Published in final edited form as:

Ann Emerg Med. 2018 June ; 71(6): 691–702.e3. doi:10.1016/j.annemergmed.2017.09.009.

Validity and Reliability of the Verbal Numerical Rating Scale in Children Aged 4 to 17 Years with Acute Pain

Daniel S. Tsze, MD, MPH^a, Carl L. von Baeyer, PhD^b, Vartan Pahalyants, BA^a, and Peter S. Dayan, MD, MSc^a

^aDepartment of Pediatrics, Division of Pediatric Emergency Medicine, Columbia University College of Physicians and Surgeons. New York, New York, USA

^bFaculty of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada

Abstract

Objective—The verbal numerical rating scale (vNRS) is the most commonly used self-report measure of pain intensity. It is unclear how the validity and reliability of vNRS scores vary across children's ages. We aimed to determine the validity and reliability of the vNRS in children presenting to the ED across a comprehensive spectrum of age.

Methods—Cross-sectional study of children aged 4–17 years. Children self-reported their pain intensity using the vNRS and Faces Pain Scale–Revised at two serial assessments. We evaluated convergent validity (strong validity defined as correlation coefficient ≥ 0.60); agreement (difference between concurrent vNRS and FPS-R scores); known-groups validity (difference in score between children with painful versus nonpainful conditions); responsivity (decrease in score after analgesic administration); and reliability (test-retest at two serial assessments) in the total sample and subgroups based on age.

Results—We enrolled 760 children; 27 did not understand the vNRS and were removed. Of the remainder, Pearson correlations were strong to very strong (0.70–0.92) in all years of age except 4- and 5-year-olds, and agreement was strong in 8-year-olds. Known-groups validity and responsivity were strong in all years of age. Reliability was strong in all age subgroups, including each year of age from 4–7 years.

Conclusion—Convergent validity, known-groups validity, responsivity and reliability of the vNRS were strong in children 6–17 years of age. Convergent validity was not strong in 4- and 5-

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Meetings: The current data has not been previously presented.

Clinical Trials Registration Number: N/A

Conflict of Interest: None of the authors have any conflicts of interest to disclose.

Author Contributions: DST, CLvB, and PSD conceptualized and designed the study. DST supervised the conduct of the study and data collection. DST and VP undertook recruitment of participating patients. DST and CLvB conducted the statistical analyses for the study. DST drafted the manuscript, and all authors contributed substantially to its revision. DST takes responsibility for the paper as a whole.

year-olds. Our findings support the use of the vNRS in children 6 years and older, but not for use in 4- and 5-year-olds.

INTRODUCTION

Background

Pain is one of the most common reasons that a child presents to the emergency department (ED).^{1–3} The appropriate treatment of pain is dependent on the ability to readily and accurately assess a child's pain intensity. The verbal numerical rating scale (vNRS) is the most frequently used self-report measure of pain intensity in older children and adults with acute pain due to its simplicity and ease of use.^{4–11} It involves verbally asking for an estimate of pain intensity using numbers from 0 (no pain) to 10 (maximal pain), and requires no equipment to administer or score.

Importance

The vNRS is frequently used in children with acute pain aged ≥ 8 years, and has strong validity and reliability in this population.^{10,11} However, it is unclear whether the vNRS has strong psychometric properties in younger children, or if these properties may vary based on patient characteristics. For example, children aged < 7 years are often considered unable to accurately use the vNRS because they lack the necessary developmental skills, although recent studies suggest otherwise.^{12–14} In addition, the validity and reliability of vNRS scores may vary based on patient characteristics such as sex, race and ethnicity, or a child's primary language, which have been shown to be related to a child's ability to describe pain, as well as their perception of, sensitivity to, and experience with pain.^{15–21} Therefore, it is important to identify any differences in validity and reliability based on patient characteristics, particularly in younger children aged < 7 years, as such findings would impact the generalizability and implementation of the vNRS in children presenting to the ED with acute pain.

Goals of This Investigation

We aimed to determine the validity and reliability of the vNRS in children presenting to the ED across a comprehensive spectrum of age (4–17 years) and other patient characteristics. Our main hypothesis was that children 4 to 17 years of age would demonstrate strong convergent validity (Pearson correlation coefficient ≥ 0.60) when comparing the vNRS to the FPS-R within older (8–17 years) and younger (4–7 years) age groups. Our secondary aims were a) to determine agreement (another measure of convergent validity), known-groups validity, responsiveness, and reliability within older and younger age groups; and b) to determine convergent validity (Pearson correlation coefficient), agreement, known groups validity, responsiveness, and reliability within each year of age for 4–7 years, for the purpose of identifying a potential lower age limit of validity for the vNRS. Our exploratory aims were to evaluate these types of validity and reliability in subgroups based on patient characteristics of sex, race/ethnicity, and primary language.

METHODS

Study Design and Setting

We conducted an observational cross-sectional study in a pediatric ED with an annual census of approximately 55,000 visits. The institutional review board approved this study with verbal informed consent.

Selection of Participants

From April 2014 to March 2016, we enrolled a convenience sample of children 4 to 17 years of age with painful and nonpainful conditions as identified by the triage nurse and confirmed by the study team by asking children themselves if they had “any pain” or “any hurt”.

Children who responded in the affirmative were considered to have a painful condition. We excluded children if they had: developmental delay or neurologic impairment; intoxication; altered mental status; a medical condition necessitating multiple painful procedures (e.g., malignancy); chronic disease associated with pain (e.g., sickle cell disease); or did not speak English or Spanish.

Patients were enrolled based on study team availability, which was primarily from 9am to midnight on weekdays. To avoid enrolling only patients who did (or did not) appear to have a strong understanding of the pain scales, we completed enrollment for every patient who was approached, identified as being eligible, and consented to participate in the study.

Measurements

The vNRS was administered by asking, “On a scale from zero to ten, where zero means no pain and ten means the most or worst pain, how much pain do you have right now?” The interaction was verbal, using no materials or equipment. We also asked children to self-report their pain intensity using the Faces Pain Scale – Revised (FPS-R) to determine convergent validity of the vNRS. The FPS-R is a self-report pain measure with strong validity and reliability in 4- to 17-year-olds (Figure 1).^{22–24} Each child was shown the faces and standardized instructions were read in English or Spanish (www.iasp-pain.org/FPSR). For both the vNRS and FPS-R, the word “hurt” was used interchangeably with “pain”, depending on what seemed most understandable for each child. Children were documented as not understanding the vNRS or FPS-R if they did not respond, or responded with a non-numerical response or a number outside the 0–10 range, when asked twice.

Procedures

Study team members were trained to collect and record data in a standardized fashion, using the same procedures and standardized data collection forms for all patients. A study team member prospectively collected data that included demographics such as age, sex, race/ethnicity, and primary language.

Two serial assessments of pain intensity were performed for each child: first on the vNRS, then on the FPS-R. For children with painful conditions, an analgesic was administered at the treating physician’s discretion after the first assessment, and the child reassessed 30–60 minutes later. For children with nonpainful conditions, or children with painful conditions

who did not receive an analgesic, the second pain assessment was performed 30–60 minutes after the first. During the second assessment, if the child had received an analgesic, they were asked: “Is your pain much less, a little less, about the same, a little worse, or much worse compared to before you got your medicine?” If the child did not receive an analgesic, they were first asked the same question but using the phrase “...compared to how you felt about [number of minutes since last assessment] ago?” Where this question was not understood, we used, “...compared to how you felt the last time I asked you how much pain/hurt you had?”

Outcomes and Analyses

Children identified as not understanding the vNRS were removed from analyses and described separately. For the remaining children, we evaluated the validity of the vNRS based on the typical criteria of convergent validity, known-groups validity, and responsivity. We also assessed the test-retest reliability, which describes the overall consistency of a measure under similar conditions and across time.

Convergent validity refers to the degree to which two different scales that are purported to measure the same construct (e.g. pain) produce similar results. We assessed convergent validity in patients who understood both the vNRS and FPS-R in two ways: First, we determined the Pearson correlation coefficient between vNRS and FPS-R scores; and second, we determined the simple agreement between these scores. The strength of correlation was described as: 0–0.19 = “very weak”; 0.20–0.39 = “weak”; 0.40–0.59 = “moderate”; 0.60–0.79 = “strong”; and 0.80–1.0 = “very strong”.²⁵ Correlation was also evaluated in subgroups based on age after adjusting for anchor bias, a tendency for young children (especially those <5 years old) to select the extremes on scales.^{15,24,26} This adjustment involved recalculating Pearson correlation coefficients after removing children who scored 10/10 on both the vNRS and FPS-R for the same assessment. We evaluated simple agreement by determining the proportion of children who had a difference between the first vNRS and FPS-R scores of less than 2 points. We determined a priori that agreement was strong if at least 80% of scores fell within 20% of scale range (i.e., $\pm 2/10$).^{24,27}

Known-groups validity is demonstrated when a measure can discriminate between groups of individuals with versus without a particular known trait. We assessed known-groups validity in children with painful and non-painful conditions who were matched by year of age and sex. The expectation was that self-reported vNRS scores would discriminate children with painful conditions from those without painful conditions (i.e. nonpainful condition). We used the independent samples t-test to compare the first mean vNRS scores in children with painful conditions to those with nonpainful conditions, with the expectation that vNRS scores should be greater in children with painful conditions by at least a clinically meaningful difference (i.e. 1/10).²⁸

Responsivity to pain-producing or pain-relieving events is another way to demonstrate construct validity. We expected that vNRS scores should decrease after analgesic administration. We determined responsivity in children with painful conditions who received an analgesic and reported a change in pain that was “much less” or “a little less” by

comparing the means of the first and second (i.e. post-analgesic) vNRS scores using the paired sample t-test. We assessed children with nonpainful conditions in a similar fashion, but with the expectation that there should be no difference between first and second scores.

Reliability describes the overall consistency of a measure under similar conditions and across time. We determined test-retest reliability in children who reported that their pain was “about the same” by comparing the vNRS scores from the first and second assessments. First, we determined the Pearson correlation coefficient of the two assessments. Second, we described the absolute differences between the two scores to evaluate the degree to which pain scores were consistent from one assessment to the next. We determined a priori that reliability was strong if the Pearson correlation was >0.60 and if more than 80% of children had an absolute difference of ≤ 1 .

The subgroups of patient ages that were analyzed for convergent validity, known-groups validity, responsiveness, and reliability included: younger (4–7 years) and older (8–17 years) (age group); and 4-, 5-, 6-, and 7-year-old children (each year of age within the younger group). The subgroups of other patient characteristics analyzed included: female and male (sex); Hispanic or Latino, Black or African-American, White, Other (race/ethnicity); and English and Spanish (primary language). We also conducted an analysis of agreement and reliability based on grouped initial pain scores, with the groups representing categories of no pain, mild, moderate, and severe pain intensity (0, 1–3, 4–6, and 7–10, respectively).^{29–32}

Sample size was based on the desired precision of our estimate of convergent validity (i.e., correlation) between the vNRS and FPS-R for children in each of the older and younger age groups. We aimed to enroll at least 300 patients in each of the older and younger age groups in order to have a 95% confidence interval (CI) of no more than ± 0.02 for correlation in each of the two age groups. By enrolling at least 100 patients in each of the younger years of age (4, 5, 6, and 7 years), we would achieve a 95% CI of ± 0.04 for correlation within each younger year of age. Enrolling at least 300 patients in each of the older and younger age groups was also sufficient to achieve an $\alpha = 0.05$ and power = 0.90 for a) detecting a clinically meaningful difference in pain of 1/10 within each of the two age groups and within each younger year of age and b) for evaluating known-groups validity and responsiveness.²⁸ We did not base our sample size on patient characteristics other than age. The software used for statistical analysis was SPSS (version 24; IBM Corporation, Armonk, NY).

RESULTS

Characteristics of Study Subjects

We enrolled 760 children; 27 children were removed from analyses because they did not understand the vNRS. The number (and proportion) of children removed from analysis in each age group was 14 (14%) 4-year-olds; 7 (7%) 5-year-olds; 5 (5%) 6-year-olds; and 1 (1%) 7-year-old. All children ≥ 8 years understood the vNRS. Figure 2 shows the number of children analyzed for each type of validity and reliability. Table 1 shows the characteristics of the patients analyzed. Supplemental Table 1 shows the number of older children enrolled by year of age. The mean age (SD) and range for the total sample was 8.9 (4) years; for

children with painful conditions, it was 9.6 (4.1) years; and for children with nonpainful conditions, it was 7.9 (3.6). The difference in years of age between children with painful and nonpainful conditions was statistically significant (1.7 years, 95% CI 1.2, 2.3). The median times to pain score reassessment (IQR) for children with painful and nonpainful conditions were 37 (32, 47) minutes and 34 (30, 41) minutes, respectively. The difference in time to pain score reassessment (minutes) between children with painful and nonpainful conditions was statistically significant (3.5 minutes, 95% CI 1.5, 5.6). Supplemental Table 2 shows the types of analgesic administered to children with painful conditions based on their initial vNRS pain score.

Main Results

Convergent Validity—Pearson correlation coefficients between the vNRS and FPS-R in 731 children are shown in Table 2. The correlations for the overall sample and for the older age group were very strong. The correlations were strong in the younger age group, including all years of age except for 4-year-olds, for whom correlation was moderate. Table 3 shows the correlations in each year of age after adjusting for anchor bias. The correlations remained very strong in children 8–17 years of age; decreased but remained strong in the 6- and 7-year-olds, were moderate in the 5 year-olds, and weak in the 4-year-olds.

We evaluated agreement in 731 children and found 81.9% had strong agreement (Table 4). There was poor (i.e. not strong) agreement between VNRS and F-PSR for children in every year of younger age, those of non-Hispanic Black race/ethnicity, and children who spoke Spanish as their primary language (Table 4). After adjusting for anchor bias in subgroups based on age, agreement remained unchanged (data available upon request). When agreement was evaluated based on category of pain intensity (i.e., none, mild, moderate, or severe pain), all children who reported no pain demonstrated strong agreement. More children reporting moderate pain intensity demonstrated poor agreement than those who reported mild or severe pain intensity (Table 5).

Known-Groups Validity—We compared all available pairs of children with painful and nonpainful conditions who could be matched by year of age and sex (269 pairs, 538 children total). Pain scores in the total sample were higher in the children presenting with painful conditions (Supplemental Table 3). There was a similar difference in pain scores between the painful and nonpainful groups when we analyzed the subgroups based on age group, year of age in the younger age group, sex, race/ethnicity, and primary language (Supplemental Table 3).

Responsivity—We evaluated 217 children with painful conditions who received an analgesic and reported a change in pain of “much less” or “a little less” (78 and 139 children, respectively) at the second pain assessment, and found a difference in vNRS scores of 3.5 (95% CI 3.1, 3.9) between before versus after the analgesic was given (Supplemental Table 4). There was a similar difference identified in all subgroups. In 302 children with nonpainful conditions, there was no difference in vNRS scores between the first and second pain assessments in the total sample or any of the subgroups (data available upon request).

Reliability—We evaluated 300 children who reported their pain was “about the same” between the first and second vNRS assessments, and found the Pearson correlation coefficient between these two assessments for the total sample to be 0.87; 0.74 for the younger and 0.97 for the older age groups. The correlation coefficients for 4-year-olds, 6-year-olds, and other race/ethnicity were 0.69, 0.66, and 0.70, respectively. The correlation coefficients for the remaining subgroups were all greater than 0.82.

The median absolute differences between the first and second vNRS assessments was 0 (IQR 0, 1) in the total population and 0 (IQR 0, 1) and 0 (IQR 0, 0.5) in the younger and older age groups, respectively. The absolute difference was ≤ 1 for 89.7% of children. Table 6 shows that more than 80% of children in all subgroups based on patient characteristics had an absolute difference ≤ 1 ; only children with an initial pain score of 4 to 6 did not.

LIMITATIONS

Limitations of the study include enrolling a convenience sample rather than consecutive patients, though our sample included a diverse representation of conditions and a wide distribution of pain intensities. Although selection bias could have occurred due to convenience sampling, bias was less likely as investigators were unable to discern each child’s ability to use or understand the pain scales prior to being approached for the study. Furthermore, study team members did not enroll based on any preceding interaction, or after speaking, with the child. All patients who were approached, identified as being eligible, and consented to participate in the study, completed enrollment. The majority of children with painful conditions received non-opioid oral analgesics or did not receive any analgesics, which could be interpreted to reflect a population with lower pain severity. However, in light of the large proportion of children who reported pain scores that typically represent severe pain intensity ($\geq 7/10$), it is likely that this pattern of analgesic administration is reflective of the practice of using non-opioid analgesics more frequently for children compared to adults or the undertreatment of pain in children in general.^{33–35} We did not randomize the order in which the two pain scales were presented when assessing convergent validity, which could have introduced order effects in which the first score influenced or constrained the second. However, we elected to consistently administer the vNRS first in order to avoid such influence or constraints on the vNRS scores, since in clinical practice, only one scale would be used and therefore would not be subject to order effects. Pain scores were reported to a member of the study team and not recorded in a blinded fashion, so subjects could have been subtly influenced to respond in ways that were consistent with the study team member’s expectations. We did not plan our sample size to account for the analyses performed involving comparisons within subgroups of sex, race/ethnicity, and primary language. However, we were still able to observe strong convergent validity, known-groups validity, responsiveness, and reliability in these subgroups based on our a priori definitions. A large proportion of our sample was Hispanic or Latino, which limited our ability to enroll sufficient numbers of children of other race/ethnicities for more definitive analyses. The use of a priori cut-offs as our criterion for validity and reliability may have resulted in the loss of variance due to recoding a continuous variable as binary or categorical. As a result, there may have been some potentially meaningful fine gradations within categories that were lost. For example, there might be differences between pain scores of 7, 8, 9, and 10 that may not

have been able to be identified because they were grouped and analyzed as the single category of “severe”. Finally, our evaluation of reliability could have been affected by asking children if their pain was “about the same” rather than “exactly the same” or “the same”, potentially introducing imprecision into our estimates.

DISCUSSION

In this cross-sectional study, we found that convergent validity, known-groups validity, and responsivity of the vNRS was strong in children 8 to 17 years of age, as well as in children 6 and 7 years of age. However, the degree of convergent validity was variable in children 4 and 5 years of age. Validity was strong in subgroups based on patient characteristics other than age, except for poor agreement in certain subgroups of race/ethnicity and primary language. Reliability was strong in all subgroups, including by age.

The minimum age at which the vNRS demonstrates strong validity for both clinical and research use has not been sufficiently studied.^{11,36} The ability to use the vNRS is related to age, insofar as age is a reflection of children’s appreciation that numbers are related to quantity, and their possession of necessary abilities such as seriation, quantity estimation, and classification.^{10,14,37–39} The minimum age at which the vNRS should be used has been frequently posited as 8 years, a cutoff adopted in multiple prior studies.^{10,12,37,40,41} Study of vNRS scores in children less than 8 years of age has been limited, with some evidence obtained in non-clinical (i.e. schools) and post-operative settings to suggest that children as young as 6 years old may be able to use the vNRS.^{12–14} Our results corroborate these findings in a clinical setting with patients experiencing acute pain by demonstrating strong Pearson correlations, known-groups validity, and responsivity in 6- and 7-year-olds. We also evaluated the validity and reliability of vNRS scores in 4- and 5-year-olds, which have not previously been reported. Although known-groups validity and responsivity were strong in this youngest cohort, Pearson correlations were not strong when anchor bias was taken into account. Our findings suggest that the vNRS may be used in children 6 and 7 years of age to assess pain intensity, change in pain, and presence of pain. However, the vNRS may be limited in children 4 and 5 years to simply identifying the presence of pain and that a change in pain occurred, but not for indicating the magnitude of pain or the degree to which a change in pain occurred.

Agreement between the vNRS and FPS-R observed in children younger than 8 years of age was poor. This finding, when considered in isolation, does not necessarily mean that vNRS scores have poor convergent validity in this age group. Correlation between two scales is commonly used as the sole criterion for determining convergent validity in many studies.^{10,12,13,22,40,42–46} In children 6 and 7 years old, correlation was strong but agreement was poor between vNRS and FPS-R scores. These findings imply that the two scales measure the same construct (i.e. pain), but do not support interchangeability between the two scales. This means that vNRS pain scores obtained, or pain management goals expressed, in 6- and 7-year-olds have strong validity but should not be transposed to or compared with norms or scores based on other scales.^{10,47} In addition, the variability in convergent validity observed over the spectrum of age demonstrates how children acquire skills and abilities sequentially: vNRS and FPS-R scores have both poor correlation and agreement in 4- and 5-year-olds;

strong correlation but poor agreement in 6- and 7-year-olds; and eventually both strong correlation and agreement in children 8 years and older.

When examining validity in subgroups based on patient characteristics, we found that children of Black or African-American race/ethnicity, and those whose primary language was Spanish, demonstrated poor agreement despite demonstrating strong convergent validity, known-groups validity, and responsivity. Although these findings are exploratory, they may be a reflection of how race, ethnicity, and language are associated with the perception and report of pain, and how these differences may become manifest when comparing scores obtained using different self-report measures of pain.^{19,20} Given that the other measures of validity were strong, it is reasonable to still recommend the use of the vNRS in these subgroups for within-patient comparisons over time, but interchangeability between the vNRS and FPS-R may be limited in these specific subgroups.

The validity and reliability we describe in children younger than 8 only apply to those who appeared to understand the vNRS, with the proportion of children who were excluded due to not understanding the vNRS decreasing with age. We applied a rudimentary criterion for inclusion and did not formally screen children to see if they had the appropriate numerical skills presumed to be prerequisite for use of the vNRS, as has been done in prior studies.^{13,14} Therefore, it is possible that we included children in analyses that did not have the appropriate numerical skills to use the vNRS. However, we intentionally chose this method of screening patients because we felt that it best represented a typical clinical encounter in which providers often do not have the skills, time or resources to carry out a formal screening process before conducting a pain assessment. Despite this limitation, there was still strong known-groups validity and responsivity demonstrated in children 4 to 7 years of age, and strong convergent validity at 6 and 7 years of age. Therefore, in children less than 8 years of age who passed our rudimentary screen of comprehension, the validity and reliability of the vNRS appear sufficiently strong for use in children 6 and 7 years of age, but limited in children 4 and 5 years of age as discussed above.

We demonstrated a strong degree of test-retest reliability across all subgroups of patient characteristics, but not in children whose initial pain score was 4–6/10. Poor reliability in scores representing moderate pain intensity may be due to opportunities for larger variations in scores in the middle compared to the ends of the scale, or greater variability between individuals in their interpretation of moderate-intensity pain compared to the more intuitive and straightforward constructs of no pain or severe pain. This finding warrants caution when interpreting reported changes in pain score, as well as the incorporation of other subjective measures of change in pain (e.g., satisfaction with analgesic administered, desire for additional analgesics), in children who report pain scores usually associated with moderate pain intensity.

In summary, the vNRS demonstrated strong convergent validity, known-groups validity, responsivity, and reliability in children 6 to 17 years of age. Convergent validity was not strong in 4- and 5-year-olds, but other measures of validity and reliability were strong in these ages. Our findings support the use of the vNRS in children 6 years and older, but not for use in 4- and 5-year-olds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Leonor Suarez, Allison Hyland, Jeffrey Sung, and Caitlin Oldenkamp for their assistance with patient enrollment; and Stephanie Raine Lee for her review of the manuscript.

Grants: This publication was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1 TR000040.

References

1. Johnston CC, Gagnon AJ, Fullerton L, Common C, Ladores M, Forlini S. One-week survey of pain intensity on admission to and discharge from the emergency department: a pilot study. *J Emerg Med.* 1998; 16(3):377–382. [PubMed: 9610963]
2. Tanabe P, Buschmann M. A prospective study of ED pain management practices and the patient's perspective. *J Emerg Nurs.* 1999; 25(3):171–177. [PubMed: 10346837]
3. Cordell WH, Keene KK, Giles BK, Jones JB, Jones JH, Brizendine EJ. The high prevalence of pain in emergency medical care. *Am J Emerg Med.* 2002; 20(3):165–169. [PubMed: 11992334]
4. Bijur PE, Latimer CT, Gallagher EJ. Validation of a verbally administered numerical rating scale of acute pain for use in the emergency department. *Acad Emerg Med.* 2003; 10(4):390–392. [PubMed: 12670856]
5. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2005; 113(1–2):9–19. DOI: 10.1016/j.pain.2004.09.012 [PubMed: 15621359]
6. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* 2011; 152(10):2399–2404. DOI: 10.1016/j.pain.2011.07.005 [PubMed: 21856077]
7. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage.* 2011; 41(6):1073–1093. DOI: 10.1016/j.jpainsymman.2010.08.016 [PubMed: 21621130]
8. Chien C-W, Bagraith KS, Khan A, Deen M, Strong J. Comparative responsiveness of verbal and numerical rating scales to measure pain intensity in patients with chronic pain. *J Pain.* 2013; 14(12):1653–1662. DOI: 10.1016/j.jpain.2013.08.006 [PubMed: 24290445]
9. Bahreini M, Jalili M, Moradi-Lakeh M. A comparison of three self-report pain scales in adults with acute pain. *J Emerg Med.* 2015; 48(1):10–18. DOI: 10.1016/j.jemermed.2014.07.039 [PubMed: 25271179]
10. Bailey B, Daoust R, Doyon-Trottier E, Dauphin-Pierre S, Gravel J. Validation and properties of the verbal numeric scale in children with acute pain. *Pain.* 2010; 149(2):216–221. DOI: 10.1016/j.pain.2009.12.008 [PubMed: 20188471]
11. Castarlenas E, Jensen MP, von Baeyer CL, Miró J. Psychometric Properties of the Numerical Rating Scale to Assess Self-Reported Pain Intensity in Children and Adolescents: A Systematic Review. *Clin J Pain.* Aug.2016 doi: 10.1097/AJP.0000000000000406
12. Miró J, Castarlenas E, Huguet A. Evidence for the use of a numerical rating scale to assess the intensity of pediatric pain. *Eur J Pain.* 2009; 13(10):1089–1095. DOI: 10.1016/j.ejpain.2009.07.002 [PubMed: 19726211]
13. Sánchez-Rodríguez E, Miró J, Castarlenas E. A comparison of four self-report scales of pain intensity in 6- to 8-year-old children. *Pain.* 2012; 153(8):1715–1719. DOI: 10.1016/j.pain.2012.05.007 [PubMed: 22703691]
14. Castarlenas E, Miró J, Sánchez-Rodríguez E. Is the verbal numerical rating scale a valid tool for assessing pain intensity in children below 8 years of age? *J Pain.* 2013; 14(3):297–304. DOI: 10.1016/j.jpain.2012.12.004 [PubMed: 23380267]

15. Chambers CT, Johnston C. Developmental differences in children's use of rating scales. *J Pediatr Psychol.* 2002; 27(1):27–36. [PubMed: 11726677]
16. Fowler-Kerry S, Lander J. Assessment of sex differences in children's and adolescents' self-reported pain from venipuncture. *J Pediatr Psychol.* 1991; 16(6):783–793. [PubMed: 1798014]
17. Stanford EA, Chambers CT, Craig KD. A normative analysis of the development of pain-related vocabulary in children. *Pain.* 2005; 114(1–2):278–284. DOI: 10.1016/j.pain.2004.12.029 [PubMed: 15733654]
18. Stanford EA, Chambers CT, Craig KD. The role of developmental factors in predicting young children's use of a self-report scale for pain. *Pain.* 2006; 120(1–2):16–23. DOI: 10.1016/j.pain.2005.10.004 [PubMed: 16359800]
19. Fortier MA, Anderson CT, Kain ZN. Ethnicity matters in the assessment and treatment of children's pain. *Pediatrics.* 2009; 124(1):378–380. DOI: 10.1542/peds.2008-3332 [PubMed: 19564322]
20. Ortega HW, Velden HV, Lin C-W, Reid S. Ethnicity and reported pain scores among children with long-bone fractures requiring emergency care. *Pediatr Emerg Care.* 2012; 28(11):1146–1149. DOI: 10.1097/PEC.0b013e31827134f6 [PubMed: 23114236]
21. Schmitz A-K, Vierhaus M, Lohaus A. Pain tolerance in children and adolescents: sex differences and psychosocial influences on pain threshold and endurance. *Eur J Pain.* 2013; 17(1):124–131. DOI: 10.1002/j.1532-2149.2012.00169.x [PubMed: 22715044]
22. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain.* 2001; 93(2):173–183. [PubMed: 11427329]
23. Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A systematic review of faces scales for the self-report of pain intensity in children. *Pediatrics.* 2010; 126(5):e1168–e1198. DOI: 10.1542/peds.2010-1609 [PubMed: 20921070]
24. Tsze DS, von Baeyer CL, Bulloch B, Dayan PS. Validation of Self-Report Pain Scales in Children. *Pediatrics.* 2013; 132(4):e971–e979. DOI: 10.1542/peds.2013-1509 [PubMed: 23999954]
25. Evans, JD. *Straightforward Statistics for the Behavioral Sciences.* Pacific Grove, CA: Brooks/Cole; 1996.
26. von Baeyer CL, Forsyth SJ, Stanford EA, Watson M, Chambers CT. Response biases in preschool children's ratings of pain in hypothetical situations. *Eur J Pain.* 2009; 13(2):209–213. DOI: 10.1016/j.ejpain.2008.03.017 [PubMed: 18479952]
27. von Baeyer CL. Reported lack of agreement between self-report pain scores in children may be due to a too strict criterion for agreement. *Pain.* 2012; 153(10):2152–3. –authorreply2153–4. DOI: 10.1016/j.pain.2012.06.029 [PubMed: 22858170]
28. Voepel-Lewis T, Burke CN, Jeffreys N, Malviya S, Tait AR. Do 0–10 numeric rating scores translate into clinically meaningful pain measures for children? *Anesth Analg.* 2011; 112(2):415–421. DOI: 10.1213/ANE.0b013e318203f495 [PubMed: 21127278]
29. Breivik H, Borchgrevink PC, Allen SM, et al. Assessment of pain. *Br J Anaesth.* 2008; 101(1):17–24. DOI: 10.1093/bja/aen103 [PubMed: 18487245]
30. Jensen MP, Smith DG, Ehde DM, Robinsin LR. Pain site and the effects of amputation pain: further clarification of the meaning of mild, moderate, and severe pain. *Pain.* 2001; 91(3):317–322. [PubMed: 11275389]
31. Jones KR, Vojir CP, Hutt E, Fink R. Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents. *J Rehabil Res Dev.* 2007; 44(2):305–314. [PubMed: 17551881]
32. Hoffman DL, Sadosky A, Dukes EM, Alvir J. How do changes in pain severity levels correspond to changes in health status and function in patients with painful diabetic peripheral neuropathy? *Pain.* 2010; 149(2):194–201. [PubMed: 20303665]
33. Schechter NL, Allen DA, Hanson K. Status of pediatric pain control: a comparison of hospital analgesic usage in children and adults. *Pediatrics.* 1986; 77(1):11–15. [PubMed: 2415909]
34. Eland, JM., Anderson, JE. The experience of pain in children. In: Jacox, A., editor. *Pain: a Source Book for Nurses and Other Health Professionals.* Boston: Pain: A Source Book for Nurses and Other Health Professionals; 1977. p. 453–473.

35. Tsze DS, Asnis LM, Merchant RC, Amanullah S, Linakis JG. Increasing computed tomography use for patients with appendicitis and discrepancies in pain management between adults and children: an analysis of the NHAMCS. *Ann Emerg Med*. 2012; 59(5):395–403. DOI: 10.1016/j.annemergmed.2011.06.010 [PubMed: 21802777]
36. von Baeyer CL. Numerical rating scale for self-report of pain intensity in children and adolescents: recent progress and further questions. *Eur J Pain*. 2009; 13(10):1005–1007. DOI: 10.1016/j.ejpain.2009.08.006 [PubMed: 19766028]
37. von Baeyer CL. Children's self-report of pain intensity: what we know, where we are headed. *Pain Res Manag*. 2009; 14(1):39–45. [PubMed: 19262915]
38. Champion, GD., Goodenough, B., von Baeyer, CL., Thomas, W. Measurement of pain by self-report. In: Finley, GA., McGrath, PJ., editors. *Measurement of Pain in Infants and Children*. Vol. 10. Seattle: IASP Press; 1998. p. 123-160.
39. Chan JY-C, von Baeyer CL. Cognitive developmental influences on the ability of preschool-aged children to self-report their pain intensity. *Pain*. 2016; 157(5):997–1001. DOI: 10.1097/j.pain.0000000000000476 [PubMed: 26716994]
40. Pagé MG, Katz J, Stinson J, Isaac L, Martin-Pichora AL, Campbell F. Validation of the numerical rating scale for pain intensity and unpleasantness in pediatric acute postoperative pain: sensitivity to change over time. *J Pain*. 2012; 13(4):359–369. DOI: 10.1016/j.jpain.2011.12.010 [PubMed: 22424915]
41. Connelly M, Neville K. Comparative prospective evaluation of the responsiveness of single-item pediatric pain-intensity self-report scales and their uniqueness from negative affect in a hospital setting. *J Pain*. 2010; 11(12):1451–1460. DOI: 10.1016/j.jpain.2010.04.011 [PubMed: 20646966]
42. DeLoach LJ, Higgins MS, Caplan AB, Stiff JL. The visual analog scale in the immediate postoperative period: intrasubject variability and correlation with a numeric scale. *Anesth Analg*. 1998; 86(1):102–106. [PubMed: 9428860]
43. Bulloch B, Tenenbein M. Validation of 2 pain scales for use in the pediatric emergency department. *Pediatrics*. 2002; 110(3):e33. [PubMed: 12205283]
44. von Baeyer CL, Spagrud LJ, McCormick JC, Choo E, Neville K, Connelly MA. Three new datasets supporting use of the Numerical Rating Scale (NRS-11) for children's self-reports of pain intensity. *Pain*. 2009; 143(3):223–227. DOI: 10.1016/j.pain.2009.03.002 [PubMed: 19359097]
45. Miró J, Huguet A. Evaluation of reliability, validity, and preference for a pediatric pain intensity scale: the Catalan version of the faces pain scale–revised. *Pain*. 2004; 111(1–2):59–64. DOI: 10.1016/j.pain.2004.05.023 [PubMed: 15327809]
46. Myrvik MP, Drendel AL, Brandow AM, Yan K, Hoffmann RG, Panepinto JA. A Comparison of Pain Assessment Measures in Pediatric Sickle Cell Disease: Visual Analog Scale Versus Numeric Rating Scale. *J Pediatr Hematol Oncol*. 2015; 37(3):190–194. DOI: 10.1097/MPH.0000000000000306 [PubMed: 25575295]
47. Ruskin D, Lalloo C, Amaria K, et al. Assessing pain intensity in children with chronic pain: convergent and discriminant validity of the 0 to 10 numerical rating scale in clinical practice. *Pain Res Manag*. 2014; 19(3):141–148. [PubMed: 24712019]

Figure 1.

The Faces Pain Scale – Revised (FPS-R). Each face represents an increasing degree of pain moving from left to right, scored 0-2-4-6-8-10. Hicks CL et al. *Pain*.2001;93:176.

www.iasppain.org/FPSR. ©2001, International Association for the Study of Pain®.

Reproduced with permission.

Figure 2.
Patients analyzed for each type of validity and reliability.

Table 1

Patient characteristics

	Total Sample Analyzed, n = 733 n (%)	Children with Painful Conditions Analyzed, n = 431 n (%)	Children with Nonpainful Conditions Analyzed, n = 302 n (%)
Age group			
Younger, 4–7 years old	373 (50.9)	182 (42.2)	191 (63.2)
Older, 8–17 years old	360 (49.1)	249 (57.8)	111 (36.8)
Younger age			
4 years old	86 (11.7)	42 (9.7)	44 (14.6)
5 years old	93 (12.7)	45 (10.4)	48 (15.9)
6 years old	95 (13.0)	46 (10.7)	49 (16.2)
7 years old	99 (13.5)	49 (11.4)	50 (16.5)
Sex			
Female	356 (51.4)	209 (48.5)	147 (48.7)
Male	377 (48.6)	222 (51.5)	155 (51.3)
Race/Ethnicity			
Hispanic or Latino	586 (80.0)	341 (79.1)	245 (81.1)
Black or African-American	99 (13.5)	59 (13.7)	40 (13.2)
White	33 (4.5)	23 (5.3)	10 (3.3)
Other ¹	15 (2.0)	8 (1.9)	7 (2.4)
Primary language			
English	662 (90.3)	391 (90.7)	271 (89.7)
Spanish	71 (9.7)	40 (9.3)	31 (10.3)
Painful condition			
Soft tissue injury		111 (25.8)	
Abdominal pain		92 (21.3)	
Ear/throat pain		58 (13.4)	
Headache		49 (11.4)	
Fracture		25 (5.8)	
Back pain		15 (3.5)	
Chest pain		13 (3.0)	
Abscess		13 (3.0)	
Laceration		9 (2.1)	
Other		46 (10.7)	
Analgesic administered			
Ibuprofen		152 (35.3)	
Parenteral opioid		59 (13.7)	
Acetaminophen		49 (11.4)	

	Total Sample Analyzed, n = 733 n (%)	Children with Painful Conditions Analyzed, n = 431 n (%)	Children with Nonpainful Conditions Analyzed, n = 302 n (%)
Ketorolac		13 (3.0)	
Oral opioid		4 (0.9)	
Other		13 (3.0)	
None		141 (32.7)	
Nonpainful condition			
Fever			46 (15.2)
Rash			36 (11.9)
Cough			37 (12.3)
Vomiting and/or diarrhea			33 (10.9)
Shortness of breath, wheezing			26 (8.6)
Soft tissue injury, nonpainful			27 (8.9)
Dizziness or syncope			7 (2.3)
Other			90 (29.8)
Initial pain score: vNRS			
0	209 (28.5)	9 (2.1)	200 (66.2)
1–3	139 (18.9)	77 (17.9)	62 (20.5)
4–6	155 (21.2)	134 (31.1)	21 (7)
7–10	230 (31.4)	211 (48.9)	19 (6.3)

¹Other includes American Indian or Alaska Native, Asian, More than one, and Don't Know.

Table 2

Pearson correlation coefficients of vNRS and FPS-R scores of first assessment (Convergent validity: Correlation)

	n	Pearson Correlation Coefficient (95% CI)
Total	731	0.80 (0.77, 0.82)
Age group		
Younger, 4–7 years	371	0.68 (0.62, 0.73)
Older, 8–17 years	360	0.92 (0.90, 0.93)
Age, years		
4	86	0.59 (0.43, 0.71)
5	93	0.71 (0.59, 0.80)
6	94	0.73 (0.62, 0.81)
7	98	0.70 (0.58, 0.79)
8	30	0.92 (0.84, 0.96)
9	42	0.87 (0.77, 0.93)
10	33	0.90 (0.81, 0.95)
11	42	0.93 (0.87, 0.96)
12	40	0.93 (0.87, 0.96)
13	35	0.90 (0.81, 0.95)
14	42	0.89 (0.80, 0.94)
15	39	0.96 (0.93, 0.98)
16	33	0.96 (0.92, 0.98)
17	24	0.93 (0.84, 0.97)
Sex		
Female	356	0.81 (0.77, 0.84)
Male	375	0.80 (0.76, 0.83)
Race/Ethnicity		
Hispanic or Latino	584	0.81 (0.78, 0.84)
Black or African American	99	0.76 (0.66, 0.83)
White	33	0.93 (0.86, 0.97)
Other ^I	15	0.86 (0.62, 0.95)
Primary language		
English	661	0.81 (0.78, 0.83)
Spanish	70	0.78 (0.67, 0.88)

^I Other includes American Indian or Alaska Native, Asian, More than one, and Don't Know.

Table 3

Comparison of Pearson correlation coefficients of vNRS and FPS-R in the younger age group and individual years of age between the total sample and anchor bias-adjusted population (Convergent validity: Correlation)[/]

	Total Sample		Anchor Bias-Adjusted Population	
	n	Pearson Correlation Coefficient (95% CI)	n	Pearson Correlation Coefficient (95% CI)
Younger age group				
4 years old	86	0.59 (0.43, 0.71)	75	0.27 (0.05, 0.47)
5 years old	93	0.71 (0.59, 0.80)	86	0.58 (0.42, 0.71)
6 years old	94	0.73 (0.62, 0.81)	85	0.69 (0.56, 0.79)
7 years old	98	0.70 (0.58, 0.79)	94	0.62 (0.48, 0.73)
Older age group				
8 years old	30	0.92 (0.84, 0.96)	27	0.89 (0.77, 0.95)
9 years old	42	0.87 (0.77, 0.93)	41	0.85 (0.74, 0.92)
10 years old	33	0.90 (0.81, 0.95)	31	0.89 (0.78, 0.95)
11 years old	42	0.93 (0.87, 0.96)	37	0.91 (0.83, 0.95)
12 years old	40	0.93 (0.87, 0.96)	33	0.90 (0.81, 0.95)
13 years old	35	0.90 (0.81, 0.95)	34	0.89 (0.79, 0.94)
14 years old	42	0.89 (0.80, 0.94)	40	0.88 (0.78, 0.94)
15 years old	39	0.96 (0.93, 0.98)	37	0.96 (0.92, 0.98)
16 years old	33	0.96 (0.92, 0.98)	32	0.96 (0.92, 0.98)
17 years old	24	0.93 (0.84, 0.97)	20	0.91 (0.78, 0.96)

[/] Anchor bias-adjusted population derived by removing children who scored 10/10 on both the vNRS and FPS-R for the same assessment.

Table 4

Percentage of children with a difference of less than 2 out of 10 between vNRS and FPS-R scores of first assessment (Convergent validity: Agreement)¹

	n	Percentage of Children (95% CI)
Total	731	81.9 (79, 84.7)
Age group		
Younger, 4–7 years	371	74.9 (70.2, 79.3)
Older, 8–17 years	360	89.2 (85.5, 92.2)
Age, years		
4	86	70.9 (60.1, 80.2)
5	93	77.4 (67.7, 85.4)
6	94	73.4 (63.3, 82)
7	98	77.6 (68, 85.4)
8	30	90 (73.5, 97.9)
9	42	92.9 (80.5, 98.5)
10	33	84.8 (68.1, 94.9)
11	42	88.1 (74.4, 96)
12	40	87.5 (73.2, 95.8)
13	35	85.7 (69.7, 95.2)
14	42	88.1 (74.4, 96)
15	39	94.9 (82.7, 99.4)
16	33	93.9 (79.8, 99.3)
17	24	83.3 (62.6, 95.3)
Sex		
Female	356	80.9 (76.4, 84.9)
Male	375	82.9 (78.7, 86.6)
Race/Ethnicity		
Hispanic or Latino	584	82.5 (79.2, 85.5)
Black or African-American	99	72.7 (62.9, 81.2)
White	33	97 (84.2, 99.9)
Other ²	15	86.7 (59.5, 98.3)
Primary language		
English	661	82.8 (79.7, 85.6)
Spanish	70	74.3 (62.4, 84)

¹Only children who understood both the vNRS and FPS-R were evaluated for agreement (n=731). Subgroups that have a percentage ≥ 80% have strong agreement.

²Other includes American Indian or Alaska Native, Asian, More than one, and Don't Know.

Table 5

Percentage of children with a difference of less than 2 out of 10 between first vNRS and FPS-R scores based on category of pain intensity (Convergent validity: Agreement) /

	No or Mild Pain		Moderate Pain		Severe Pain	
	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)
Total	382	88.2 (84.6, 91.3)	170	68.8 (61.3, 75.7)	179	81 (74.5, 86.5)
Age group						
Younger, 4–7 years	235	85.5 (80.4, 89.8)	75	46.7 (35.1, 58.6)	61	68.9 (55.7, 80.1)
Older, 8–17 years	147	92.5 (87, 96.2)	79	86.3 (77.7, 92.5)	145	87.3 (79.9, 92.7)
Age, years						
4	54	79.6 (66.5, 89.4)	16	31.3 (11, 58.7)	16	81.3 (54.4, 96)
5	59	89.8 (79.2, 96.2)	19	47.4 (24.4, 71.1)	15	66.7 (38.4, 88.2)
6	59	88.1 (77.1, 95.1)	18	38.9 (17.3, 64.3)	17	58.8 (32.9, 81.6)
7	63	84.1 (72.7, 92.1)	22	63.6 (40.7, 82.8)	13	69.2 (38.6, 90.9)
8	16	93.8 (69.8, 99.8)	7	71.4 (29, 96.3)	7	100 (59, 100)
9	25	92 (74, 99)	15	93.3 (68.1, 99.8)	2	100 (15.8, 100)
10	21	85.7 (63.7, 97)	4	100 (39.8, 100)	8	75 (34.9, 96.8)
11	18	94.4 (72.7, 99.9)	9	66.7 (29.9, 92.5)	15	93.3 (68.1, 99.8)
12	17	88.2 (63.6, 98.5)	10	80 (44.4, 97.5)	13	92.3 (64, 99.8)
13	15	86.7 (59.5, 98.3)	8	87.5 (47.3, 99.7)	12	83.3 (51.6, 97.9)
14	15	86.7 (59.5, 98.3)	16	93.8 (69.8, 99.8)	11	81.8 (48.2, 97.7)
15	12	100 (73.5, 100)	15	86.7 (59.5, 98.3)	12	100 (73.5, 100)
16	14	93.9 (66.1, 99.8)	7	85.7 (42.1, 99.6)	12	100 (73.5, 100)
17	10	70 (34.8, 93.3)	4	75 (19.4, 99.4)	10	100 (69.2, 100)
Sex						
Female	179	88.3 (82.6, 92.6)	87	63.2 (52.2, 73.3)	90	83.3 (74, 90.4)
Male	203	88.2 (82.9, 92.3)	83	74.7 (64, 83.6)	89	78.7 (68.7, 86.6)
Race/Ethnicity						

	No or Mild Pain		Moderate Pain		Severe Pain	
	n	% (95% CI)	n	% (95% CI)	n	% (95% CI)
Hispanic or Latino	304	89.5 (85.5, 92.7)	136	66.9 (58.3, 74.4)	144	82.6 (75.4, 88.4)
Black or African-American	53	81.1 (68, 90.6)	20	60 (36.1, 80.9)	26	65.4 (44.3, 82.8)
White	16	93.8 (69.8, 99.8)	12	100 (73.5, 100)	5	100 (47.8, 100)
Other ²	9	77.8 (40, 97.2)	2	100 (15.8, 100)	4	100 (39.8, 100)
Primary language						
English	348	89.1 (85.3, 92.2)	153	69.3 (31.3, 76.5)	160	81.9 (75, 87.5)
Spanish	34	79.4 (62.1, 91.3)	17	64.7 (38.3, 85.8)	19	73.7 (48.8, 90.9)

¹ Categories of pain severity: No pain = 0; mild pain = 1–3; moderate pain = 4–6; severe pain = 7–10. Pain intensity score obtained by averaging vNRS and FPS-R scores. Groups that have a percentage 80% have strong agreement.

² Other includes American Indian or Alaska Native, Asian, More than one, and Don't Know.

Table 6

Absolute difference in vNRS score in children who reported their pain was “about the same” from the first to second assessment (Reliability)

	Number of children with the following absolute difference in vNRS score between first and second assessment, n (%) ^f				
	n	0	1	2	>2
Total	300	217(72.3%)	269 (89.7%)	281 (93.7%)	19 (6.3%)
Age group					
Younger, 4–7 years old	153	109 (71.2%)	134 (87.6%)	139 (90.8%)	14 (9.2%)
Older, 8–17 years old	147	108 (73.5%)	135 (91.8%)	142 (96.6%)	5 (3.4%)
Age, years					
4	41	31 (75.6%)	34 (82.9%)	36 (87.8%)	5 (12.2%)
5	31	22 (71.0%)	28 (90.3%)	29 (93.5%)	2 (6.5%)
6	44	31 (70.5%)	38 (86.4%)	40 (90.9%)	4 (9.1%)
7	37	25 (67.6%)	34 (91.9%)	35 (94.6%)	2 (5.4%)
8	12	9 (75%)	11 (91.7%)	11 (91.7%)	1 (8.3%)
9	17	9 (52.9%)	13 (76.5%)	16 (94.1%)	1 (5.9%)
10	10	9 (90%)	10(100%)	10 (100%)	0
11	15	10 (66.7%)	13 (86.7%)	14 (93.3%)	1 (6.7%)
12	14	10 (71.4%)	13 (92.9%)	14 (100%)	0
13	17	11 (64.7%)	16 (94.1%)	17 (100%)	0
14	16	12 (75%)	14 (87.5%)	15 (93.8%)	1 (6.2%)
15	20	16 (80%)	19 (95%)	19 (95%)	1 (5%)
16	14	13 (92.9%)	14 (100%)	14 (100%)	0
17	12	9 (75%)	12(100%)	12 (100%)	0
Sex					
Female	142	106 (74.6%)	126 (88.7%)	135 (95.1%)	7 (4.9%)
Male	158	111 (70.3%)	143 (90.5%)	146 (92.4%)	12 (7.6%)
Race/Ethnicity					
Hispanic or Latino	236	172 (72.9%)	210 (89.0%)	221 (93.6%)	15 (6.4%)

	n	Number of children with the following absolute difference in vNRS score between first and second assessment, n (%) ¹			
		0	1	2	>2
Black or African American	42	30 (71.4%)	39 (92.9%)	40 (95.2%)	2 (4.8%)
White	13	9 (69.2%)	12 (92.3%)	12 (92.3%)	1 (7.7%)
Other ²	9	6 (66.7%)	8 (88.9%)	8 (88.9%)	1 (11.1%)
Primary language					
English	270	198 (73.3%)	243 (90.0%)	254 (94.1%)	16(5.9%)
Spanish	30	19 (63.3%)	26 (86.7%)	27 (90.0%)	3 (10.0%)
Initial pain score					
0	143	133 (93.0%)	138 (96.5%)	138 (96.5%)	5 (3.5%)
1–3	69	37 (53.6%)	61 (88.4%)	65 (94.2%)	4 (5.8%)
4–6	42	20 (47.6%)	32 (76.2%)	37 (88.1%)	5 (11.9%)
7–10	46	27 (58.7%)	38 (82.6%)	41 (89.1%)	5 (10.9%)

¹ The children who have an absolute difference in vNRS score of 1 may have a difference of any value between, and including, 0 to 1. Children with an absolute difference of 2 may have a difference of any value between, and including, 0 to 2. Children with an absolute difference of > 2 may have a difference of any value greater than 2.

² Other includes American Indian or Alaska Native, Asian, More than one, and Don't Know.