquard for participating in the concept study. We thank the following neonatal intensive care units for their participation: Robert Debré Hospital, Paris; Trousseau Hospital, Paris; Mother and Child Hospital, Colmar.

Commentary
The fallacy that neonates are incapable of experiencing pain has been laid to rest. Neonates were thought to be incapable of pain perception because of an underdeveloped nervous system, immature pain receptors, and the lack of myelinisation. Furthermore, it was believed that neonates were not able to remember painful experiences, and hence were not capable of interpreting noxious stimuli as painful. Developments in the field have shown that the basic somatosensory pathways and the anatomical and functional capabilities necessary for pain perception are developed at or
A new tool for assessing prolonged pain in preterm infants

F41

before birth. Although the mechanisms involved in memory are not well understood, the persistence of altered behaviour after painful events suggests that newborn infants respond differently after exposure to a single or multiple painful experiences. If memory is defined as experience dependent changes in behaviour, then existing data support the persistence of memories of pain in premature and full term newborns.3

Neonatal pain perception has been shown by a number of research studies. From a behavioural perspective, motor responses, changes in facial expression, and crying characteristics were associated with the pain from invasive procedures. Physiological pain responses include increases in heart rate, respiratory rate, intracranial pressure, and palmar sweating as well as decreases in oxygen saturation and vagal tone. Biochemical responses to pain associated with surgical stress have been observed and include increases in cortisol, adrenaline (epinephrine), noradrenaline (norepinephrine), growth hormone, and endorphins as well as suppression of insulin.4,5 Contextual factors such as gestational age, severity of illness, and behavioural state have been associated with the modulation of behavioural pain responses after a standardised noxious stimulus.

Recent progress in the study of neonatal pain supports: (a) the neurophysiological and clinical basis for an increased sensitivity to pain in neonates, (b) the prolonged hyperalgesia that follows acute painful stimuli in preterm neonates, leading to prolonged or chronic pain, and (c) the potential association of neuro-behavioural and developmental sequelae resulting from repetitive painful experiences during neonatal intensive care. Neonatal experiences from insufficiently relieved pain can have long term effects reflected in altered pain thresholds and other behaviours during the infancy and childhood of ex-preterm neonates.4 Persistent effects of early painful experience may be indicated by the increased prevalence of psychosomatic complaints that ex-premature children have later in life compared with ex-full term neonates.

Neonates are subjected to clinical practice to different types of pain, acute and prolonged. The diagnosis of pain in neonates remains, however, a major challenge for parents and health professionals alike. The lack of specific indicators for pain and the poor training of clinicians in recognising pain are commonly cited factors. Accurate assessment of neonatal pain is nevertheless imperative for ensuring comfort during the diagnostic process and in evaluating the effectiveness of pain treatment modalities.

A state of the art study4,5 identified 16 methods for the assessment of acute pain in neonates. These measures, for the most part, have inadequately developed psychometric properties and clinical utility. No tools have been developed as yet for assessing prolonged pain in preterm babies. There is undoubtedly a need for the development of such an assessment tool. The authors of this paper are therefore to be commended for addressing this subject and for their systematic approach to tool development.

In the development of assessment tools, reliability, validity, and clinical utility are of particular importance. Adequate attention has been paid in this study to establishing the reliability, content, and construct validity of the EDIN. More attention needs to be given to validity issues such as criterion related, discriminant, and clinical validity as well as to sensitivity, specificity, and clinical utility for use of the tool in practice.

Neonates need to be comfortable and as free from pain as possible in order to grow and develop normally. Valid and reliable assessment of pain is the major prerequisite for attaining this goal.

H H ABU-SAAD

Professor of Nursing Science
Director, Centre for Nursing Research,
Maastricht University
PO Box 616
6200 MD Maastricht
The Netherlands
h.huyer@zin.ruimaas.nl