OPEN

Psychometric Properties of the Numerical Rating Scale to Assess Self-Reported Pain Intensity in Children and Adolescents

A Systematic Review

Elena Castarlenas, PhD,†‡ Mark P. Jensen, PhD,§ Carl L. von Baeyer, PhD,‖¶ and Jordi Miró, PhD*†‡#*

Objectives: The Numerical Rating Scale-11 (NRS-11) is one of the most widely used scales to assess self-reported pain intensity in children, despite the limited information on its psychometric properties for assessing pain in pediatric populations. Recently, there has been an increase in published findings regarding the strengths and weaknesses of the NRS-11 as a measure of pain in youths. The purpose of this study was to review this research and summarize what is known regarding the reliability and validity of the NRS-11 as a self-report measure of pediatric pain intensity.

Methods: A literature search was conducted using PubMed, PsycINFO, CINAHL, and the Psychology and Behavioral Sciences Collection from their inception to February 2016.

Results: A total of 382 articles were retrieved, 301 were screened for evaluation, and 16 were included in the review. The findings of reviewed studies support the reliability and validity of the NRS-11 when used with children and adolescents.

Discussion: Additional research is needed to clarify some unresolved questions and issues, including (1) the minimum age that children should have to offer valid scores of pain intensity and (2) the development of consensus regarding administration instructions, in particular with respect to the descriptors used for the upper anchor. On the basis of available information, the NRS-11 can be considered to be a well-established measure for use with pediatric populations.

Key Words: pain intensity, assessment, children, numerical rating scale, NRS-11

(*Clin J Pain* 2017;33:376–383)

Valid and reliable pain assessment is a basic requirement for effective clinical care of youths with acute and chronic pain. Although the multidimensional nature of the pain experience is well recognized in children and adolescents, pain intensity is the preferred pain-related domain when it comes to making pain management decisions (e.g., increasing or decreasing a drug dose) or deciding whether a treatment has been successful. A number of self-report measures of pain intensity have been developed to be used with children and adolescents, and the 0 to 10 Numerical Rating Scale (NRS-11) is among the most commonly used measures in this context.

The NRS-11 has been extensively studied in research with adults.[1] Research provides support for the validity of self-reported NRS-11 pain intensity scores through their positive and significant correlations with other measures of pain intensity[2,3] and responsivity to treatments aimed at diminishing or eliminating pain.[4,5] Moreover, NRSs seem to be preferred over other pain intensity measures by most of the patients in different populations and cultures,[6] for example, when compared with other pain intensity scales such as the Verbal Descriptor Scale or the Visual Analog Scale (VAS).[5,7]

However, until recently, there has been very limited research on the psychometric properties of the NRS-11 for use in children and adolescents with acute or chronic pain. This inconsistency was highlighted by several authors who encouraged clinicians and researchers to work to address this limitation. For example, Stinson et al[8] excluded the NRS-11 from their review on self-report pediatric pain intensity measures because they found it had not yet been adequately evaluated in research. The Ped IMMPACT group[9] stated that "Despite the widespread use of the NRS in clinical practice, the lack of research on the NRS in children and adolescents (…) precluded a recommendation for its use" (p775). von Baeyer[10] reasserted this idea and

Received for publication April 1, 2016; revised August 3, 2016; accepted June 21, 2016.

From the *Unit for the Study and Treatment of Pain—ALGOS; †Department of Psychology, Research Center for Behavior Assessment (CRAMC); ‡Institut d'Investigació Sanitària Pere Virgili; #Chair in Pediatric Pain URV-Fundación Grünenthal, Universitat Rovira i Virgili, Catalonia, Spain; §Department of Rehabilitation Medicine, University of Washington, Seattle, WA; ‖Departments of Psychology and Pediatrics, University of Saskatchewan, Saskatoon, SK; and ¶Faculty of Medicine, University of Manitoba, Winnipeg, MB, Canada.

Supported in part, by Obra Social de Caixabank, the Spanish Ministry of Economy and Competitiveness (grants PSI2012-32471 and PSI2015-70966-P), Catalonia, Spain, and Universitat Rovira i Virgili (PFR program), Catalonia, Spain. E.C. work is supported by grant PSI2014-60180-JIN of the Spanish Ministry of Economy and Competitiveness, Catalonia, Spain. J.M.'s work is supported by the Institució Catalana de Recerca i Estudis Avançats (ICREA-Acadèmia) and Fundación Grünenthal, Catalonia, Spain. The authors declare no conflict of interest.

Reprints: Jordi Miró, PhD, Departament de Psicologia, Universitat Rovira i Virgili, Carretera de Valls, s/n, 43007 Tarragona, Catalonia, Spain (e-mail: jordi.miro@urv.cat).

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website, www.clinicalpain.com.

DOI: 10.1097/AJP.0000000000000406

concluded that: "There is some evidence that NRS is used more often than any other pain scale (…). Surprisingly, there is a gap in the literature with respect to the use of the NRS in children" (p42).

Recently, a number of research groups have worked to address the knowledge gap regarding the psychometric properties of the NRS-11 when used for assessing pain intensity in youth. The aim of this paper is to review this research and summarize what is now known regarding the reliability and validity of the NRS-11 as a self-report measure of pediatric pain intensity.

METHODS

Search Strategy and Identification of Studies

The electronic databases PubMed, PsycINFO, CINAHL, and the Psychology and Behavioral Sciences Collection (EBSCOhost) were searched to identify validation studies of the NRS-11 when used to assess self-reported pain intensity in pediatric populations published through February 2016. The keywords included in the search were the following: "numeric* rating scale" AND "pain intensity" AND "child* OR adolescent OR pediatric OR paediatric" (asterisks were used as wildcards to search terms with the same word root). Reference lists of the selected papers were also reviewed to identify potentially eligible articles not found in the main search.

Selection Criteria

An article was included in this review if: (1) it was published in a peer-reviewed journal; (2) it was a validation study, that is, it intentionally addressed the evaluation of 1 or more psychometric characteristics of the NRS-11 when used to assess pain intensity in children or adolescents; and (3) the pain intensity scores were self-reported by the youths participating in the study. Following PRISMA guidelines,[11] the articles were evaluated by 2 researchers to check whether they met the inclusion criteria (Supplemental Digital Content, http://links.lww.com/CJP/A371). If there was any disagreement about eligibility for inclusion, this was discussed until consensus was reached. No additional help was needed to make a consensual decision.

RESULTS

The search strategy retrieved a total of 379 references, and 3 additional papers were identified by checking reference lists. Eighty-one of the papers identified were excluded as they were duplicated and 208 because the studies were conducted only with adult samples. Thus, 93 articles were suitable for eligibility. Figure 1 describes the selection process followed in this study.

Sixteen studies were identified that examined the psychometric properties of the NRS-11 when used to self-report pain intensity in pediatric samples; the first appeared in 2007. Eleven of the studies examined the properties of the NRS-11 in acute pain in clinical and nonclinical populations. Specifically, 5 studies were conducted with surgical patients, 1 study was conducted with children undergoing an immunization injection, and all other studies were conducted with schoolchildren who were asked to report the intensity of the most frequent pain during a certain period of time. In addition, 5 studies included samples of children and adolescents with chronic pain problems referred to pain clinics and 2 with young people with sickle cell disease. Finally, in 1 study the NRS-11 was used to measure pain intensity in youth with some type of physical disability.

Most papers included samples of children and adolescents aged between 8 and 18 years, but 2 works included children as young as 6 years of age[12,13] and 2 included participants who were 7 years old.[14,15] Finally, the work by Miró et al[16] included youth with physical disabilities aged 8 to 20 years.

In the sections that follow, we discuss the available evidence with respect to each of the following psychometric properties of the NRS-11: reliability, convergent and discriminant validity, concurrent and predictive validity, responsivity/sensitivity to change, and interpretability. We also include the information available about the agreement (concordance) between the NRS-11 and other self-report measures of pain intensity. Table 1 provides additional and detailed information about the reviewed studies.

Reliability

Reliability of pain measures is usually estimated by computing test-retest stability in samples for whom little change in pain is anticipated (eg, those who do not receive treatment) or are self-described as having little to no change in pain. We were able to identify only 1 study that examined reliability of the NRS-11 pain intensity scores in this way. Bailey et al[17] asked a sample of 202 children, between the ages of 8 and 17 years with abdominal pain being seen in an emergency department, to rate their current pain intensity twice; first after they arrived and then again either just after an intervention to reduce the pain or right before discharge if no intervention was implemented. The average time between assessments was 120 ± 75 minutes. To analyze the agreement between the 2 measurements, they used the Bland-Altman Method,[18] which provides a plot that depicts the difference in score between the 2 measurements against the average of the 2 measurements for each participant. Results showed a high agreement (limits of agreement of -0.9 and 1.2, 95% CI) between pain intensity scores in those 44 children who did not report any change in pain intensity between the first and the second measurement, supporting the tentative conclusion that the NRS-11 shows high levels of test-retest stability in an 8- to 17-year-old sample of children. Such a conclusion is deemed tentative because (1) it is based on evidence from only a single study and (2) the high agreement was found for a discreet subsample that was selected for reporting no change in pain.

Validity

Several studies support different types of validity of the NRS-11 when used with children and adolescents.

Convergent and Discriminant Construct Validity

Construct validity has been evaluated examining both convergent and discriminant validity studies. The convergent validity of the NRS-11 has been supported by the moderate-to-high correlations between the NRS-11 pain intensity scores and scores from other self-report scales that also measure pain intensity, such as the Faces Pain Scale-Revised (FPS-R) (r range across studies = 0.75 to 0.93),[13,14,19,20] the VAS (r range across studies = 0.73 to 0.95),[13,14,20] the Verbal Rating Scale (VRS) (r range, 0.48 to 0.0.79),[16,21] and the Colour Analogue Scale (CAS) (r range across studies = 0.58 to 0.84).[13,22] In contrast to the

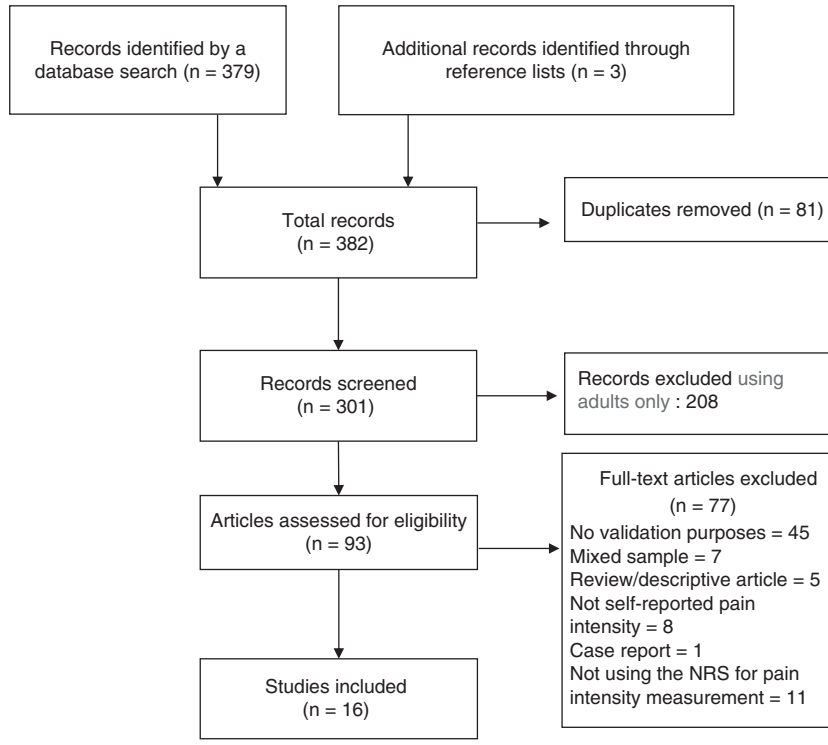

FIGURE 1. Flow chart followed to identify the studies that were included in accordance with PRISMA guidelines.

moderate-to-high correlations found in the 8 studies cited above, only a modest correlation was found in 1 study between the NRS-11 and the Wong-Baker FACES Pain Rating Scale ($r = 0.33$).[16]

In addition, Bailey et al[17] compared the pain intensity reports on the NRS-11 and on a VRS with 3 pain intensity levels: mild, moderate, and severe pain. The results of this study support the convergent validity of the NRS-11 ratings. That is, pain intensity ratings on the NRS-11 of children who rated their pain intensity as mild in the VRS were lower than ratings of those who rated pain intensity as moderate, which in turn were lower than those who rated it as severe. Furthermore, the convergent validity of an electronic version of the NRS-11 has been reported in a recent study by showing high correlations with other self-report pain intensity scales (the FPS-R, the CAS, or the VAS) administered using the traditional paper-and-pencil format (r range, 0.59 to 0.79).[23]

The discriminant validity of the NRS-11 was supported by findings indicating moderate correlations between scores on the NRS and scores on pain unpleasantness ($r = 0.63$)[20] and functional disability measures ($rs = 0.35$ and 0.43).[24] Discriminant validity of NRS-11 pain intensity scores has also been examined using Fisher z-transformation to compare the magnitude of the correlation coefficient between ratings on the NRS-11 and other self-report pain intensity scales with those between ratings on the NRS-11 and other nonpain intensity measures. Available results show that correlation coefficients between the scores on the NRS-11 and other pain intensity scales (such as the FPS-R, the VAS, or the CAS) were significantly larger than those between the scores on the NRS-11 and measures of pain-related affect (z range across studies = 2.05 to 7.62, all $P < 0.05$)[13,19,22] and disability ($z = 7.62$, $P < 0.001$).[13,19] An electronic version of the NRS-11 has also shown adequate discriminant validity properties in relation with measures of fatigue (z range, 2.4 to 8.8, all Ps > 0.05).[23]

Concurrent and Predictive Criterion-related Validity

Available studies also show good results for criterion-related validity (concurrent and predictive) of the pain intensity scores on the NRS-11. For example, scores on the NRS-11 have been shown to be moderately correlated (at least $r \geq 0.3$ to 0.5) with other measures at the same time of measurement to analyze concurrent validity, while using measures of disability (r range, 0.22 to 0.39),[16,19] psychological functioning (r range, 0.46 to 0.66),[16,19] and pain interference ($r = 0.62$),[16] or at 2 different time points of measurement to study the predictive validity of NRS-11 scores using measures of disability ($r = 0.39$) and quality of life ($r = 0.46$).[19] In addition, an electronic version of the NRS-11 has shown adequate concurrent validity properties when correlated with measures of pain catastrophizing ($r = 0.30$).[23]

Responsivity or Sensitivity to Change

Finally, data regarding the responsivity or sensitivity to change with the NRS-11, relative to other measures, is inconsistent. Bailey et al[17] compared the median of NRS-11 scores reported by 46 children before and after a pain-relieving intervention in the emergency department and found large (and statistically significant) differences between them ($NRS\text{-}11_{before} = 6.5$, $NRS\text{-}11_{after} = 4.0$; $W = 2$, $P < 0.001$). Connelly and Neville,[20] in a study with 29 children aged 9 to 18 undergoing surgical repair of pectus excavatum, compared the relative responsivity of NRS-11, VAS, and FPS-R. They assessed pain intensity 5

TABLE 1. Summary of Studies Reporting the NRS-11 Psychometric Properties, by Chronological Order of Publication

				Reliability	Validity					Other Characteristics	
References	Sample N (Age Range) Type of Pain	Method of Administration	NRS-11 Anchors	Test-Retest Stability	Convergent	Discriminant	Concurrent	Predictive	Responsivity/ Sensitivity	Interpretability	Concordance
Bailey et al[28]	87 (8-18) Acute abdominal pain	Verbal	0 = no pain 10 = worst possible pain								✓
von Baeyer et al[14]	69 (7-17) Postoperative pain 29 (9-17) Postoperative pain 236 (11-17) Vaccination pain	Verbal & Printed	0 = no pain 10 = most or worst pain		✓						✓
Miró et al[19]	175 (8-12) Most frequent pain last 3 mo (acute/chronic) 63 (6-16) Postoperative pain	Verbal	0 = no pain 10 = very much pain		✓	✓	✓	✓			
Bailey et al[17]	202 (8-17) Acute pain	Verbal	0 = no pain 10 = the worst possible pain	✓	✓				✓	✓	✓
Connelly and Neville[20]	29 (9-18) Postoperative pain	Verbal	0 = no pain 10 = worst pain possible		✓	✓			✓		✓
Voepel-Lewis et al[15]	113 (7-16) Postoperative pain	Verbal	0 = no pain or hurt 10 = the most or worst pain/hurt							✓	
Pagé et al[24]	83 (8-18) Postoperative pain	Verbal	0 = no pain at all 10 = worst possible pain		✓	✓			✓		✓
Sánchez-Rodríguez et al[12]	126 (6-8) Most frequent pain last 3 mo	Verbal	0 = no pain 10 = very much pain								
Castarlenas et al[13]	126 (6-8) Most frequent pain last 3 mo	Verbal	0 = no pain 10 = very much pain		✓	✓	✓				
Myrvik et al[26]	28 (8-18) Chronic pain	Printed	0 = no pain 10 = the worst pain possible							✓	
Hirschfeld et al[27]	153 (-) Chronic pain	Verbal	0 = no pain 10 = maximal pain							✓	
Ruskin et al[22]	143 (8-17) Chronic pain	Verbal	0 = no pain 10 = strongest or worst pain you can imagine		✓	✓					✓
Castarlenas et al[29]	191 (12-18) Most frequent pain last 3 mo	Electronic format	0 = no pain 10 = very much pain								✓
Myrvik et al[26]	28 (8-18) Chronic pain	Verbal	0 = no pain 10 = worst pain possible		✓					✓	
Sánchez-Rodríguez et al[23]	180 (12-19) Most frequent pain last 3 mo	Electronic format	0 = no pain 10 = very much pain		✓	✓	✓				✓
Miró et al[16]	113 (8-20) Chronic pain	Verbal	0 = no pain 10 = pain as bad as could be		✓		✓				

times per day during the 3 days following surgery. Using hierarchical linear models, they found statistically significant improvements in pain as measured by the VAS and FPS-R, but not as measured by the NRS-11, suggesting that the NRS-11 was the least sensitive for detecting changes in pain intensity over the 3 days following surgery. Finally, a recent study by Pagé et al[24] showed that the NRS-11 detected changes in pain intensity over a 2-week period in a sample of 83 youths with scoliosis and osteotomy who had a planned major surgery.

Interpretability

Another issue of interest that has attracted the attention of researchers is the interpretability of the NRS-11, that is, the meaning of scores obtained by this measure. We identified 5 studies that examined this characteristic. Bailey et al[17] examined the median scores of the NRS-11 across the categories of the VRS, finding that average values of 3, 6, and 8 on the NRS-11 are associated with the categories of mild, moderate, and severe pain on the VRS, respectively. Hirschfeld and Zernikow[25] also conducted a study to

identify the cut points for the NRS-11 in a sample of children and adolescents with chronic pain. However, pain intensity information was based on parent proxy report for more than a third of the participants, those under the age of 11 years. As the results are presented differentiated by pain typology but collapsed by age, it is not possible to analyze the information specifically for those who used the NRS-11 as a self-report measure, and therefore this paper was excluded from this review.

Two studies have examined the relationship between NRS-11 ratings and other pain-related variables such as the perceived need of medicine or patient satisfaction. Voepel-Lewis et al[15] studied the interpretability of NRS-11 scores in a sample of children undergoing surgery. In addition to rating their pain intensity, the participants were asked to report the perceived need for medicine and the perceived satisfaction with treatment to establish the NRS-11 cut points related to these 2 outcomes. Results showed that scores >4 in the NRS-11 discriminated those children who perceived they needed medicine, whereas scores >6 discriminated those children who were dissatisfied with treatment from those who were not. However, Voepel-Lewis et al[15] pointed out that the percentage of false positives and negatives in this study were both fairly high, indicating that these may not reflect the levels associated with what individual children judge to be mild, moderate, and severe pain. Following a similar procedure, Myrvik et al[26] found in a sample of 28 pediatric patients with sickle cell disease that pain scores from 8 to 10 on the NRS-11 identified children who needed pain medication, and scores from 9 to 10 identified children who were dissatisfied with the treatment.

Finally, a number of investigators have sought to identify the minimum clinically significant difference (MCSD) of the NRS-11, that is, the smallest difference in pain ratings that reflects changes in a clinical intervention and which is perceived as beneficial by patients. For example, Bailey et al[17] analyzed the median difference between the first and the second assessment when participants considered their pain level as "a lot worse," "a little worse," "the same," "a little better," or "a lot better." Results showed that the MCSD (representing a change of "a little better" or "a little worse") for the NRS-11 was 1 out of 10. In addition, Voepel-Lewis et al[15] also established the difference that reflected a change in pain intensity at $-1/+1$ ("feeling a little better" or "feeling worse," respectively). Hirschfeld et al[27] studied the MCSD in a sample of adolescents with chronic pain and their results also suggested that a change of 1/10 or a 12.5% decrease on the NRS-11 represents a meaningful change of pain intensity in this population. Finally, Myrvik et al[26] found that a decrease of 0.9 in the NRS-11 scores represent a clinically significant reduction of pain intensity in a sample of children and adolescents with sickle cell disease.

Agreement Between the NRS-11 and Other Pain Intensity Scales

A high correlation between 2 pain intensity self-report measures does not necessarily mean there is good agreement between them. Correlation analyzes the association between one variable and another, and it is not the best alternative to study the comparability or interchangeability of scales. For example, scores of pain intensity of 0, 2, 4 obtained with the scale A and scores of 6, 8, 10 obtained with the scale B have a perfect association ($r = 1.0$), but there is not concordance. Having information about agreement or concordance shows whether 2 scores obtained with different scales are comparable in magnitude and consequently, interchangeable.

The agreement between scores from the NRS-11 and scores from other self-report pain intensity scales has been documented; results are mixed. Although 2 studies support equivalence,[14,24] most studies suggest that the NRS-11 cannot be considered to be equivalent to other scales such as the FPS-R,[12] the Wong-Baker FACES Pain Rating Scale,[28] the VAS,[12,17,28] or the CAS.[12,22,28] Nevertheless, the distribution of scores obtained with the NRS-11 seems to follow a similar pattern to those obtained with other pain intensity scales, such as the FPS-R or the VAS,[14,20] with the exception that ratings on the NRS-11 tend to be higher than those of other scales.[12,19,20] Ratings from an electronically administered NRS-11 have shown agreement with scores from the traditionally administered (paper-and-pencil) version of this self-report scale when using the Bland-Altman method at 80% CI.[23,29]

DISCUSSION

The aim of this review was to examine the available evidence regarding the psychometric properties of the NRS-11 as a measure of self-reported pain intensity in children and adolescents, updating the review by von Baeyer.[30] We were able to identify 16 studies that purposely studied the psychometric properties of the NRS-11 pain intensity scores. Generally speaking, results show that the NRS-11 provides valid and reliable scores when used with pediatric populations. This represents a fundamental advance in the literature about the measurement of pediatric pain intensity, as little psychometric information about the NRS-11 was available just a few years ago. The proliferation of works around this topic reflects the interest of researchers and clinicians to respond to the need, as noted in the Introduction, to provide evidence that the NRS pain intensity scores would prove to be reliable and valid when used to measure children's pain intensity.

Most of the clinical decisions about pain management are based on changes in pain levels and having some data about this domain represents important progress. The use of the NRS for self-report of pain intensity has been criticized[31] on the grounds that it does not capture the complexity of pain. But to our knowledge, no one has claimed that pain intensity is the only important aspect of pain to consider. Pain intensity is one among numerous domains requiring consideration in the assessment of people with pain,[9,32–34] and the NRS seems to be an efficient tool for this purpose.

The increasing number of studies published on the psychometric properties and characteristics of the NRS-11, when used in pediatric populations, demonstrates the interest of clinicians and researchers in this scale. On the basis of reviewed studies, at this point, there is enough evidence to consider the NRS-11 as a "well-established measure" for assessing pain intensity in children and adolescents according to 3 criteria for this category[35(p913)]: (1) it has been presented in at least 2 peer-reviewed articles by different investigators or investigatory teams; (2) studies provide sufficient detail about the measure to allow critical evaluation and replication; and (3) there is detailed information indicating good validity and reliability in at least 1 peer-reviewed article.

Despite the increasing number of studies evaluating the psychometric properties and the performance of the NRS-11 scores in youths, however, some issues about the use of this scale still remain to be clarified.[30] Two related issues that need to be addressed have to do with the administration of the NRS-11: (1) What are the most appropriate administration instructions when the scale is presented to children? (2) What should be the descriptor indicating the high pain endpoint? The results of an email survey[30] revealed that clinicians use a large number of alternatives both to explain the scale and to describe the upper anchor when they ask children to report their pain intensity. There is no controversy about how the lower anchor should be explained (ie, as "No pain"). However, the same cannot be said for the upper anchor. The most common upper anchor used in the studies included in this review is "The worst pain possible," but numerous other phrases are also used, such as "Very much pain" (Table 1), which is the upper anchor established for the FPS-R.

Preliminary results regarding the most appropriate top anchor for the NRS-11 have been presented by Young and von Baeyer.[36] These authors conducted a study to determine whether the ability of children to use the NRS-11 is influenced by the expressions or words used to explain the anchors. In their study, 98 university students graded 6 different top anchors based on severity, concreteness, and clarity (ie, "Most hurt," "Hurt as bad as it could be," "Worst hurt you can imagine," "Most hurt possible," "Very much hurt," and "Hurt as bad as breaking your arm"). Then, 133 children enrolled in grades 1 to 5 participated in assessing the calibration and comprehension of the anchors as applied to hypothetical pain vignettes. According to the results of their study, "Most hurt" was the least useful endpoint due to the high mean error score. The remaining alternatives seemed to have similar utility; the authors concluded that each one could therefore be considered roughly equivalent.

Still, having a standard upper anchor would be useful for the field and allow for greater comparability across studies. Further research assessing how children understand such anchor phrases, and working toward a consensus of researchers and clinicians on standardizing these phrases, is needed to improve pain assessment with the NRS-11 and other pediatric self-report scales. Differences between languages should be taken into account: an anchor phrase that has good psychometric qualities in one language may, when translated into another language, be meaningless or hard to understand. Thus, standard anchor phrases will need to be identified that are valid across languages and cultures, which will allow for cross-cultural comparisons. If this is not found to be possible, then different anchor phrases may be needed for different regions. Research is now needed to help determine what these anchors might be.

In addition, future studies should determine whether the method of presentation of the NRS-11 (ie, verbally or in printed format) has an influence on the pain intensity scores reported by participants. Research in adults suggests that written versions of the numerical scale reduce the amount of error, relative to verbal presentations.[37] Using the printed form, the anchors are clarified and the prevalence of out-of-range scores (eg, 25 out of 10) and fractional values (eg, 5½) is reduced because the respondent sees the boundaries of the scale and the possible alternative scores.

One additional issue that deserves further clarification is that of the lowest age at which the NRS-11 could be reliably and validly used. Most studies have been conducted with samples of children 8 years of age or older. However, we have conducted studies to analyze the psychometric properties of the NRS-11 when used in even younger children. Preliminary data[12,13] have shown adequate construct and criterion-related validity from pain intensity scores even when the NRS-11 is used with children as young as 6 years of age. To confirm this, research studies are needed that not only include participants who are as young as 6 (or younger), but that includes an adequate number of participants at the lower ages to have adequate statistical power to detect any age effects.

It is well known that certain cognitive skills are required if children are to understand the meaning of the numbers of the NRS-11, and provide accurate ratings of pain intensity so that efficient pain management can be guaranteed. Several activities have been suggested to identify those children who are able to give accurate self-reports and those who are not.[38,39] Children about 2 years old are often able to say numbers by rote learning repetition and even verbalize them in the correct order but they have not acquired yet a quantitative understanding of the numbers, a skill essential to use a self-report pain intensity scale.[40] Two of the works presented in this review used screening tasks to determine whether children were able to provide a reliable and valid pain intensity score. Voepel-Lewis et al[15] used seriation and classification tasks, and Castarlenas et al[13] used counting, seriation, and comparison tasks. Although the use of screening tasks is a good addition in these studies, it is unclear whether they are capable of discriminating between children who can and cannot reliably use the NRS-11. A screening task should be brief but also be able to evaluate several numerical skills; future studies should be able to clarify whether screening tasks used are sensitive enough to distinguish between children who are able to reliably and validly use the NRS-11 to report on their pain intensity.

Present pain intensity is not the only variable measured by NRS-11 scores. Children are frequently asked about their usual, maximum, or minimum pain intensity over a specified period of time such as the last week or month. In addition, pain intensity at rest is sometimes differentiated from pain experienced during specified activities. These judgments place additional cognitive requirements such as memory for past pain and the ability to construct and report an average or usual score over a period of time. In the current review, these aspects of pain intensity were not differentiated; more than a half of the studies referred either to present pain intensity or to the pain experienced in a procedure just completed.

Most samples in studies reviewed here included as participants children who were experiencing acute pain. Although there is no clear reason to expect that these findings would not generalize to children with chronic pain, additional studies that include samples of children with chronic pain are needed to demonstrate this.

Another area of interest is the study of the properties of the NRS-11 when administered with the so-called information and communication technologies, for example, smartphones. With the progressive implementation of the information and communication technologies (eg, in online treatments), studies of feasibility and usability, as well as compliance, will be explicitly needed. In fact, a recent study with adolescents aged 12 to 18 suggests that scores obtained with a NRS-11 incorporated in an

electronic device are comparable with scores on the traditional verbal version of this scale.[29]

To summarize, and in light of the findings from this review, the results indicate support for the reliability and validity of the NRS-11 as a self-report scale of pain intensity in many populations of children and adolescents as young as 8 years old, and perhaps even as young as 6 years old. Thus, the findings indicate that NRS-11 can be viewed as a "well-established measure," both for the assessment of acute and chronic pain in children and adolescents. These positive properties, together with the advantages of the NRS-11 over other pain scales (eg, no physical materials are needed for its administration, it can be administered orally), support its recommendation for use with children and adolescents. Finally, we would like to emphasize that pain intensity is only one of many important pain-related domains that can and should be assessed in clinical and research settings. Careful attention should also be paid to other important pain-related domains such as pain location, pain extent, and the impact of pain on psychological and physical function, among others. A comprehensive assessment of the person with pain, across different levels and units of analysis, is fundamental to understanding the pain as well as providing the best possible care.

REFERENCES

1. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;113:9–19.
2. 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;10:390–392.
3. Bahreini M, Jalili M, Moradi-Lakeh M. A comparison of three self-report pain scales in adults with acute pain. *J Emerg Med*. 2015;48:10–18.
4. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain*. 2011;152:2399–2404.
5. Chien CW, Bagraith KS, Khan A, et al. Comparative responsiveness of verbal and numerical rating scales to measure pain intensity in patients with chronic pain. *J Pain*. 2013;14:1653–1662.
6. 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;41:1073–1093.
7. Gagliese L, Weizblit N, Ellis W, et al. The measurement of postoperative pain: a comparison of intensity scales in younger and older surgical patients. *Pain*. 2005;117:412–420.
8. Stinson JN, Kavanagh T, Yamada J, et al. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain*. 2006;125:143–157.
9. McGrath PJ, Walco GA, Turk DC, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. *J Pain*. 2008;9:771–783.
10. von Baeyer CL. Children's self-report of pain intensity: what we know, where we are headed. *Pain Res Manag*. 2009;14:39–45.
11. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264–269. W64.
12. 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;153:1715–1719.
13. 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;14:297–304.
14. von Baeyer CL, Spagrud LJ, McCormick JC, et al. Three new datasets supporting use of the Numerical Rating Scale (NRS-11) for children's self-reports of pain intensity. *Pain*. 2009;143:223–227.
15. Voepel-Lewis T, Burke CN, Jeffreys N, et al. Do 0-10 numeric rating scores translate into clinically meaningful pain measures for children? *Anesth Analg*. 2011;112:415–421.
16. Miró J, Castarlenas E, de la Vega R, et al. Validity of three rating scales for measuring pain intensity in youths with physical disabilities. *Eur J Pain*. 2016;20:130–137.
17. Bailey B, Daoust R, Doyon-Trottier E, et al. Validation and properties of the verbal numeric scale in children with acute pain. *Pain*. 2010;149:216–221.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
19. 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;13:1089–1095.
20. 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;11:1451–1460.
21. Pagé MG, Campbell F, Isaac L, et al. Reliability and validity of the Child Pain Anxiety Symptoms Scale (CPASS) in a clinical sample of children and adolescents with acute postsurgical pain. *Pain*. 2011;152:1958–1965.
22. 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;19:141–148.
23. Sánchez-Rodríguez E, de la Vega R, Castarlenas E, et al. An app for the assessment of pain intensity: validity properties and agreement of pain reports when used with young people. *Pain Med*. 2015;16:1982–1992.
24. Pagé MG, Katz J, Stinson J, et al. Validation of the numerical rating scale for pain intensity and unpleasantness in pediatric acute postoperative pain: sensitivity to change over time. *J Pain*. 2012;13:359–369.
25. Hirschfeld G, Zernikow B. Variability of "optimal" cut points for mild, moderate, and severe pain: neglected problems when comparing groups. *Pain*. 2013;154:154–159.
26. Myrvik MP, Drendel AL, Brandow AM, et al. A comparison of pain assessment measures in pediatric sickle cell disease: visual analog scale versus numeric rating scale. *J Pediatr Hematol Oncol*. 2015;37:190–194.
27. Hirschfeld G, Wager J, Schmidt P, et al. Minimally clinically significant differences for adolescents with chronic pain-variability of ROC-based cut points. *J Pain*. 2014;15:32–39.
28. Bailey B, Bergeron S, Gravel J, et al. Comparison of four pain scales in children with acute abdominal pain in a pediatric emergency department. *Ann Emerg Med*. 2007;50:379–383. e1-e2.
29. Castarlenas E, Sánchez-Rodríguez E, Vega Rde L, et al. Agreement between verbal and electronic versions of the numerical rating scale (NRS-11) when used to assess pain intensity in adolescents. *Clin J Pain*. 2015;31:229–234.
30. 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;13:1005–1007.
31. Green SM, Krauss BS. The numeric scoring of pain: this practice rates a zero out of ten. *Ann Emerg Med*. 2015;67:573–575.
32. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, eds. *Handbook of Pain Assessment*, 3rd ed. New York: Guilford Press; 2011:21–44.

33. Miró J. *Dolor Crónico: Procedimientos de Evaluación E Intervención Psicológica*. Bilbao: Desclée de Brouwer; 2003.
34. McGrath PJ, Stevens BJ, Walker SM, et al. *Oxford Textbook of Paediatric Pain*. Oxford, UK: Oxford University Press; 2013.
35. Cohen LL, La Greca AM, Blount RL, et al. Introduction to special issue: evidence-based assessment in pediatric psychology. *J Pediatr Psychol*. 2008;33:911–915.
36. Young MA, von Baeyer CL. Selection of anchors to improve children's understanding of the numerical pain rating scale. *13th World Congress on Pain, International Association for the Study of Pain*. Montreal: IASP Press; 2010.
37. Pasero C, McCaffery M. *Pain Assessment and Pharmacologic Management*. St Louis, MO: Elsevier Mosby; 2011.
38. Besenski LJ, Forsyth SJ, von Baeyer CL. Screening young children for their ability to use self-report pain scales. *Pediatr Pain Lett*. 2007;9:1–7.
39. von Baeyer CL, Uman LS, Chambers CT, et al. Can we screen young children for their ability to provide accurate self-reports of pain? *Pain*. 2011;152:1327–1333.
40. Chan JY-C, von Baeyer CL. Cognitive developmental influences on the ability of preschool-age children to self-report their pain intensity. *Pain*. 2016;157:997–1001.

