Measuring Pediatric Physical Function

Nancy L. Young, B.Sc.P.T., M.Sc., and James G. Wright, M.D., M.P.H., F.R.C.S.C.

Division of Orthopaedic Surgery and Clinical Epidemiology Unit, The Hospital for Sick Children, Toronto, Ontario, Canada

Summary: Most pediatric orthopaedic interventions are intended to improve or preserve physical function, yet their outcomes have been assessed using primarily surrogate measures (e.g., radiographic indices) that may not accurately represent patients' function. Physical function may be more appropriately measured with activity-based scales, but these have been infrequently applied in surgical studies. The purpose of this study was to identify existing activity-based physical-function scales appropriate for pediatric orthopaedics, to present criteria useful for scale selection, and to discuss the special problems of measuring physical function in children. Twenty-one scales relevant to pediatric orthopaedics are described according to their target population, purpose, method of administration, content, and quality of standardization. These scales have been further classified according to a new taxonomy. The unique aspects of measuring physical function in children are discussed and include the effect of age and development, method of reporting, and question formats. Standardized measures of physical function based on physical-activity ability exist and should be used more frequently to assess pediatric orthopaedic interventions. Key Words: Measurement—Orthopaedics—Outcome assessment—Physical function.

Most pediatric orthopaedic treatment strategies are intended either to reduce existing physical disability (which has been defined as limitation in physical activity (65)) or to prevent future physical disability. Orthopaedic surgeons, however, have relied primarily on surrogate measures of physical function (e.g., range of motion and radiographic measurement) to assess the effectiveness of interventions. Surrogate measures are popular because they are easy to use, bear immediate relevance, and appear to be reliable. However, surrogate measures are not necessarily reliable (69), and more important, may not accurately represent changes in physical disability [e.g., despite radiographic curve correction in children with spina bifida, walking ability may decrease (40)]. Thus, although surrogate measures may be important in day-to-day clinical management, physical-function measures based on activities are arguably the most cogent outcomes of orthopaedic interventions.

Many activity-based physical-function scales capable of assessing the effects of clinical management have been developed in the rehabilitation field but have not been widely used by orthopaedic surgeons. Surgeons' infrequent use of such scales may be attributed to a lack of awareness of the existence of appropriate scales, difficulty obtaining scales, or unfamiliarity with the factors to be considered in selecting scales. In addition, because measurement of pediatric function has many special problems, such as the impact of growth and development, which few of the existing scales have completely addressed, surgeons may have deemed these scales inadequate for inclusion in clinical practice or research.

The purpose of this article is to aid surgeons in using activity-based measures by (a) cataloguing activity-based pediatric physical-function measures with direct application to pediatric orthopaedic practice; (b) explaining the options relevant to selection of appropriate scales; and (c) discussing some of the special problems of measuring physical function in pediatrics.

Physical function is defined in this article as the ability to use the musculoskeletal system to interact with the environment in a purposeful way for the performance of activities of daily living, mobility (e.g., manual dexterity, transfers, ambulation), and leisure activities (16,42). Physical function is a distinct subcomponent of more global health measures such as functional status (16), health status (1,42,52,56), and quality of life (42,63). Because of the nature of orthopaedic interventions, the physical-function components of patients' status is the most...
MEASURING PHYSICAL FUNCTION

likely to be affected and therefore is the focus of this article.

Standardization refers to the reliability, validity, and the responsiveness or discriminative ability of a measure. Reliability (reproducibility or consistency) is the degree to which the scale will yield similar answers when the measure is repeated (60). Validity (or accuracy) is the extent to which the scale measures what it is intended to measure. Scales intended to evaluate change must also be able to detect clinically important change, which is termed responsiveness (or sensitivity). Scales intended to measure patients' status must be able to classify subjects correctly, which is termed discrimination and is quantified in terms of specificity and sensitivity (51).

METHODS

Physical-function measures were identified for review from three sources: a computerized literature search, the experience and resources of authors and colleagues, and the reference lists of publications identified using the first two strategies. The computer searches were conducted using the MEDLINE Index for the years 1984 to March 1993 and the allied health literature from 1984 to February 1992.

Scales were included in the review if the primary intent of the scale was to quantify activity-based physical function. Scales were excluded if they were developed for adult populations without specific documentation of pediatric application or were developmental milestone inventories (assuming a "normal" developmental sequence). Developmental scales have been extensively reviewed elsewhere (8,12-14,20-22,27,43,44,58,61,62); however, one example has been included in the Appendix to promote an understanding of where these scales fit into the taxonomy.

RESULTS

The literature review identified 147 articles. The vast majority of the scales focused on developmental assessment (e.g., Hughes Basic Gross Motor Assessment), psychiatric or behavioral assessment (Child Behavior Check List), and intellectual aptitude (Weschler Scales, Illinois Test of Psycholinguistic Abilities), as has been previously reported (32). Twenty-one scales relevant to pediatric orthopaedics remained after exclusions: 13 were developed primarily for a pediatric population; four are adult scales modified for a pediatric population; and four are adult scales that have not been modified for children but have been used to measure pediatric physical function. These 21 scales are presented in the Appendix; however, the Appendix is not expected to be exhaustive.

For ease of use, the Appendix is subdivided into three sections: scales for children with neurological impairments, scales for children with arthritis, and scales that are not specific to any disease (also called generic scales). Within each section, the scales are listed according to a taxonomic classification. Column 1 lists names of the scales and provides references intended to assist the reader in the further investigation of pertinent scales. In several cases, a publication of the original scale development was not found; thus, the reference cited may be not the original author but a secondary source that presents a sufficient degree of detail for those interested in pursuing this scale further. Columns 2 to 7 are intended to aid surgeons in choosing among the available scales by describing the scales according to criteria that should be considered in the selection of appropriate scales and are described in detail in the following paragraphs.

Purpose of the measure

Column 2 discusses the intended use of the measure. Scales may be discriminative (distinguish between groups of patients), evaluative (detect change), or predictive (forecast the results of subsequent evaluations) (28). In clinical practice, surgeons are most often interested in evaluating the effect of interventions and thus predominantly seek evaluative scales. A scale developed for one purpose is not necessarily valid if used for a different purpose or population.

Domains, format, and scoring

Column 3 lists the functional domains that the scale covers, the format of the questions (e.g., visual analog or categorical), and how the questions are scored and aggregated. In some instances, full information was not available.

Population

Column 4 describes the population(s) on which the scale has been applied. Note that a scale's discriminative ability may be specific to the age or disease group for which it was developed. The measure must also be capable of scoring the full range of ability and disability expected in the study population.

Method of administration

Column 5 describes the details of administering the measure, including by whom (e.g., clinician or self-administered), how (e.g., observation or interview), to whom (e.g., child, parent, or proxy reporter), and the time and special equipment requirements.

Standardization of the measure

Column 6 contains details of the published reliability, validity, and responsiveness testing of the scales. The specifics of each scale's standardization testing can be found in the references cited in the Appendix and should be reviewed before selecting a scale (9,46,60,68,69). One must seek a scale that offers sufficient standardization, bearing in mind
that few, if any, scales have been completely standardized and that use of untested or inappropriate outcome variables will decrease the likelihood of detecting clinically important differences.

**Taxonomic classification**

Column 7 classifies the scales according to a new taxonomy (organizational framework), and the classification number is recorded in brackets in the Appendix. The taxonomy was created to simplify functional scale selection for surgeons and clinical researchers and is shown in Table 1. Under the taxonomy, scales are first classified by whether they are direct or indirect measures and then as capability or performance measures.

Direct measures are scored on the basis of personal observation of an activity or behavior. Capability measures assess what the child can do. Performance measures assess what the child does do. Finally, indirect measures are further subdivided according to method of administration into self-report scales and interviewer-administered scales.

Examples of physical-function measures of each of the four main types are as follows. The play-performance scale (31), which requires that a clinician directly observe skills the child uses spontaneously in a normal play environment, is a direct-performance measure (1.1). Clinical gait assessment requires a clinician’s direct observation of what the child is able to do when it is demanded of him/her, and thus is a direct-capability measure (1.2). The Klein-Bell ADL scale is an indirect-performance measure, because it is scored based on report of previously observed spontaneous activity in a normal environmental context (2.1). Finally, surveys of what the child can do under hypothetical circumstances (often ideal circumstances) are indirect-capability measures (2.2).

Direct methods may be considered more valid because they eliminate the biases of the reporter but may be sensitive to environmental changes and are rarely practical. Indirect methods offer enhanced feasibility, may have greater consistency of administration, and if self-administered, eliminate interviewer bias. The difficulties inherent in self-report measures include uncertain comprehension or interpretation of the questions and response bias (59).

**DISCUSSION**

Physical-function measures are infrequently used by pediatric orthopaedists, possibly because of lack of awareness of existing scales, limited availability of the scales, difficulties in making appropriate selections, or because of the special problems inherent in the measurement of pediatric function. The identification and description of 21 scales that are potentially appropriate for pediatric orthopaedics attempt to address the first two obstacles. We hope the taxonomy presented in this article and the discussion of criteria for scale selection will aid surgeons in choosing among the available scales. Finally, the special challenges of measuring physical function in children must be addressed. Clinical application of physical-function scales should offer substantial benefits to research, provided there is some recognition of the effects of age, growth and development, the impact of the parent in reporting, and the framing of questions and response options.

**Age, growth, and development**

Prime considerations when evaluating the appropriateness of a pediatric scale for a specific population are the age for which the scale is applicable and the effect of development on sequential scores. Due to development, age has a distinct impact on ability to perform certain activities and on their relative importance. For example, tricycle riding is an important part of physical function at age 4 but not at age 8, even though the motor skills required are still present.

At least two methods may accommodate for the effects of age. First, a comprehensive scale may be developed that covers physical function across all age groups, such as the Rand Health Insurance Scale. This method is simple because only one scale is required for all children, but it may not be responsive to clinically important change. A variation of this method is to have a single scale but correct for the effect of development on sequential scores. Due to some recognition of the effects of age, growth and development, the impact of the parent in reporting, and the framing of questions and response options.

**TABLE 1. Physical function measure taxonomy**

<table>
<thead>
<tr>
<th>1. Direct measurement (clinical observation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Performance based (does do)</td>
</tr>
<tr>
<td>1.2 Capability based (can do)</td>
</tr>
<tr>
<td>2. Direct measurement (report of parent, patient, or proxy)</td>
</tr>
<tr>
<td>2.1 Performance based (does do)</td>
</tr>
<tr>
<td>2.11 Interviewer administered</td>
</tr>
<tr>
<td>2.12 Independent/self-administered</td>
</tr>
<tr>
<td>2.2 Capability based (can do)</td>
</tr>
<tr>
<td>2.21 Interviewer administered</td>
</tr>
<tr>
<td>2.22 Independent/self-administered</td>
</tr>
</tbody>
</table>

*Scales must be selected on the basis of their purpose and population. It was not feasible to put this into the taxonomy, as there is considerable overlap in purposes and populations.*
on the same scale, a 5-year-old might have a maximum of 120 because of advanced motor skill level. If a child scores 80 at age 4, has surgery, and subsequently scores 96 at age 5 (both standardized scores are equal to 80% of expected), then no improvement can be attributed to the intervention beyond that which would have occurred with development. (Note, however, that the absence of a decline may be a clinically important finding indicative of success.) The process of age-adjusted scores requires normative data on the populations in question, and expected scores for disabled children are rarely available.

The second option is to use scales that are appropriate for limited age groups. This, however, requires multiple scales to accommodate various age groupings and makes measuring the effect of an intervention in children who cross over into a new age category during the trial very difficult. Thus age-specific scales are not recommended unless a translation between scales for different age groups has been clearly determined.

Self, parent, and proxy report

If the decision has been made to evaluate physical function indirectly, then the source of information must be selected: child, parent, or a proxy. When the focus of the intervention and research is the child, then the child should be the source of information. Parent report is required for patients whose communicative capacity is impaired by age, illness, or cognitive ability. Alternatively, proxy report can be used and may be advantageous when strong parental bias is suspected.

Context

The environmental conditions are particularly important when measuring physical function because they define whether capability or performance is being measured and affect the outcome (e.g., the degree of motivation, environmental distractions, and the presence of parents may significantly affect children's physical function). Additionally, physical function can be measured in multiple ways depending on the wording of the questions. Questions may ask about quality or quantity of function, each potentially yielding a different outcome. For example, physical function can be measured on a scale of independence, which can be affected by physical function, availability of supports, and willingness to accept assistance. The social construction of childhood is such that most children have readily available supports and may also be willing to accept assistance; therefore, independence measures may overestimate children's disability. Thus physical-function measures should not be adopted without consideration of contextual issues. Children's ability to comprehend certain question formats (such as visual analog scales) may also change as a function of age and requires consideration.

In summary, this article has addressed the problems of availability, difficulties in selecting appropriate measures, and conceptual and methodological issues unique to measuring physical function in children. Appropriate scales can be selected using the references, standards, and taxonomy provided. Clinicians are encouraged to include activity-based function outcome measures in clinical and research practice, provided that they evaluate the existing scales carefully with regard to population, purpose, and standardization.

Future research will be required to determine the relationship between performance and capability, the agreement between parents and children, and the preferred context to measure physical function. These issues do not have a single correct answer, but none of the difficulties precludes the use of these activity-based measures. Finally, because orthopaedic interventions are intended to improve (or maintain) function, evaluations of surgical therapy should include measures of physical function, which can then be interpreted on the basis of clinical expectations.

REFERENCES

13. Gesell Institute of Human Development. The Gesell Pre-


33. Law M. Copy of the Klein-Bell ADL Scale with age norms applied. Hamilton, Ontario: McMaster University, 1992.


53. Singh G. Copy of Childhood Health Assessment Questionnaire and guidelines. Palo Alto: Stanford University School of Medicine, 1984.


MEASURING PHYSICAL FUNCTION


APPENDIX

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Purpose of scale</th>
<th>Domains, format, &amp; scoring</th>
<th>Method of administration</th>
<th>Standardization of measure</th>
<th>Taxonomic class (as per Table 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional Independence Measure for Children (WeeFIM)</td>
<td>Burden of care; discriminative, evaluable</td>
<td>Degree of assistance required (provided by a caregiver or assistive device)</td>
<td>Trained clinician observation</td>
<td>No patient data reported. Developers state that face validity and reliability were established in &gt;50 facilities but no reference cited.</td>
<td>Clinician observation of capability</td>
</tr>
<tr>
<td>Motor Control Assessment (MCA) (57)</td>
<td>Motor control skills (not functional ability); evaluable</td>
<td>2 to 5-yr olds</td>
<td>Clinician observation</td>
<td>Validity: correlation with Physical Abilities score = 0.9 Reliability: ICCs: internist = 0.99, for internist = 0.97</td>
<td>Clinician observation of capability</td>
</tr>
<tr>
<td>Tufts Assessment of Motor Performance (TAMP) (12)</td>
<td>Physical function and motor performance; evaluable</td>
<td>6 yrs upward including adults (161, primarily neurologically impaired)</td>
<td>Clinician observation</td>
<td>Intrarater reliability using a videotaped assessment exceeded 0.85 (ICC) for all domains/dimension combinations Factor analysis of data on 206 subjects used to determine empirically item groupings: dynamic balance, fasteners, ambulation, manipulation, mat mobility, typing, grasp/release</td>
<td>Clinician observation of capability</td>
</tr>
</tbody>
</table>

(Continued)
### Scale name | Purpose of scale | Domains, format, & scoring | Population | Method of administration | Standardization of measure | Taxonomic class (as per Table 1) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Klein-Bell ADL scale</strong> (14,36,37); original adult scale cited as (29,30)</td>
<td>ADL function; evaluative, discriminative</td>
<td>6 domains: dressing, bathing/hygiene, elimination, functional mobility, eating, emergency communication</td>
<td>All ages</td>
<td>Test population 10 CP and 10 normals</td>
<td>Clinician observation ~1 h to administer all items</td>
<td>Validity: discriminated between normal and CP subjects p &lt; 0.001</td>
</tr>
<tr>
<td><strong>Barthel Index</strong> (14,36, Original ref. cited as (39))</td>
<td>Activities of daily living; discriminative, predictive, evaluative</td>
<td>ADLs, Ordinal scale</td>
<td>Applied to adult and adolescent chronically disabled patients</td>
<td>Expert clinician observation 1 h to complete</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Karnofsky Scale</strong> (23,25,45)</td>
<td>Global rating of physical capacity; evaluative, predictive</td>
<td>Based primarily on mobility level</td>
<td>Undescribed cancer population (generally poor description of samples)</td>
<td>Physician report 2 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vineland Adaptive Behavior Scales</strong> (14, Original ref. cited as (55))</td>
<td>Developmental Assessment Tool Included as example of developmental scale classification</td>
<td>4 domains: communication, daily living skills, socialization, and motor skills (impairment)</td>
<td>0-18 yrs</td>
<td>Normative data based on a large sample of disabled children</td>
<td>Trained clinician interview of parent 20-60 min</td>
<td></td>
</tr>
<tr>
<td><strong>Quality of Life</strong> (47); Original ref. cited as (24)</td>
<td>Quality of life; discriminative, evaluative</td>
<td>3 domains: mobility (3-level ordinal scale), social activity (5-level ordinal scale), physical activity (4-level ordinal scale)</td>
<td>Adult tool applied to children 25 boys and 19 girls with CF Ages 7-16, mean 16.5 ± 6.9 yrs</td>
<td>Interview administered to parent or patient depending on age</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Scale name</th>
<th>Purpose of scale</th>
<th>Domains, format, &amp; scoring</th>
<th>Population</th>
<th>Method of administration</th>
<th>Standardization of measure</th>
<th>Taxonomic class (as per Table 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canadian Occupational Performance Measure (COMP)</td>
<td>Evaluative</td>
<td>Domains: self-care, productivity and leisure</td>
<td>Not age specific</td>
<td>Clinician administered</td>
<td>No evidence of validity or reliability included (14)</td>
<td>Interview measure of capability and performance [2.11 &amp; 2.23]</td>
</tr>
<tr>
<td>(14,34,35)</td>
<td>Subjects generate their own items</td>
<td>Dimensions: importance of activities, level of performance &amp; satisfaction with performance</td>
<td>Developed for adults and applied to children</td>
<td>Items spontaneously elicited from each patient</td>
<td>Interview of parent or child</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Useful for comparison within individual patients rather than between patients</td>
<td>Scoring: 10-category ordinal scales</td>
<td></td>
<td></td>
<td>Environmental demands</td>
<td></td>
</tr>
<tr>
<td>Pediatric Evaluation of Disability Inventory (PEDI)</td>
<td>Physical Function &amp; Independence Measure; evaluative</td>
<td>Domains: self-care, mobility, social function, Scoring dimensions: functional capability, caregiver assistance, environmental modifications</td>
<td>Chronically ill and disabled children 0.5-7 yrs old</td>
<td>Parent report</td>
<td>Concurrent validity: moderately high correlations with Battelle Developmental Inventory Screening Test for self-care and mobility domains but not social function domains (10)</td>
<td>Parent report of performance [2.12]</td>
</tr>
<tr>
<td>Development edition (10,17,18)</td>
<td>Expert reviewers preferred to class the tool as discriminative rather than evaluative (17)</td>
<td>Scored: 100 in 10 points increments</td>
<td>Parent report 20 min to 1 h to complete</td>
<td></td>
<td>Significant differences between normals and disabled samples (10)</td>
<td></td>
</tr>
<tr>
<td>Play Performance Scale (31,32)</td>
<td>Play, evaluative</td>
<td>Concepts based on Karnofsky’s Scale</td>
<td>1-16 yrs olds</td>
<td>Parent report</td>
<td>Construct validity: detected significant difference between patients and siblings, and significantly related to global measure of performance from nurses and researchers (r = 0.73 and r = 0.92, respectively) (32)</td>
<td>Parent report of performance [2.12]</td>
</tr>
<tr>
<td>Rand Health Insurance Study Scale (HIS) (6,7)</td>
<td>Physical activity; discriminative</td>
<td>4 domains: mobility, physical activity, role activity, and self-care</td>
<td>Ages 0-13 yrs (n = 2,152 children in 6 U.S. cities) (7)</td>
<td>Researcher administered</td>
<td>Construct validity: comparison of HIS classification of able/disabled with 11 other scales showed significant differences for all 11 scales; however, actual differences were small and the sample large</td>
<td>Parent report of performance [2.12]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Children’s tool similar to AIMS (7)</td>
<td>Ages 4-16 yrs (n = 3,294 children in Ontario) (4)</td>
<td>Parent report</td>
<td>Found a 57/1,000 prevalence of disability (4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Healthy populations (n = 116 pediatric trauma survivors age 8.7 ± 4.4) (64)</td>
<td>Healthy populations (n = 116 pediatric trauma survivors age 8.7 ± 4.4) (64)</td>
<td></td>
<td>Wesson et al. found the HIS not to be able to discriminate severity in a trauma population (54)</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### APPENDIX (Continued)

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Purpose of scale</th>
<th>Domains, format, &amp; scoring</th>
<th>Population</th>
<th>Method of administration</th>
<th>Standardization of measure</th>
<th>Taxonomic class (as per Table I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childhood Health Assessment Questionnaire (CHAQ) (53,54)</td>
<td>Functional status, evaluative</td>
<td>8 domains: dressing, grooming, arising, eating, walking, hygiene, reach, grip, and activities</td>
<td>IRA (n = 62) 1-19 yr olds Derived from an adult tool</td>
<td>Parent or patient self-administered 10 min</td>
<td>No documentation of validity or reliability in children</td>
<td>Parent or self-report of capability [2.22]</td>
</tr>
<tr>
<td>Gross Motor Function Measure (GMFM) (49,50)</td>
<td>Gross motor abilities; discriminative</td>
<td>85-98 items Gross motor skills Assess the quantity of skill capability not quality 4-point ordinal scale per item</td>
<td>Cerebral palsy (n = 111) and acute head injuries (n = 25) (control subjects were 34 normals &lt;5 yrs old)</td>
<td>Trained clinician observation</td>
<td>Highly structured assessment</td>
<td>Time estimated at &gt;1 h</td>
</tr>
<tr>
<td>Gross Motor Performance Measure (GMPM) (2.3)</td>
<td>Qualitative aspects of gross motor function; evaluative</td>
<td>22 items Assesses the quantity of skill capability not quality 4-point ordinal scale per item</td>
<td>Cerebral palsy (n = 111) and acute head injuries (n = 25) (control subjects were 34 normals &lt;5 yrs old)</td>
<td>Trained clinician observation</td>
<td>Content validity assessed as good by expert rating</td>
<td>Clinician observation of capability [1.2]</td>
</tr>
<tr>
<td>Seated Postural Control Measure (SPCM) (11)</td>
<td>Seating function</td>
<td>22 items 1-19 yrs of age Neurologically impaired (n = 45) Functional movements = 17 items</td>
<td>Clinician observation 20 min 3- or 4-level ordinal scale for each item</td>
<td>Time estimated at &gt;1 h</td>
<td>Report face validity but no details of methodology Reliability: interrater = 0.35 for alignment and 0.85 for function Test-retest reported as r = 0.35 for alignment and 0.22 for function (poor)</td>
<td>Clinician observation of capability [1.2]</td>
</tr>
<tr>
<td>Children's Adaptive Behavior Checklist (CAB) (14) original ref. cited as (26)</td>
<td>Adaptive behavior; discriminative, evaluative</td>
<td>5 subdomains: language, independent function, family role performance, economic vocational activity, and socialization Ordinal scale 5-11 yr olds Developmentally disabled</td>
<td>Trained psychologist-educational evaluator Interview of child 45 min</td>
<td>Adequate content, construct, and criterion validity</td>
<td>Interview-administered measure (not clear whether performance or capability) [2.11 or 2.21]</td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Purpose of scale</th>
<th>Domains, format, &amp; scoring</th>
<th>Population</th>
<th>Method of administration</th>
<th>Standardization of measure</th>
<th>Taxonomic class as per Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile Arthritis Functional Assessment Scale (JAFAS) (36)</td>
<td>Speed of physical function; discriminative</td>
<td>10 items from previous scales</td>
<td>JRA (n = 71)</td>
<td>Clinician observation</td>
<td>Convergent validity: with number of involved joints r = 0.4, p = 0.003, with Steinbrocker class r = 0.59, p = 0.0001, with disease activity r = 0.32, p = 0.01</td>
<td>Clinician observation of capability [0.12]</td>
</tr>
<tr>
<td>Juvenile Arthritis Functional Assessment Report (JAFAR)</td>
<td>Independent performance of activities; discriminative</td>
<td>Same as above with the addition of parents (JAFAR-P not administered for controls)</td>
<td>JAFAR-C interviewer-administered to children</td>
<td>Construct validity: differences between patients and controls (p &lt; 0.001), No differences between parents' reports (p = 0.54), No significant age correlation</td>
<td>JAFAS-C self-report of performance [0.12]</td>
<td></td>
</tr>
<tr>
<td>Juvenile Arthritis Function Index (JASI) (66.67)</td>
<td>JRA physical function; evaluative, discriminative</td>
<td>Intended for juvenile rheumatoid arthritis patients Ages 6-18 (n = 30)</td>
<td>Self-report 272 items (current version has been reduced to 100 items)</td>
<td>Face validity assessed by 17 clinicians, Test-retest reliability ICC = 0.99 (67)</td>
<td>JAFAR-P parent report of performance [0.12]</td>
<td></td>
</tr>
<tr>
<td>Arthritis Impact Measurement Scales (AIMS) (15)</td>
<td>Physical limitations; discriminative</td>
<td>2 dimensions: physical disability (modified), pain 9 items (from the original 45 items) 2-5 response options per item</td>
<td>Adult tool, modified for children with juvenile arthritis (n = 77) Ages 2-17 yrs, mean = 9.3 yrs (70% girls) (97% white)</td>
<td>Clinician administered interview of parents in a clinic setting</td>
<td>Coulton et al. state attainment of convergent validity; however, this conclusion is not well supported by data, Correlations with diagnostic category = 0.24-0.26, and joint count = 0.31-0.35 Discriminated between active and inactive disease groups at p &lt; 0.01</td>
<td>Interview measure of capability [0.21]</td>
</tr>
</tbody>
</table>

* See also WesFIM, MCA, PEDI, TAMP, and COPM.
* See also CHAQ and COPM.