Outcome Measures for Pediatric Pain: Practical Guidance on Clinical Use in Juvenile Arthritis

Chitra Lalloo,1 Giulia Mesaroli,1 Mallika Makkar,2 and Jennifer Stinson1

INTRODUCTION

Pain is the most common and distressing symptom experienced by children and adolescents with arthritis. In a survey of North American pediatric rheumatologists, 77% of respondents agreed that there are pediatric patients who continue to experience moderate to severe pain despite adequate treatment with disease-modifying therapy and nonsteroidal anti-inflammatory drugs (1). This clinical experience is supported by evidence from electronic diary (e-diary) pain studies. For instance, Bromberg et al (2) found that arthritis patients (n = 59) continued to report pain in 66% of e-diary entries despite being under treatment with a disease-modifying antirheumatic drug (79%) or a biologic agent (47%). Across the 1-month study period, no participants were completely pain free, and 86% reported at least one episode of severe pain.

The assessment of pain in children with arthritis is a common challenge for clinicians. A recent qualitative study found that health care professionals report gaps in pain-specific knowledge and skills to assess pain in children with arthritis (3). In the context of assessment, it is important to recognize pain as a multidimensional experience that is comprised of sensory, affective, and evaluative components (3). The sensory dimension is related to quality (what pain feels like), intensity (how much pain hurts), location (spatial distribution of pain), and duration (how long pain lasts). The affective dimension is related to the emotional impact of pain, such as the extent to which pain is perceived as unpleasant or distressing. The evaluative dimension describes the degree to which pain is perceived to interfere with physical, psychological, role, and social functioning.

Guidance on pain assessment was offered by a working group comprised of representatives from the American College of Rheumatology, the American Academy of Pediatrics, the American Board of Pediatrics, and the Association of Rheumatology Health Professionals (4). This working group recommended that “pain should be assessed in all patients at the first visit [to a pediatric rheumatologist after a diagnosis has been made] and at each subsequent visit that occurs at least 7 days apart” (4). The group also emphasized that pain must be assessed using valid and reliable tools that are developmentally appropriate for the individual patient.

This article will provide practical information to help clinicians assess pain in their patients with pediatric arthritis. Information about outcome measures for pediatric fibromyalgia can be found in the article by Daffin et al in this issue. Each outcome measure is described individually, and Tables 1 and 2 summarize and draw comparisons across different measures. Our aim is to support clinicians in selecting an appropriate measure to assess pain in their patients with pediatric arthritis.

MEASURES OF PAIN INTENSITY

VISUAL ANALOG SCALE (VAS)

Description

Purpose. The VAS was designed to assess self-reported pain intensity.

Content or domains. The VAS is a single-domain measure focused on pain intensity.

Number of items. The VAS consists of one item.

Response options/scale. Respondents are asked to mark their pain intensity on a line that is typically 100 mm in length. The lower anchor is labeled “no pain.” The upper anchor label is varied, including “worst pain possible,” “worst possible pain,” and “pain as bad as it could be.” The recommended upper anchor is “worst pain possible” (5).

Recall period for items. The VAS is typically used to assess current pain.

Cost to use. The VAS is free to use.

1Chitra Lalloo, PhD, Giulia Mesaroli, MScPT, Jennifer Stinson, RN-EC, PhD, CPNP: The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada; Mallika Makkar, BSc: The Hospital for Sick Children, Toronto, Ontario, Canada.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Chitra Lalloo, PhD, Peter Gilgan Centre for Research and Learning, 686 Bay Street, Toronto, ON MSG 0A4. E-mail: chitra.lalloo@sickkids.ca.

Submitted for publication January 31, 2020; accepted in revised form April 4, 2020.
OUTCOME MEASURES FOR PAIN ASSESSMENT IN JUVENILE ARTHRITIS

How to obtain. No standardized paper version is available, but one can be easily constructed. A validated electronic version is freely available through a smartphone application called Painometer V2 (6–8) (http://app.dolorinfantil.urv.cat/en/painometer/).

Practical application

Method of administration. Paper or electronic delivery options are available.

Scoring. The 100-mm line has a corresponding score ranging from 0 to 100 or 0 to 10. The score is calculated as the measured distance in millimeters between the left anchor of the scale and the mark drawn by the participant.

Score interpretation. Higher scores correspond with higher pain intensity. Typically, clinicians interpret a score of 10 to 30 as mild pain, 40 to 60 as moderate pain, and 70 to 100 as severe pain.

Respondent time to complete. The VAS takes under 1 minute to complete.

Administrative burden. The VAS requires low burden for in-person assessments. This tool cannot be administered over the telephone. The only training required is how to use a ruler to generate a score, which typically takes less than 1 minute.

Translations/adaptation. Translations are available in Catalan (7), French (9), Thai (9), Portuguese (10), and Spanish (10).

Psychometric information

Floor and ceiling effects. Insufficient data are available.

Reliability. As per a 2019 systematic review (5), few pediatric studies have examined the reliability of the scores generated by the VAS, and no studies have been conducted in children with arthritis. Only two studies were identified, ranging in quality from poor to excellent on the COnsensus-based Standards for the selection of health Measurement InStuments (COSMIN) checklist (11). The first study (12), rated as excellent quality, was conducted in 151 patients (aged 8–17 years) presenting to a pediatric emergency department. The VAS repeatability coefficient was determined to be 12 mm for the 100 patients who reported that their pain had not changed during pain assessments completed at baseline, 3-minute, and 6-minute time points. This finding indicates that when a patient reports the same pain intensity, a clinician can be expected to observe a repeated VAS score to fall between ±12 mm of the initial score 95% of the time (12). The second study (13), rated as poor quality, was conducted in 100 children (aged 3–18 years) presenting to a sickle cell anemia clinic. Test-retest reliability was assessed wherein retest scores that were within ±10 mm of the initial score were assigned a reliability score of 1. Otherwise, the assigned reliability score was zero. Across the sample, test-retest reliability was reported as 45% (13).

Validity. The systematic review by Birnie et al (5) identified one publication of excellent quality that reported validity data for the VAS. Another study (7) found evidence of moderate-to-high convergent validity, adequate discriminant validity with fatigue ratings, and adequate concurrent validity with pain catastrophizing ratings. The study sample was comprised of 180 school children aged 12 to 19 years (7).

Responsiveness. The systematic review by Birnie et al (5) did not identify any relevant responsiveness studies in children with arthritis. However, in a sample of 456 children (aged 6–17 years) presenting to an emergency department, Le May et al found evidence of moderate responsiveness for the VAS (14). As hypothesized, mean differences in pain scores were significantly lower 60 minutes following the administration of medication (mean −1.43, SD 1.97; P < 0.0001). The standardized response mean was 0.72 (14).

Minimally important differences. The minimally important difference of the VAS is 10 to 12 mm (12,15).

Generalizability. Limited psychometric research has been conducted in pediatric arthritis. The generalizability of data from emergency department settings and patients with other health conditions is unknown. Existing evidence provides a weak recommendation for using this tool to assess persistent pain in children aged 6 years and older (5). There is a weak recommendation against using this tool in children younger than 6 years old with persistent pain (5).

Use in clinical trials. The VAS has been used in multiple clinical trials (16,17).

Critical appraisal of overall value to the rheumatology community

Strengths include simplicity, high accessibility, low administrative burden, paper and electronic delivery options, and ease of scoring. Cautions include inconsistent scale anchors and limited available psychometric data for pediatric arthritis.

FACES PAIN SCALE REVISED (FPS-R)

Description

Purpose. The FPS-R was designed to assess self-reported pain intensity.
Content or domains. The FPS-R is a single-domain measure focused on pain intensity.

Number of items. The FPS-R consists of one item.

Response options/scale. There are six response options consisting of drawn human faces. Each option exhibits a facial expression representing a different level of pain.

Recall period for items. The FPS-R is used to assess current pain.

Cost to use. The FPS-R is free to use.

How to obtain. The FPS-R is available for download from the International Association for the Study of Pain (IASP) website (https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1823&navItemNumber=1119).

Practical application

Method of administration. Paper or electronic delivery options are available. The recommended script for clinicians is “These faces show how much something can hurt. This face [point to face on far left] shows no pain. The faces show more and more pain [point to each from left to right] up to this one [point to face on far right]—it shows very much pain. Point to the face that shows how much you hurt [right now].” Do not use words such as “happy” or “sad” when explaining the tool to the child. The FPS-R is designed to measure how children feel inside, not how their face looks.

Scoring. Each face is assigned a different score, as follows (from left to right): 0, 2, 4, 6, 8, or 10.

Score interpretation. A score of 0 corresponds with no pain, whereas a score of 10 corresponds with very much pain.

Respondent time to complete. The FPS-R takes under 1 minute to complete.

Administrative burden. The FPS-R requires low burden because there is a validated script to follow.

Translations/adaptation. The FPS-R is available for download on the IASP website in 68 different languages apart from English, including French, Spanish, and Chinese.

Psychometric information

Floor and ceiling effects. Insufficient data are available.

Reliability. There have been no published reliability studies specific for children with arthritis (5). Tsze et al assessed test-retest reliability of the FPS-R in a sample of 40 children aged 4 to 17 years with an existing painful condition (eg, abdominal pain or back pain) (18). Pain assessments were conducted at least 30 minutes apart. Among children who reported that their pain was about the same prior to the second pain assessment, the repeatability coefficient was ±0.53 using the Bland-Altman method. The authors described these data as indicating an acceptable degree of test-retest reliability (18).

Validity. Hicks et al demonstrated evidence of convergent validity of the FPS-R in a hospitalized sample of 90 children aged 4 to 12 years with a variety of painful conditions, including arthritis (19). In comparison with the Color Analog Scale (20) and Visual Analog Scale, Pearson’s correlations with the FPS-R ranged from $r = 0.84$ to 0.94, which exceed the common conventional standard of $r$ being more than 0.70. More recently, Sánchez-Rodríguez et al demonstrated convergent validity ($r = 0.89$) between electronic and paper versions in a sample of 180 schoolchildren aged 12 to 19 years (7).

Responsiveness. Tsze et al evaluated responsivity by comparing initial mean FPS-R scores with postanalgesic scores using the paired sample $t$-test (18). In a sample of 314 children aged 4 to 17 years with an existing painful condition, the authors identified a difference in FPS-R scores in response to analgesic, demonstrating adequate responsiveness.

Minimally important differences. One face on the FPS-R is considered the minimal clinically significant difference, whereas two faces are considered the ideal clinically significant difference (21). However, these estimates are based on psychometric data from children with acute pain. No relevant studies have been completed in children with arthritis-related pain (5).

Generalizability. Limited psychometric research on the FPS-R has been conducted in pediatric arthritis. Most studies have been conducted in acute pain settings. Existing evidence provides a weak recommendation against using this tool in children aged 4 years and older with persistent pain (5).

Use in clinical trials. The FPS-R has been used in clinical trials (22).

Critical appraisal of overall value to the rheumatology

Strengths include simplicity, high accessibility, low administrative burden, paper and electronic delivery options, and ease of scoring. Cautions include potential for confusion around emotional
effect versus pain intensity and limited available psychometric data for pediatric arthritis.

**ELEVEN-POINT NUMERICAL RATING SCALE (NRS-11)**

**Description**

**Purpose.** The NRS-11 is a self-report tool used to assess pain intensity.

**Content or domains.** The NRS-11 is a single-domain measure of pain intensity.

**Number of items.** The NRS-11 has one item.

**Response options/scale.** The NRS-11 is numbered from 0 (typically labeled “no pain”) to 10 (typically labeled “worst pain you can think of”).

**Recall period for items.** The NRS-11 has a recall period of 7 days.

**Cost to use.** The NRS-11 is free to use.

**How to obtain.** The standard NRS-11 can be easily constructed. The validated Patient-Reported Outcome Measurement Information System (PROMIS) Pediatric Pain Intensity NRS-11 is freely available from the HealthMeasures website (http://www.healthmeasures.net/explore-measurement-systems/promis/obtain-administer-measures).

**Practical application**

**Method of administration.** Paper or electronic delivery options are available.

**Scoring.** Responses are scored from 0, representing no pain, to 10, representing the “worst pain you can think of.” Detailed instructions on scoring PROMIS Pain Intensity NRS-11 are available online (23).

**Score interpretation.** Although other PROMIS measures typically require conversion of raw scores into standardized T scores, the single-item Pain Intensity NRS-11 relies only on raw scores. In both the PROMIS and the standard NRS-11, a higher score reflects higher pain intensity.

**Respondent time to complete.** The NRS-11 takes under 1 minute to complete.

**Administrative burden.** Time to administer, score, and interpret the NRS-11 is minimal. No special software or equipment is required.

**Translations/adaptation.** French, Icelandic, and Spanish translations of the PROMIS Pain Intensity NRS-11 are available by request on the HealthMeasures website (https://www.healthmeasures.net/index.php?Itemid=992). A parent proxy form is also available. The PROMIS Pain Intensity NRS-11 was also adapted for use in Catalan-speaking children with acute pain (24,25).

**Psychometric information**

**Floor and ceiling effects.** Insufficient data are available.

**Reliability.** A systematic review of pediatric pain intensity measures identified three studies that evaluated the reliability of NRS-11 in a pediatric setting (5). The methodology of two studies was rated as poor per the COSMIN checklist and will not be described here. A third study, rated as good per the COSMIN checklist, assessed test-retest reliability in 300 children aged 4 to 17 years who presented to a hospital emergency department (26). This study focused on children who reported little to no change between pain assessments conducted at baseline and 30 to 60 minutes postbaseline. The Pearson correlation coefficient for test-retest reliability across the whole sample was reported as 0.87. Importantly, test-retest reliability was higher in children above the age of 6 years (0.82-0.97) than in children aged 4 (0.69) to 6 years (0.66) (26).

**Validity.** Tsze et al reported that the NRS-11 had strong convergent validity compared with the FPS-R (Pearson’s correlation coefficient 0.80) in a sample of children aged 4 to 17 years with acute pain (26). Sánchez-Rodríguez et al evaluated validity of an electronic NRS-11 (eNRS-11) delivered through the Painometer mobile application among 180 participants aged 12 to 19 years (7). The methods used to assess validity in this study were rated as excellent (5). Electronic and verbal NRS-11 scores did not differ by more or less than one point at 80% confidence interval (7). The eNRS-11 also showed strong convergent validity significant at a P value of less than 0.001 with the verbal NRS-11, Color Analog Scale, FPS-R, and mechanical VAS (correlation coefficients of 0.90, 0.77, 0.58, and 0.79, respectively) (7).

**Responsiveness.** The systematic review by Birnie et al identified five studies that assessed the responsiveness of NRS-11 to changes in pediatric pain (5). The methodology of two of these studies was rated as poor and fair, respectively, and will not be described here. The study by Connolly and Neville, rated as good in terms of methodology, evaluated the responsiveness of NRS-11 within a sample of 29 children aged...
9 to 18 years with postoperative pain (27). Pain scores were taken daily at five time points 3 hours apart for 3 days following surgery. Scores on the NRS-11 did not show significant responsiveness to change, remaining consistently higher than those from the FPS-R and the VAS over time. Interestingly, NRS-11 scores were positively correlated with age and intensity of anxious feelings before surgery (27). The study by Tsze et al, rated as good in terms of methodology, performed paired t-tests on 217 children’s NRS-11 scores before and after they received an analgesic for acute pain (26). They reported that the NRS-11 showed strong responsiveness to change in pain intensity even in children as young as 4 years old.

Minimally important differences. Hirschfeld et al evaluated minimally clinically important differences of the verbal NRS-11 in a sample of 153 adolescents (mean age 15.5 years) undergoing multidisciplinary rehabilitation for chronic pain (28). They reported a minimal clinically significant difference of one point change or 12.5% on the scale. They described a loss in sensitivity to change when higher cut points, such as a two-point change or 50%, were used.

Generalizability. The NRS-11 has been used in a variety of clinical populations, including in patients with acute pain, chronic pain, and postoperative pain, and with children as young as 4 years old (5). There is a weak recommendation for using the NRS-11 in children 6 years old and older with chronic pain (5).

Use in clinical trials. The NRS-11 is used in the Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry (29).

Critical appraisal of overall value to the rheumatology

The standard NRS-11 and the PROMIS Pain Intensity NRS-11 are easy to administer and score and are readily accessible, with both paper and electronic options. The psychometric properties of NRS-11 have been evaluated thoroughly in pediatric pain. Electronic delivery is unlikely to compromise the validity of scores. Pain scores from the NRS-11 are comparable with other commonly used measures (eg, the FPS-R and VAS). However, it may not provide reliable assessments of pain in children under 6 years of age (5). Based on the limited research that has been conducted, the NRS-11 may not be well suited to assessing change in pain levels over time. Further research is needed in this area.

The format of the PROMIS Pain Intensity NRS-11 is more standardized than other numerical rating scales, which may vary significantly in their anchor labels and administration. However, multiple psychometric properties have not yet been assessed in pediatric arthritis.

MEASURES OF PAIN IMPACT
PROMIS PAIN INTERFERENCE SHORT FORM

Description

Purpose. The PROMIS Pain Interference is a self-report tool used to assess pain interference.

Content or domains. The PROMIS Pain Interference is a single-domain measure of pain interference.

Number of items. The PROMIS Pain Interference consists of 20 items.

Response options/scale. Each item uses the same five-point standardized frequency scale, labeled never, almost never, sometimes, often, and almost always for points 1 to 5, respectively.

Recall period for items. The PROMIS Pain Interference has a recall period of 7 days.

Cost to use. The PROMIS Pain Interference is free to use.

How to obtain. The PROMIS Pain Interference is downloadable from the HealthMeasures website (http://www.healthmeasures.net/explore-measurement-systems/promis/obtain-administer-measures).

Practical application

Method of administration. Paper or electronic delivery options are available.

Scoring. Responses to each question are scored from 1, representing never, to 5, representing always, and then summed to yield a total raw score. The total raw score must be converted to a standardized T score using lookup tables provided by PROMIS or using the HealthMeasures scoring service (https://www.assessmentcenter.net/ac_scoring_service) or a similar score calculation service. T scores are based on an initial reference sample that included 269 patients with juvenile idiopathic arthritis; the population mean is 50, and the SD is 10 (30).

Score interpretation. A higher T score suggests higher pain interference. For example, a T score of 70 suggests a level of pain interference that is two SDs above the population mean.

Respondent time to complete. The PROMIS Pain Interference takes under 5 minutes to complete.
Administrative burden. Training is required to use the lookup tables. Scoring using lookup tables may be more time-consuming. Score calculation services may be more appropriate for large samples.

Translations/adaptation. The PROMIS Pain Interference has been translated into 15 languages, including French, Spanish, and traditional Chinese. Translations are available on request from the HealthMeasures website (https://www.healthmeasures.net/index.php?Itemid=992).

Psychometric information

Floor and ceiling effects. Insufficient data are available.

Reliability. Brandon et al found test-retest reliability correlation coefficients at 0.9 or more in a sample of 228 patients aged 8 to 17 years with pediatric arthritis (31). They also reported that the PROMIS Pain Interference had statistically significant ($P < 0.05$) ability to discriminate between patients with active and inactive arthritis.

Validity. Kashikar-Zuck et al assessed the validity of the PROMIS Pain Interference in a sample of 145 pediatric patients undergoing interdisciplinary rehabilitation for chronic pain (32). Using fixed-effects models to assess construct validity, they found that PROMIS Pain Interference scores decreased significantly over time (pain interference slope = $-0.06; P < 0.01$). This result was convergent with scores from the Functional Disability Inventory (FDI) (slope = $-0.72; P < 0.01$).

Responsiveness. Kashikar-Zuck et al used unconditional growth models to demonstrate that the PROMIS Pain Interference captures significant responsiveness to change and improvement in pain interference levels (32).

Minimally important differences. The minimally important differences of the PROMIS Pain Interference are not yet clearly established. Morgan et al surveyed pediatric rheumatologists as well as adolescents with arthritis and their parents to establish minimally important differences for the PROMIS Pain Interference in pediatric arthritis (33). Cut scores tended to vary greatly between groups. Clinicians tended to rate even lower levels of pain interference more severely than patients and parents did.

Generalizability. The PROMIS Pain Interference has been used in pediatric rheumatology and pediatric chronic pain and across a wide range of pediatric age groups. However, no specific recommendations for use from PROMIS are currently available.

Use in clinical trials. The PROMIS Pain Interference is used in the Childhood CARRA registry (29). See also the article by Lee et al (34).

Critical appraisal of overall value to the rheumatology community

The PROMIS Pain Interference Short Form is readily accessible with both paper and electronic options and multiple official translations. The PROMIS Pain Interference can also be administered using computerized adaptive testing (CAT), an algorithm-based method in which participant responses guide the system’s choice of subsequent items from the full item bank (35). Although the PROMIS Pain Interference Short Form requires patients to respond to all 20 items, the CAT version only requires 5 to 12 item responses in order to calculate a total pain interference score. Therefore, the CAT version may be more efficient, given that it provides automatic score conversion, and may provide more accurate assessments of pain interference owing to its scaling feature. Pragmatically, the use of CAT requires a computer and specialized software, whereas the Short Form may be easier to administer in clinical practice. In terms of scoring, the Short Form version may be the most time-consuming, given the need for manual score conversion. Although studies in pediatric chronic pain suggest that the PROMIS Pain Interference is able to detect change in pain interference over time, more research on the psychometric properties of both the PROMIS Pain Interference Short Form and CAT in pediatric arthritis populations is needed.

**BATH ADOLESCENT PAIN QUESTIONNAIRE (BAPQ)**

Description

**Purpose.** The BAPQ is a multidimensional measure that assesses the impact of pain on overall life.

**Content or domains.** The BAPQ has seven domains: the impact of pain on social function, physical function, depression, generalized anxiety, pain-specific anxiety, family function, and development.

**Number of items.** The BAPQ has 61 items in total.

**Response options/scale.** All domains except for development have a five-point Likert scale (never, hardly ever, sometimes, often, and always). Development is scored on a separate five-point Likert scale (very behind, a little behind, same, a little ahead, and very ahead).

**Recall period for items.** The BAPQ has a recall period of 2 weeks.
**Cost to use.** The BAPQ is free to use.

**How to obtain.** The BAPQ is available for download from the Bath Centre for Pain website (https://www.bath.ac.uk/publications/bath-adolescent-pain-questionnaire-bapq/).

**Practical application**

**Method of administration.** The BAPQ is a self-report measure that can be administered by paper.

**Scoring.** All seven domains are scored separately, and some domains are grouped together to form a total of four subscales (daily functioning, emotional functioning, family functioning, and development). Calculating a total score for the BAPQ is not recommended. All 61 items in the measure are weighted equally, with a minimum of zero points and a maximum of four points per item. In each domain, a number of items require reverse scoring. After reverse scoring is completed, the domain score is calculated through summation. Subscales of daily functioning and emotional functioning require the administrator to combine scores from several domains.

**Score interpretation.** In all four subscales of the BAPQ, higher scores indicate more impaired functioning.

**Respondent time to complete.** The BAPQ takes 10 to 15 minutes to complete.

**Administrative burden.** Scoring is manual for the paper version, and it is time-consuming because it requires some reverse scoring and combining scores from grouping domains to achieve subscale scores.

**Translations/adaptation.** The BAPQ was developed in English and has not been translated to other languages. The BAPQ paper version has been adapted for electronic use with a computerized version, although this is not widely available (36). The BAPQ has also been adapted for parent proxy report, assessing parents’ perceptions of the impact of pain on their adolescent’s overall life (37).

**Psychometric information**

**Floor and ceiling effects.** Insufficient data are available.

**Reliability.** There are two known studies examining the psychometric properties of the BAPQ. Eccleston et al examined the psychometric properties of the BAPQ in a population of 222 adolescents (aged 11-18 years) with chronic pain (38). Participants were recruited from two outpatient pediatric clinics focused on rheumatology and chronic pain, respectively. Of the total sample, 21.2% of participants had juvenile idiopathic arthritis; other diagnoses included headache, low back pain, and complex regional pain syndrome. Test-retest reliability was evaluated in a subset of 30 participants over a 17-day period. Reliability coefficients were calculated for each of the seven domains and separated between the rheumatology participants and those with chronic pain. Within the rheumatology sample, test-retest reliability ranged from 0.66 to 0.94 across the domains, with high correlations for the social function, physical function, and pain-specific anxiety domains. Internal consistency was also evaluated by subscale interrelatedness, with coefficients ranging from 0.34 to 0.79 in the rheumatology sample. The second study by Gagnon et al examined the psychometric properties in 110 children and adolescents (aged 8-18 years) in a pain management program (39). This study also examined subscale interrelatedness, with subscale coefficients ranging from 0.81 to 0.87.

**Validity.** Convergent validity was explored in two studies. Eccleston et al found good convergent validity when comparing specific BAPQ domains with existing measures of pain and function (for example, $r = 0.77$ for the physical function domain compared with the FDI, and $r = 0.81$ for the pain-specific anxiety domain with the Pain Catastrophizing Scale [PCS]). No domains correlated with pain severity (38). Gagnon et al found similar results with correlations of the BAPQ subscales with the FDI, PCS, Children’s Depression Inventory, and Spence Children’s Anxiety Scale. Neither study examined convergent validity related to the development subscale (39).

**Responsiveness.** Insufficient data are available.

**Minimally important differences.** Insufficient data are available.

**Generalizability.** There is limited research examining the psychometric properties of this tool in children with arthritis-related pain. One study examined the reliability and validity in adolescents with rheumatic disease, particularly juvenile idiopathic arthritis (38). Original psychometric testing was completed in children from the United Kingdom, and another study was able to reproduce similar results with reliability and validity in children from the United States (36,37). There is no evidence to support using this tool in children younger than 8 years of age.

**Use in clinical trials.** The BAPQ has been used in several clinical studies involving children with chronic pain (40–43). It has not been used in any randomized controlled trials or trials in the arthritis population.
Critical appraisal of overall value to the rheumatology community

Strengths of the BAPQ include its multidimensional nature, paper and electronic options, accessibility, and generalizability to the pediatric arthritis population. Cautions include interpretability of scoring and its unknown ability to detect change over time.

CONCLUSIONS

Pain is a multifaceted construct, and its assessment can be used to guide treatment decisions. Therefore, in clinical settings, it is important to measure pain itself as well as the impact of pain on different aspects of function. A clinician’s decision to use one measurement tool over another may be influenced by numerous factors. For example, a unidimensional tool such as the PROMIS Pain Interference may be favored for routine pain assessment, whereas a multidimensional tool such as the BAPQ may be better suited for children with persistent pain. Clinicians should also consider the accessibility of measurement tools, the amount of time available for assessment, and the type of pain to be assessed (eg, acute versus chronic) as well as the age or developmental level of the patient.

All the measures presented herein are widely accessible and free to use. They can all be administered electronically as well as on paper. Parent proxy forms are also available for the BAPQ and PROMIS measures. Parent proxies allow the parental perspective to be part of the assessment and are a suitable option if the patient is unable to provide a self-report. See Tables 1 and 2 for a summary of the measures.

There are key limitations of the field to highlight. The psychometric properties of the measures discussed have largely not been assessed in pediatric arthritis. Furthermore, some measures are more complicated and time-consuming to score. The PROMIS Pain Interference requires raw scores to be converted into T scores, and the BAPQ requires manually computing subscale scores through reverse scoring and score compilation.

Numerous measurement tools exist for the purposes of pain assessment. The VAS, FPS-R, PROMIS Pain Intensity NRS-11, PROMIS Pain Interference, and BAPQ are presented here as five examples of tools that clinicians can consider for assessing pain in their patients with pediatric arthritis. Information about each tool has been provided in this article through individual descriptions as well as in summary Tables 1 and 2. Clinicians are encouraged to consult these descriptions and compare the measures in order to select the most appropriate measure for routine pain assessment in their pediatric patients with arthritis.

AUTHOR CONTRIBUTIONS

All authors drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Kimura Y, Walco GA, Sugarman E, Conte PM, Schanberg LE. Treatment of pain in juvenile idiopathic arthritis: a survey of pediatric rheumatologists. Arthritis Rheum 2006;55:81–5.
2. Bromberg MH, Connelly M, Anthony KK, Gil KM, Schanberg LE. Self-reported pain and disease symptoms persist in juvenile idiopathic arthritis despite treatment advances: an electronic diary study. Arthritis Rheumatol 2014;66:462–9.
3. Lee RR, Rashid A, Thomson W, Cordingley L. “Reluctant to assess pain”: a qualitative study of health care professionals’ beliefs about the role of pain in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2020;72:69–77.
4. Lovell DJ, Passo MH, Beukelman T, Bowyer SL, Gottlieb BS, Henrickson M, et al. Measuring process of arthritis care: a proposed set of quality measures for the process of care in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63:10–6.
5. Birnie KA, Hundt AS, Laloo C, Nguyen C, Stinson JN. Recommendations for selection of self-report pain intensity measures in children and adolescents: a systematic review and quality assessment of measurement properties. Pain 2019;160:5–18.
6. De la Vega R, Roset R, Castarlenas E, Sánchez-Rodríguez E, Solé E, Miró J. Development and testing of painometer: a smartphone app to assess pain intensity. J Pain 2014;15:1001–7.
7. Sánchez-Rodríguez E, de la Vega R, Castarlenas E, Roset R, Miró J. An app for the assessment of pain intensity; validity properties and agreement of pain reports when used with young people. Pain Med 2015;16:1982–92.
8. Sánchez-Rodríguez E, Castarlenas E, de la Vega R, Roset R, Miró J. On the electronic measurement of pain intensity: can we use different pain intensity scales interchangeably? [Original Report]. J Health Psychol 2017;22:1658–67.
9. Newman CJ, Lolekha R, Limkittikul K, Luangxay K, Chotpitayasunondh T, Chanthanavanich P. A comparison of pain scales in Thai children. Arch Dis Child 2005;90:269–70.
10. Algós. Painometer v2. URL: http://app.dolorinfantil.urv.cat/wp-conent/uploads/2014/02/painometer_brochure.pdf.
11. Terver CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651–7.
12. Bailey B, Gravel J, Daoust R. Reliability of the Visual Analog Scale in children with acute pain in the emergency department. Pain 2012;153:839–42.
13. Luffy R, Grove SK. Examining the validity, reliability, and preference of three pediatric pain measurement tools in African-American children. Pediatr Nurs 2003;29:54–9.
14. Le May S, Ballard A, Khadra C, Gouin S, Plint AC, Villeneuve E, et al. Comparison of the psychometric properties of 3 pain scales used in the pediatric emergency department: Visual Analogue Scale, Faces Pain Scale-Revised, and Colour Analogue Scale. Pain 2018;159:1508–17.
15. Powell CV, Kelly AM, Williams A. Determining the minimum clinically significant difference in visual analog pain score for children. Ann Emerg Med 2001;37:28–31.
16. Hanns L, Cordingley L, Galloway J, Norton S, Carvalho LA, Christie D, et al. Depressive symptoms, pain and disability for adolescent patients with juvenile idiopathic arthritis: results from the Childhood Arthritis Prospective Study. Rheumatology (Oxford) 2018;57:1381–9.
17. Norgaard M, Twilt M, Andersen LB, Herlin T. Accelerometry-based monitoring of daily physical activity in children with juvenile idiopathic arthritis. Scand J Rheumatol 2016;45:179–87.
18. Tsze DS, von Baeyer CL, Bulloch B, Dayan PS. Validation of self-report pain scales in children. Pediatrics 2013;132:e971–9.
19. Hicks CL, von Baeyer CL, Spafford PA, van Kortlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. Pain 2001;93:173–83.
20. McGrath PA, Seifert CE, Speechley KN, Booth JC, Stitt L, Gibson MC. A new analogue scale for assessing children's pain: an initial validation study. Pain 1996;64:435–43.
21. Tsze DS, Hirschfeld G, von Baeyer CL, Bulloch B, Dayan PS. Clinically significant differences in acute pain measured on self-report pain scales in children. Acad Emerg Med 2015;22:415–22.
22. Lomholt JJ, Thastum M, Christensen AE, Leegaard A, Herlin T. Cognitive behavioral group intervention for pain and well-being in children with juvenile idiopathic arthritis: a study of feasibility and preliminary efficacy. Pediatr Rheumatol Online J 2015;13:35.
23. HealthMeasures. Pain intensity. 2017. URL: http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Pain_Intensity_Scoring_Manual.pdf.
24. Castarlenas E, Sánchez-Rodríguez E, Vega Rde L, Roset R, Miró J. 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;31:229–34.
25. 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? [Original Report]. J Pain 2013;14:297–304.
26. Tsze DS, von Baeyer CL, Pahalyants V, Dayan PS. Validity and reliability of the verbal numerical rating scale for children aged 4 to 17 years with acute pain. Ann Emerg Med 2018;71:691–702.
27. 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;11:1451–60.
28. Hirschfeld G, Wager J, Schmidt P, Zernikow B. Minimally clinically significant differences for adolescents with chronic pain-variability of ROC-based cut points. J Pain 2014;15:32–9.
29. Beukelman T, Kimura Y, Ilowite NT, Mieszkalski K, Natter MD, Burrell MC. A new analogue scale for assessing children's pain: an initial validation study. Pediatr Rheumatol Online J 2017;15:30.
30. Quinn H, Thissan D, Liu Y, Magnus B, Lai JS, Amtmann D, et al. Using item response theory to enrich and expand the PROMIS® pediatric self report banks. Health Qual Life Outcomes 2014;12:160.
31. Brandon TG, Becker BD, Bevans KB, Weiss PF. Patient-Reported Outcomes Measurement Information System tools for collecting patient-reported outcomes in children with juvenile arthritis. Arthritis Care Res (Hoboken) 2017;69:398–402.
32. Kashkar-Zuck S, Carle A, Barnett K, Goldschneider KR, Sherry DD, Mara CA, et al. Longitudinal evaluation of Patient-Reported Outcomes Measurement Information Systems measures in pediatric chronic pain. Pain 2016;157:339–47.
33. Morgan EM, Mara CA, Huang B, Barnett K, Carle AC, Farrell JE, et al. Establishing clinical meaning and defining important differences for Patient-Reported Outcomes Measurement Information System (PROMIS®) measures in juvenile idiopathic arthritis using standard setting with patients, parents, and providers. Qual Life Res 2017;26:565–86.
34. Lee RR, Shoop-Worrall S, Rashid A, Thomson W, Cordingley L. “Asking too much?”: A randomised N-of-1 trial exploring patient preferences and measurement reactivity to frequent use of remote multi-dimensional pain assessments in children and young people with Juvenile Idiopathic Arthritis. JMIR 2019;22:e14503.
35. HealthMeasures. Pain interference. 2019. URL: http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Pain_Interference_Scoring_Manual.pdf.
36. Jordan A, Begem FM, Austin L, Edwards RT, Connell H. A usability and feasibility study of a computerized version of the Bath Adolescent Pain Questionnaire: the BAPQ-C. BMC Pediatr 2020;20:6.
37. Jordan A, Eccleston C, McCracken LM, Connell H, Clinch J. The Bath Adolescent Pain-Parental Impact Questionnaire (BAP- PIQ): development and preliminary psychometric evaluation of an instrument to assess the impact of parenting an adolescent with chronic pain. Pain 2008;137:478–87.
38. Eccleston C, Jordan A, McCracken LM, Sleed M, Connell H, Clinch J. The Bath Adolescent Pain Questionnaire (BAPQ): development and preliminary psychometric evaluation of an instrument for assessing the impact of chronic pain in adolescents. Pain 2009;118:263–70.
39. Gagnon C, Amstutz D, Revivo G. Psychometric evaluation of the Bath Adolescent Pain Questionnaire (BAPQ) in a US-based clinical population. J Pain 2011;12 Suppl:17.
40. Revivo G, Amstutz DK, Gagnon CM, McCormick ZL. Interdisciplinary pain management improves pain and function in pediatric patients with chronic pain associated with joint hypermobility syndrome. PM R 2019;11:150–7.
41. Benore E, D’Auria A, Banez GA, Worley S, Tang A. The influence of education on clinical response to pediatric chronic pain rehabilitation. Clin J Pain 2015;31:375–83.
42. Cohen LL, Vowles KE, Eccleston C. Adolescent chronic pain-related functioning: concordance and discordance of mother-proxy and self-report ratings. Eur J Pain 2010;14:862–6.
43. Benore E, Brenner A, Banez GA, Wang L, Worley S. It takes two: parent functioning within the pediatric chronic pain experience and interdisciplinary rehabilitation treatment. Rehabil Psychol 2018;63:205–14.
### Table 1. Practical applications*

| Measure                        | Number of Items | Content/ Domains                                                                 | Method of Administration | Recall Period | Response Format                                                                 | Range of Scores  | Score Interpretation            | Availability of Normative Data | Cross-Cultural Validation |
|-------------------------------|----------------|---------------------------------------------------------------------------------|--------------------------|---------------|---------------------------------------------------------------------------------|------------------|--------------------------------|------------------------------|-----------------------------|
| Visual Analog Scale           | 1              | Pain intensity                                                                  | Paper or electronic     | Typically used to assess current pain | Single 100-mm line anchored by no pain and worst pain possible; respondent marks their pain level on the line | 0-10 or 0-100   | Higher score indicates greater pain intensity. | Not available                | Available in Catalan, French, Thai, Portuguese, and Spanish |
| Faces Pain Scale Revised      | 1              | Pain intensity                                                                  | Paper or electronic     | Used to assess current pain           | Six drawn human faces, each representing a different level of pain; respondent selects the face that best represents their pain level | 0-10            | Higher score indicates greater pain intensity. | Not available                | Available in 68 different languages apart from English, including French, Spanish, and Chinese |
| NRS-11                        | 1              | Pain intensity                                                                  | Paper or electronic     | 7 days                                   | Eleven-point numerical rating scale (0 = no pain; 10 = worst pain you can think of) | 0-10            | Higher score indicates greater pain intensity. | Not available                | Used to assess pain in Catalan-speaking children |
| PROMIS Pain Interference      | 20             | Pain interference                                                              | Paper or electronic     | 7 days                                   | Frequency response (1 = never, 2 = almost never, 3 = sometimes, 4 = often, and 5 = almost always) | Raw scores range from 1-5; raw scores converted to T scores (minimum = 0; maximum = 100) | Higher T score indicates greater pain interference. | Available (30); see manual (34) provided by PROMIS for score conversion | - |
| Bath Adolescent Pain Questionnaire | 61          | Four subscales (daily functioning, emotional functioning, family functioning, and development) | Paper or electronic     | 2 weeks                                  | Frequency response (0 = never; 4 = always) except for the development scale (0 = very behind; 4 = very ahead) | Minimum subscale score = 0; maximum subscale score = 44-80 | Higher scores indicate more impaired function. | Not available                | Not available |

* NRS-11 = Eleven-Point Numerical Rating Scale; PROMIS = Patient-Reported Outcomes Measurement Information System.
| Measure                          | Floor/Ceiling Effects | Reliability                                                                 | Validity                                                                 | Responsiveness                                                                 | Minimally Important Differences | Generalizability | Used in RCTs |
|---------------------------------|-----------------------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------------------------------|------------------|--------------|
| Visual Analog Scale             | Not available         | Test-retest reliability assessed in children with painful conditions; no studies in pediatric arthritis. | Convergent validity and discriminant validity assessed in school children; no studies in pediatric arthritis. | Assessed in children presenting to an emergency department; no studies in pediatric arthritis. | 10-12 mm                       | Weak recommendation for use in children aged ≥6 years with persistent pain; weak recommendation against use in children aged <6 years (5) | Yes          |
| Faces Pain Scale Revised        | Not available         | Test-retest reliability assessed in children with painful conditions; no studies in pediatric arthritis. | Convergent validity assessed in children with painful conditions, including arthritis. | Assessed in children with existing painful conditions; no studies in pediatric arthritis. | One face is considered the minimal clinically significant difference, whereas two is considered the ideal clinically significant difference | Weak recommendation against using this tool in children aged ≥4 years with persistent pain (5) | Yes          |
| NRS-11                          | Not available         | Test-retest reliability assessed in children with acute pain; no studies in pediatric arthritis. | Convergent validity assessed in children with acute pain and postoperative pain; no studies in pediatric arthritis. | Assessed in children with acute pain and postoperative pain; no studies in pediatric arthritis. | 1 point or 12.5%                  | Weak recommendation for using this tool in children aged ≥6 years | Yes          |
| PROMIS Pain Interference        | Not available         | Test-retest reliability and discriminative ability assessed in pediatric arthritis. | Construct validity assessed in pediatric chronic pain. | Responsiveness assessed in pediatric chronic pain. | Not yet clearly established; cut points found to vary between patients, parents, and clinicians in pediatric arthritis | No specific recommendations for use currently available | Yes          |
| Bath Adolescent Pain Questionnaire | Not available       | Evaluated subscale interrelatedness, test-retest reliability variable across domains | Evaluated convergent validity between subscales and FDI, CDI, PCS, and SCAS. | Not available | Not available | -- | No          |

* CDI = Children's Depression Inventory; FDI = Functional Disability Inventory; NRS-11 = Eleven-Point Numerical Rating Scale; PCS = Pain Catastrophizing Scale; PROMIS = Patient-Reported Outcomes Measurement Information System; RCT = randomized controlled trial; SCAS = Spence Children's Anxiety Scale.