Do 0–10 Numeric Rating Scores Translate into Clinically Meaningful Pain Measures for Children?

Terri Voepel-Lewis, MS, RN, Constance N. Burke, BSN, RN, Nicole Jeffreys, MD, Shobha Malviya, MD, and Alan R. Tait, PhD

BACKGROUND: Self-reported pain scores are used widely in clinical and research settings, yet little is known about their interpretability in children. In this prospective, observational study we evaluated the relationship between 0 to 10 numerical rating scale (NRS) pain scores and other self-reported, clinically meaningful outcomes, including perceived need for medicine (PNM), pain relief (PR), and perceived satisfaction (PS) with treatment in children postoperatively.

METHODS: This study included children ages 7 to 16 years undergoing surgery associated with postoperative pain. One to 4 observations were recorded in each child within the first 24 hours postoperatively. At each assessment, children rated their pain with the NRS, stated their PNM, and rated their satisfaction with pain management. Assessments were repeated within 1 to 2 hours, and children additionally rated their PR as the same, better, or worse in comparison with the earlier assessment. Receiver operator characteristic curves were developed to examine potential NRS cut-points for PNM and PS, and the minimum clinically significant difference (MCSD) in pain score associated with PR was calculated.

RESULTS: Three hundred ninety-seven observations (including 189 pairs) were recorded in 113 children. NRS scores associated with PNM were significantly higher than “no need” (median 6 vs. 3; P < 0.001). NRS scores >4 had good sensitivity (0.81) and specificity (0.70) to discriminate PNM, but with a large number of false positives and negatives (e.g., 42% of children with scores >4 did not need analgesia). The MCSD in NRS scores was −1 (95% confidence interval [CI] −0.5 to 1) or +1 (CI 0.5 to 2.7) in relation to feel “a little better” or “worse,” respectively (P < 0.001 vs. the same). NRS scores >6 had a sensitivity of 0.82 and specificity of 0.76 in discriminating dissatisfaction with treatment, yet 46% and 24% of children with scores >6, respectively, were somewhat to very satisfied with their analgesia.

CONCLUSIONS: This study provides important information regarding the clinical interpretation of NRS pain scores in children. Data further support the NRS as a valid measure of pain intensity in relation to the child’s PNM, PR, and PS in the acute postoperative setting. However, the variability in scores in relation to other clinically meaningful outcomes suggests that application of cut-points for individual treatment decisions is inappropriate. (Anesthesiology 2011;112:415–21)

Self-reported pain scores are widely recommended and used to assess pain in pediatric research and clinical settings. Guidelines on the assessment and management of acute pain, indeed, recommend that pain intensity be assessed before the initiation of and continually throughout treatment, and that dosing adjustments be made, in part, on the basis of the patient’s self-report. Recommendations for core measures in pediatric pain trials further state that pain intensity is an obvious outcome measure. Despite their widespread use, few studies have addressed the clinical meaning or interpretability of self-reported pain scores in children. A better understanding of their meaning is necessary if, as guidelines suggest, pain intensity be assessed before the initiation of and continually throughout treatment, and that dosing adjustments be made, in part, on the basis of the patient’s self-report.

A number of studies have evaluated the relationship between pain scores and other clinically relevant indicators in adults, including the patient’s perceived need for medicine (PNM) or pain relief (PR), yielding data about potential “cut-points” for treatment or minimum clinically significant differences (MCSD) in scores. Although a large body of literature has addressed the psychometric properties of self-report tools in children, few have investigated the interpretability of pain scores. One survey asked hospitalized children ages 6 to 16 years postoperatively regarding what pain score on the 0 to 6 Faces Pain Scale would be the minimum that would warrant treatment (i.e., “a pill to help the hurt/pain go away”). This study reported that a median pain score of 3 of 6 (mean 3.2 ± 1.8) was associated with the child’s PNM. Another study evaluated the change in Visual Analog Scale (VAS; 0 to 100 mm) scores associated with children’s perceived PR in the emergency department setting. A 10-mm change in VAS score (95% confidence interval [CI] 7 to 12 mm) was the minimum difference that evoked a feeling of “a bit better” or “a bit worse” in this sample. These studies provide some understanding, albeit limited, of the interpretability of pain intensity scores in children. Data related to the 0 to 10 numeric rating scale (NRS), which is likely the most commonly used self-report tool, are minimal. A recent
study from von Baeyer et al. demonstrated excellent correlations between scores on the NRS, Faces Pain Scale, and VAS, supporting the criterion validity of the NRS in children ages 7 years and older.17 Interestingly, they reported an increased frequency of scores near the lower anchor (i.e., “no pain”) with the VAS but not with the NRS, suggesting a tendency for children to shift scores when only end-point descriptors are available, a finding similar to previous studies.18,19 These data suggest a potential advantage for using the NRS, which offers more meaningful and salient numerical points along the scale.17 These investigators and others—including the recent Pediatric Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials panel—have acknowledged the need for further study of this commonly used tool in children.2,3,16,17

The purpose of this prospective, observational study was, therefore, to evaluate the relationship between pain scores derived from the NRS and other self-reported outcome measures in children in a postoperative setting. The specific aims of the study were to (a) describe the relationship between NRS pain scores and the child’s PNM; (b) describe the change in pain score associated with perception of feeling better or worse (i.e., PR); (c) describe pain scores associated with the child’s perceived satisfaction [PS] with pain treatment; (d) examine the sensitivity of NRS scores to correctly identify children who need medicine or who are dissatisfied with treatment; (e) examine factors (e.g., age) that modify the relationship between NRS pain scores and other relevant outcomes.

METHODS

Approval from the IRB MED at the University of Michigan and written consent and assent from parents and children were obtained before study recruitment. Children ages 7 to 16 years who required hospital admission after surgery associated with postoperative pain (e.g., spine fusion and splenectomy) were consecutively recruited during their preoperative assessment by trained research assistants. Children were excluded if they did not speak English or were not engaged during the survey responses.

Preoperatively, demographics and pertinent medical history, including the child’s previous surgery experience, were documented. Children were tested for their ability to self-report pain with previously described measures of the concepts of magnitude (rank/size) and order using blocks of differing sizes and a series of numbers.20,21 Children who passed all tests were deemed able to reliably self-report their pain intensity using the NRS. All children were provided with the following information regarding use of the 0 to 10 NRS to assess pain after surgery: “You will be asked to rate your pain by choosing a number from 0 to 10 that best tells us how much you are hurting, where 0 = no pain or hurt and 10 = the most or worst pain/hurt.”

Postoperative pain management was provided via patient-controlled analgesia (PCA) or intermittent analgesia per routine care at the discretion of the care provider(s). Observations were independent of the care providers’ assessments, and were not used for the purposes of clinical intervention, although concerns raised during observation sessions were conveyed to care providers. All observations were made only when children were awake (at least 1 hour after awakening), during the first 24 hours postoperatively by trained research assistants who recorded, verbatim, the child’s responses to each question, and who were blinded to interventions before or between observations. Because the aim of this study was to determine clinical interpretability (i.e., meaning of a pain score) from the child’s perspective (e.g., need for rescue or bolus), observations were not timed in relation to interventions. The following procedures are well documented in previous studies conducted in acute pain settings.5,9,12–14 At each observation the child was first asked to rate his or her pain using the 0 to 10 NRS at that point in time. Children were then asked to rate their PNM (i.e., “Do you feel the need for more pain medicine right now?”) and their PS (i.e., “How satisfied (or happy) are you, right now, with your overall pain treatment/medicine?”). To examine the relationship between changes in pain scores over time and the child’s perceived PR, assessments were repeated within 2 hours after the first observation, and children were also asked whether their pain was “a lot better,” “a little better,” “the same,” “a little worse,” or “a lot worse” than in the earlier assessment. No more than 1 paired observation was recorded on the day of or morning after surgery. Although parents were generally present at the bedside, they were not engaged during the survey responses.

Data were analyzed using SPSS statistical software (v. 17.0.3; Chicago, Illinois), and are described as n (%), mean ± sd, median, and 95% CI, as applicable. Pain scores and ordinal scale data were treated as nonparametric data, and all comparisons were 2-tailed. Mann–Whitney U tests were used to compare pain scores related to dichotomous grouping measures (e.g., PNM). Kruskal–Wallis with Dunnett C post hoc analyses were used to compare pain scores across the PS and PR scales. Because up to 4 observations were obtained per subject, Bonferroni’s correction was applied and, thus, P values <0.0125 were considered significant. Receiver operating characteristic (ROC) curve analyses were used to assess the ability of the NRS to discriminate the child’s PNM and dissatisfaction. Areas

<table>
<thead>
<tr>
<th>Table 1. Description of Sample (n = 113)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
</tr>
<tr>
<td>&lt;12 years</td>
</tr>
<tr>
<td>Female/male</td>
</tr>
<tr>
<td>ASA-PS 1–2</td>
</tr>
<tr>
<td>ASA-PS 3–4</td>
</tr>
<tr>
<td>Previous surgery</td>
</tr>
<tr>
<td>Surgical procedure</td>
</tr>
<tr>
<td>Orthopedic spine</td>
</tr>
<tr>
<td>Orthopedic peripheral</td>
</tr>
<tr>
<td>General surgery</td>
</tr>
<tr>
<td>Otology</td>
</tr>
<tr>
<td>Method of pain management</td>
</tr>
<tr>
<td>Intermittent</td>
</tr>
<tr>
<td>Patient controlled</td>
</tr>
<tr>
<td>Timing of observations (n = 397)</td>
</tr>
<tr>
<td>Evening of surgery</td>
</tr>
<tr>
<td>Morning after surgery</td>
</tr>
<tr>
<td>Afternoon day after surgery</td>
</tr>
</tbody>
</table>

ASA-PS = American Society of Anesthesiologist’s Physical Status Classification. Data are presented as n (%) or mean ± SD.
under the curve (AUC) of 0.70 to 0.80 were considered fair, and 0.80 to 1.00 as good to excellent measures of accuracy/discrimination. The sensitivity and specificity of the derived NRS cut-points related to PNM and dissatisfaction are described. Lastly, data were analyzed by age group, previous surgery experience, and sex to examine whether these factors modified the relationships between pain scores and other measures.

The sample size was determined a priori on the basis of the following assumptions. Given an estimated difference in pain scores of $6.6 \pm 2.31$ versus $4.5 \pm 2.94$ in patients who need medicine versus those who did not, at least 45 observations per group ($\alpha = 0.0125; \beta = 0.90$) were necessary. Pilot data suggested that approximately 20% of pain scores in the early postoperative period were in the severe range (i.e., 7 to 10). To obtain a sufficient sample of observations with severe pain levels (i.e., potentially need medicine), 250 observations were needed. This sample was deemed sufficiently large to detect a significant correlation between pain scores and satisfaction (estimated $r = 0.23$).

**RESULTS**

One hundred thirteen children were included and described in Table 1. Three hundred ninety-seven observations were recorded in the following manner: 2 paired sets in 86 children (i.e., 344 separate observations), 1 pair in 17 ($n = 34$), and 19 unpaired observations in 10.

**Perceived Need for Medicine (PNM)**

NRS scores associated with PNM were significantly higher than were scores associated with “no need” in this sample (median 6 vs. 3; $P < 0.001$); however, a wide range of scores was reported in each group (Fig. 1). Figure 2 displays the ROC curve examining the relationship between NRS scores and PNM. The AUC of 0.84 reflects the ability of NRS scores to discriminate the child’s PNM. A NRS score $>4$ (sensitivity 0.81; specificity 0.70) maximized the sum of sensitivity (true positives) and specificity (true negatives) in discriminating the PNM in this sample. Importantly, application of this cut-point yielded a high number of false positives, i.e., 42% of children with NRS $>4$ did not need analgesia. Pain scores associated with PNM were not significantly different between children receiving intermittent and those receiving PCA ($5.64 \pm 2.1$ vs. $6.5 \pm 2.2; P = 0.159$), nor between observations obtained on the day of surgery and those of the next day ($6.3 \pm 2.1$ vs. $6.5 \pm 2.2; P = 0.766$).

**Pain Relief (PR)**

Figure 3 displays the change in NRS scores associated with the child’s perceptions of PR as “a lot” or “a little” better or worse. These NRS changes were significantly different for all comparisons (i.e., “a lot” vs. “a little” better or worse, “a little” vs. “same,” etc.) except between “a little better” versus “the same.” The MCSD in NRS, i.e., the change needed for the child to feel a little better or worse, was $-1$ (CI $-0.5$ to 1) and $+1$ (CI 0.5 to 2.7), respectively ($P < 0.001$ vs. the same). In 27% of cases, the direction of the change in score did not match the child’s perceived PR (i.e., scores increased when the child said he or she felt better). Additionally, the MCSD in NRS scores associated with the child’s perception of feeling better was significantly lower for children whose baseline NRS was $\geq 5$ in comparison with those with NRS $\leq 5$ (median 0[CI $-2$ to 3] vs. $-2[-2.5$ to $-1.3]; P < 0.001$).

**Perceived Satisfaction (PS)**

Children stated they were very satisfied with their pain treatment during 215 (54%) of the observations, somewhat satisfied in 118 (30%) cases, not satisfied in 19 (5%), and undecided in 45 (11%). There was a significant inverse relationship between PS and NRS scores, and significant
Differences in scores between those reporting that they were very, somewhat, and not satisfied (Fig. 4; \( P < 0.001 \)). Furthermore, ROC curve analysis showed that NRS scores had good sensitivity and specificity in predicting dissatisfaction with treatment (Fig. 5). However, the majority of children rating pain were somewhat to very satisfied (46% and 24%, respectively). Although satisfaction was significantly lower for children who received intermittent rather than PCA, pain scores associated with dissatisfaction were not significantly different in these subgroups (6 ± 1.4 vs. 7.9 ± 1.6; \( P = 0.282 \)).

Differences Related to Age, Gender, and Previous Surgery
The relationships between underlying patient factors—including sex, age, and previous surgical experience—and pain scores related to the other outcomes (i.e., PNM, PR, PS) are presented in Table 2. Females, children who had no previous surgical experience, and those who were older reported satisfaction (somewhat to very satisfied) at higher pain scores than did other children.

DISCUSSION
This study is the first to evaluate the relationship between NRS pain scores and several verbal outcome measures in children. These data demonstrate that 0 to 10 NRS scores in children are, in general, reliably associated with the child’s PNM, perceived PR, and PS with treatment, supporting use of the NRS as a measure of pain intensity in clinical and research settings. Importantly, the large overlap in scores associated with these outcomes, and the number of false negatives and positives, suggest that application of cutoff scores for individual treatment decisions is inappropriate for children.

The notion of using pain score thresholds for treatment or rescue decisions has been suggested in clinical and research settings, yet previous studies have demonstrated significant variability in individual thresholds. Blumstein and Moore described poor sensitivity and specificity for the VAS in predicting the need for analgesia for adults in the emergency room. Gauthier et al. showed a significant overlap in scores related to the child’s perceptions of pain severity and thresholds for needing medicine. The variability between reported pain scores and individual treatment thresholds in these studies reflects the complexity of the pain experience for children and adults. The need for analgesia is likely dependent on many factors, including perceptions of medicine, importance of side effects, and individual pain temperament or sensitivity. Although the NRS had a general ability to detect the PNM in the present study sample, the number of false positives and negatives suggests that application of absolute cut-points to treatment decisions for individual patients is inappropriate.

Previous studies have shown that, in general, mean pain score changes of 9 to 13 mm on the VAS or 1 on the NRS represent the minimum change reflecting the patient’s perception of “more” or “less” pain. Powell et al. reported the MCSD on the VAS as 10 mm (CI 7 to 12) for children in the emergency room. The present study, similar to previous reports, found that a 1-point decrease or increase in NRS score was associated with perceptions of feeling “a little better” or “little worse,” respectively, and that a larger change (median = 2) was needed to impart perceived PR when baseline scores were higher (i.e., >5). This later finding is particularly important when calculating sample sizes and interpreting findings in clinical pain trials. Lastly, as was previously reported, there were a significant number of children whose perception (feeling better or worse) was in the opposite direction of the change in score, highlighting the complexity of pain assessment for the individual child.

The wide overlap in NRS scores in relation to satisfaction and dissatisfaction further emphasize the complexity of evaluating pain outcomes. Several studies in adults have demonstrated only poor-to-fair relationships between global satisfaction measures and pain intensity.
Similar to our findings, previous data depict a skewed distribution of satisfaction data, with most patients rating treatments as excellent or good, even those with moderate to severe pain. Importantly, the relationship between satisfaction and pain intensity may differ on the basis of the timing of assessment postoperatively, and phrasing of the satisfaction question. The present study compared a given pain score and PS with overall treatment at that time, and assessments occurred only during the first 24 hours after surgery. Further study is needed to evaluate the impact of time on this relationship in children. Lastly, although the survey question was phrased to assess satisfaction with treatment, other moderating factors, such as side effects, may have influenced the child’s response. Despite ongoing questions related to the relevance of satisfaction measures, they remain among the recommended outcome measures for acute and chronic pain trials.

Age, gender, and previous experience modified the relationship between NRS scores and PS but not PNM or PR in this sample. Older children, females, and those without previous experience reported satisfaction at higher NRS scores than did others. These findings reflect potential differences in coping ability, expectations, or anxiety. Indeed, previous investigators have demonstrated a higher pain threshold and differences in coping techniques used by older children during experimental pain situations. Others have described long-lasting effects of early (i.e., infancy) injury on somatosensory processing, as well as the effect of anxiety and conditioned fear on pain thresholds. The relationship between gender and perceptions of pain are less clear, given confounding findings in previous studies. Further study is necessary to better characterize the potential modifying effects of age, gender, and other factors on the pain experience in children.

Because these data were collected at 1 institution in English-speaking children only, the ability to generalize to other pediatric populations may be limited. The method of surveying children may have posed several limitations. Repeat observations were conducted 1 to 2 hours after the initial assessment, which may have influenced the child’s perceived PR. The fact that previous reports of the MCSD in pain scores in adults and children have yielded similar findings to ours, however, provides some external validity to our data. Other factors—such as analgesics used, their side effects, anxiety, parental presence, or unknowns—may have modified the child’s perceptions. However, the purpose of this study was to determine the...
clinical meaning of pain scores in a real postoperative setting, and as such, the setting with all of its confounders was highly relevant. Further qualitative study is necessary to better explore the effects of other factors. Lastly, this study may have been underpowered to detect threshold differences in the subgroups.

This study provides new data to describe the relationship between NRS scores and the child’s PNM, PR, and PS with treatment. Although findings support the use of the NRS in children ages 7 to 16 years for research and clinical practice, the notable variability between pain scores and other outcomes highlight the importance of individualizing pain assessment, and treatment.

**DISCLOSURES**

Name: Terri Voepel-Lewis, MS, RN.

Contribution: This author helped design the study, conduct the study, analyze the data, and write the manuscript.

Attestation: Terri Voepel-Lewis has seen the original study data, reviewed the analysis of the data, approved the final manuscript, and is the author responsible for archiving the study files.

Name: Constance N. Burke, BSN, RN.

Contribution: This author helped conduct the study.

Attestation: Constance N. Burke has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.

Name: Nicole Jeffreys, MD.

Contribution: This author helped conduct the study.

Attestation: Nicole Jeffreys has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.

Name: Shobha Malviya, MD.

Contribution: This author helped conduct the study and write the manuscript.

Attestation: Shobha Malviya reviewed the analysis of the data and approved the final manuscript.

Name: Alan R. Tait, PhD.

Contribution: This author helped analyze the data and write the manuscript.

Attestation: Alan R. Tait has seen the original study data, reviewed the analysis of the data, and approved the final manuscript.

**ACKNOWLEDGMENTS**

The authors gratefully acknowledge the contributions of the following persons: Deepa Bhat, Elsa Pechliivanidis, Jennifer Hemberg, Lauren Perlin, and Caela Hesano.

**REFERENCES**


7. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004;8:283–91


Table 2. Relationship Between Numerical Rating Scale (NRS) Scores and Outcomes Based on Underlying Patient Factors

<table>
<thead>
<tr>
<th>Gender</th>
<th>Need medicine</th>
<th>Satisfied</th>
<th>Feel “a little better”*</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>6 [5.5–6.9] (51)</td>
<td>3 [3.2–4.1] (137)</td>
<td>−1 [−1.8–0.4] (77)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>7 [6.1–7] (78)</td>
<td>4 [3.9–4.6] (196)</td>
<td>−1 [−1.7–0.4] (112)</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Previous surgery

<table>
<thead>
<tr>
<th>Gender</th>
<th>Need medicine</th>
<th>Satisfied</th>
<th>Feel “a little better”*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6 [5.85–7.1] (65)</td>
<td>3 [3.4–4.2] (196)</td>
<td>−1 [−1.7–0.41] (109)</td>
</tr>
<tr>
<td>No</td>
<td>7 [6.0–6.9] (64)</td>
<td>4 [4–4.8] (137)</td>
<td>1 [1.5–0.5] (80)</td>
</tr>
</tbody>
</table>

Age groups

<table>
<thead>
<tr>
<th>Gender</th>
<th>&lt;12 years</th>
<th>≥12 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need medicine</td>
<td>6 [5.3–6.7] (40)</td>
<td>7 [6.2–7.1] (89)</td>
</tr>
<tr>
<td>Satisfied</td>
<td>3 [2.9–3.9] (100)</td>
<td>4 [4–4.6] (233)</td>
</tr>
<tr>
<td>Feel “a little better”*</td>
<td>−1 [−2.6–0] (60)</td>
<td>−1 [−1.5–0.5] (129)</td>
</tr>
</tbody>
</table>

Data are presented as median or as change in NRS score [confidence interval]: (number of observations in the groups).
29. Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? Acad Emerg Med 1998;5:1086–90