Clinical Applications of Outcome Tools in Ambulatory Children with Cerebral Palsy

Donna J. Oeffinger, PhD\a,*, Sarah P. Rogers, MPH\b, Anita Bagley, PhD\c, George Gorton, BS\d, Chester M. Tylkowski, MD\e

An increasing demand for evidence-based decision-making has challenged the medical community to clearly show that treatment improves a person’s functional abilities within their environment. This has resulted in increased use of outcome tools in the clinical setting to supplement technical measures, such as physical examination. Outcome tools are used to measure functional performance, as a baseline descriptive assessment, to select treatment goals, and to evaluate treatment.\(^1\) Outcome tools can help maintain patient and family motivation and provide scientific evidence on outcomes that are meaningful to patients and clinicians.

Outcome tools objectively quantify some aspect (physical, mental, or emotional) of a person. They are typically developed with a specific purpose for a target population and are used in clinical care to assess or establish treatment efficacy. They should be valid, reliable, and responsive to change. Validity is the degree to which a scale

This work was supported by Grant 9140 from the Shriners Hospitals for Children Clinical Outcomes Study Advisory Board.
\(\text{a} \) Pediatric Orthopedics, Shriners Hospitals for Children, 1900 Richmond Road, Lexington, KY 40502, USA
\(\text{b} \) Clinical Outcomes Program, Shriners Hospitals for Children, 1900 Richmond Road, Lexington, KY 40502, USA
\(\text{c} \) Motion Analysis Laboratory, Shriners Hospitals for Children, 2425 Stockton Boulevard, Sacramento, CA 95817, USA
\(\text{d} \) Clinical Outcome Assessment Laboratory, Shriners Hospitals for Children, 516 Carew Street, Springfield, MA 01104, USA
\(\text{e} \) Pediatric Orthopedics, Shriners Hospitals for Children, 1900 Richmond Road, Lexington, KY 40502, USA
* Corresponding author.
E-mail address: doeffinger@shrinenet.org (D.J. Oeffinger).

doi:10.1016/j.pmr.2009.04.003
1047-9651/09/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
measures what it is intended to measure. Reliability is the extent to which a measure yields consistent values over time when no change has occurred and yields consistent values between examiners. Responsiveness establishes a tool’s ability to detect change when a change has occurred.

Since clinicians may lack an understanding of what to measure, lack information regarding appropriate tools, have difficulty assessing psychometric properties of tools, have doubts about clinical utility of tools, or lack the time or money for training and implementation, outcome tools are not always used by clinicians consistently, as intended, or at all. Clinicians tend to use the easiest, most familiar, or available measure rather than what is “best” theoretically.

With a goal of providing clinicians and clinician-researchers with more information on outcome tools used to assess ambulatory children with cerebral palsy (CP), this article provides an overview of several outcome tools commonly used in this population, provides research findings from a recent large multicenter study, and provides ways to integrate the research findings into clinical practice and clinical outcomes research. The information presented is the result of work conducted by the Functional Assessment Research Group (FARG), a group comprising individuals from 7 pediatric orthopedic facilities (Shriners Hospitals for Children [SHC]-Lexington, KY; SHC-Northern California; SHC-Springfield; SHC-Houston; SHC-Salt Lake City; Washington University, St. Louis; and the University of Virginia). The FARG mission is to conduct research contributing to the evidence base for treatment of the neuromusculoskeletal system in ambulatory children with CP.

In 2001, the World Health Organization (WHO) released the International Classification of Functioning, Disability and Health (ICF) framework. The ICF provides a comprehensive framework to approach outcome measurement of children with disabilities. The ICF dimensions include Body Functions and Structures, Activities and Participation, Environmental factors, and Personal factors. Health-related quality of life is not formally included in the ICF framework, but was assessed by FARG, and is defined as “what people ‘feel’ about their health condition or its consequences.”

Several outcome tools assessing the ICF dimensions of Body Functions and Structures and of Activities and Participation that are frequently used in pediatric orthopedics were selected for this study. These included: the Gillette Functional Assessment Questionnaire (FAQ), Gross Motor Function Measure (GMFM), Pediatric Quality of Life Inventory (PedsQL), Pediatric Outcomes Data Collection Instrument (PODCI), Pediatric Functional Independence Measure (WeeFIM), temporal-spatial gait parameters (velocity, stride length, and cadence), and energy cost during walking (O₂ cost). With the exception of FAQ Questions Two and Three and PODCI subscales of Satisfaction and Expectations, each tool has been tested for content validity and reliability. These tools encompass technical measures (temporal-spatial gait parameters, O₂ cost), clinician-rated tools (GMFM D (standing) & E (walking, running, jumping), GMFM-66), parent-reported measures (WeeFIM, FAQ, PODCI, PedsQL), and child-reported measures (PODCI, PedsQL).

The Gross Motor Function Classification System (GMFCS) is a standardized classification system that categorizes the function of individuals with CP into 5 levels, with level I being the most functional and level V the least. The GMFCS is valid and reliable with excellent inter-rater reliability. The FARG study focused on ambulatory children of GMFCS levels I to III: the child at level I has some motor limitations but does not use devices, at level II has limitations walking indoors and outdoors but does not use devices, and at level III has limitations walking indoors and outdoors and uses assistive devices.
GOALS OF THIS ARTICLE

The FARG study methodology, statistical analyses, and results are fully described in the series of articles published in *Developmental Medicine and Child Neurology (DMCN)* \(^{14,29–32}\) and *Journal of Pediatric Orthopedics*. \(^{33}\) This article will summarize the FARG study findings and succinctly answer common questions related to the clinical applications of outcome tools in children with CP. Questions related to both the cross-sectional and longitudinal study data are answered, followed by questions related to how to apply the data clinically. In addition to summarizing the study findings, examples of ways to integrate the findings into clinical practice are presented.

**Cross-sectional and Longitudinal Questions**

For ambulatory children with CP:

- Do outcome scores reflect differences based on GMFCS levels?
- What outcome tools best discriminate physical function?
- Do parent and child reports differ on outcome tools?
- Do children with hemiplegia function differently from those with diplegia?
- How do scores on outcome tools change over 1 year?
- Are outcome tools responsive to change?
- Does orthopedic surgery change function as measured by changes in scores on the studied outcome tools?
- What is a minimum clinically important change in outcome scores?

**Clinical Practice Questions**

- How do I integrate this information into my practice?
- Can I tell how the child is doing compared with others with similar severity level?
- Can I predict the normal range of scores for children with CP based on age and GMFCS level?
- How do I know if a changed score is clinically meaningful?
- Can I use outcome tools to focus on the treatment plan?

In answering these questions, comparison data are presented that will assist clinicians to understand how a child is functioning, develop individualized treatment plans, and know when a clinically meaningful change has occurred. Use of these standardized assessments can also improve communication among parents, patients, and care providers.

**METHODS**

This 6-year prospective multicenter study was entitled “A Cross-sectional and Longitudinal Assessment of Outcome Instruments in Patients with Ambulatory Cerebral Palsy.” The study was conducted at 7 pediatric orthopedic facilities across the United States that treat children from several surrounding states. Institutional Review Board approval was obtained at each site and all participants signed consent, assent as appropriate, and privacy and confidentiality forms.

**Participants**

Participants were enrolled in the study which included cross-sectional and longitudinal endpoints. Inclusion criteria were: diagnosis of CP, GMFCS levels I to III, ages 4 to 18 years, and ability to complete a gait evaluation. Exclusion criteria were: previous selective dorsal rhizotomy, lower extremity orthopedic surgery within the past year,
botulinum toxin-A injections within the past 6 months, or a currently operating baclofen pump.

Five hundred and sixty-two individuals completed the baseline assessments and 387 completed the follow-up evaluation (68.7%). Of those who did not complete the follow-up evaluation: 95 were unable to be contacted (16.9%), 29 declined (5.2%), 9 did not come for their follow-up appointment (1.6%), 4 were no longer ambulatory (0.7%), 2 had surgery less than 1 year from baseline to study completion (0.4%), and 36 did not participate for other reasons (6.4%). Six (1.6%) were excluded from analysis because of incomplete data, resulting in a final longitudinal sample of 381. There were no differences at baseline for age, height, weight, type of involvement, gender, GMFCS level, birth history, and ethnicity between those who completed the follow-up and those who did not.

Of the 562 participants who completed the cross-sectional component, there were 339 (60%) males, 223 (40%) females; 240 (43%) GMFCS level I, 196 (35%) Level II, and 126 (22%) Level III; 400 (71%) diplegic and 162 (29%) hemiplegic; and 83% were Caucasian. Mean age at baseline was 11.0 (standard deviation [SD] 4.3) years (range, 4.1–18.3).

For the 381 participants who completed the follow-up assessments, there were 230 (60%) males, 151 (40%) females; 174 (46%) GMFCS level I, 132 (34%) level II, and 75 (20%) level III; 265 (69%) diplegic and 116 (31%) hemiplegic; and predominately Caucasian (83%). Mean age at follow-up was 12.4 (SD 3.4) years (range, 5.2–20.5). Mean time between assessments was 1.4 (SD 0.4) years. There was no significant difference in mean age, height, weight, sex, or ethnicity distributions among GMFCS levels at either baseline or follow-up.

Between assessments, individuals received treatments based on their physician’s recommendations. During the study period, 87 participants (23%) had orthopedic surgery. The remaining participants had treatments that included physical therapy, bracing, and observation.

**Tools and Study Methodology**

GMFM Dimensions D and E, Parent and Child PedsQL, Parent and Child PODCI, FAQ, WeeFIM, \( \text{O}_2 \) cost, temporal-spatial gait parameters, and GMFCS level were collected at baseline and follow-up done at least 1 year later. Before study initiation, consistency among coordinators was verified; local coordinators were trained in GMFCS classification, tool administration, and data collection procedures. Data were collected in a study-specific database via direct computer entry and reviewed by the project manager for completeness and accuracy. Statistical analyses performed to address each study hypothesis are described briefly in the following results sections and they were also fully described in the previously published articles.\(^14,29-33\)

**RESULTS AND CLINICAL APPLICATIONS**

**Do Outcome Scores Reflect Differences Based on GMFCS Level?**

To gain a better understanding of the outcome tools and provide clinical comparison data on the function of individual children to their peers with CP, a descriptive analysis was completed.\(^14\) The results showed that children in GMFCS levels I, II, and III functioned differently as measured by the study outcome tools. The mean scores for each outcome tool were clearly separated by GMFCS level. The completed results for all studied tools which included the mean, SD, range, interquartile range, and 95% confidence intervals were previously reported.\(^14\)
Analysis of variance (ANOVA) was used to assess the relationships between outcome scores and GMFCS levels. It was determined that there were direct relationships between outcome scores and GMFCS level for Parent and Child PODCI Global Function, Transfers and Basic Mobility, Sports and Physical Function, Parent PODCI Upper Extremity and Physical Function, WeeFIM domains of Self Care and Mobility, FAQ Question One (walking scale), GMFM Dimensions D and E, GMFM-66 score, O₂ cost, and temporal-spatial gait parameters. For all of the functional assessments, as the child’s severity level increased (from GMFCS I to II to III), the mean functional outcome scores decreased; an example is shown in Fig. 1. However, for the quality of life measures of Parent and Child PODCI Comfort/Pain, Parent and Child PODCI Happiness, Child PODCI Expectations, and Parent and Child PedsQL Emotional Functioning, children of different GMFCS levels reported a similar quality of life.

Although there was clear separation of the means, other descriptive measures of central tendency overlapped. The large sample size in this study may have contributed to finding statistically significant differences among GMFCS levels that may not have clinical significance. The ANOVA examines if the means of the groups are different and provides little information about the variation within the groups. Therefore, the degree of separation and overlap of interquartile ranges (middle half of population, 25%-75%) among GMFCS levels were reviewed. Overlap between levels shows the heterogeneous nature of CP. There was minimal overlap between the lower end of GMFCS level I scores and the upper end of level III scores. However, there was substantial overlap between the lower end of level I scores and the upper end of level II scores. These findings are consistent with the difficulties reported in classifying patients between levels I and II.21,27 Minimal overlap was seen between the lower end of level II and the upper end of level III.

The FAQ data are nonparametric and required analyses using odds ratios. The analysis revealed that FAQ Question One showed a difference among all GMFCS levels indicating that parents’ report of their child’s walking ability is consistent with the clinician’s gross motor rating using the GMFCS. Parents reported on FAQ Question Two what they felt limited their child’s walking ability and on FAQ Question Three what skills...
their child is able to do under typical conditions. Only 27% of parents reported that
pain limits their child’s walking ability and 39% reported safety as a limiting factor. As severity level increased, a greater percentage reported safety as a problem. Regardless of GMFCS level, the majority of parents felt that balance, endurance, and weakness were limiting factors for walking. For FAQ Question Three, parents indicated that nearly all children were able to walk up and down stairs using a railing (87%), kick a ball (92%), and step over an object (90%). Only 11% of the children were able to jump rope and 15% to ice or roller skate. The parent’s perspective is important to help focus goals and clinicians should be sensitive to parents’ concerns when developing treatment plans.

Since the GMFCS was designed as a broad classification system based on function,21 one would expect children in different GMFCS levels to score differently on outcome tools. The significant differences in mean outcome scores found on the tools used in this study indicate that children classified into different GMFCS levels function differently. These findings support the use of the study outcome tools to assess function and support the GMFCS as an appropriate method of classifying severity, despite some overlap between levels. Examples of the differences among GMFCS levels for the studied outcome tools are reported in Table 1.

What Outcome Tools Best Discriminate Physical Function in Ambulatory Children with CP?

Knowing which tools best discriminate among GMFCS levels can assist in the selection of tools for clinical or research use, since it is not practical and often not feasible to administer multiple tools during a single visit.29 The best outcome tool, however, is dependent upon the clinical/research question and clinical/research endpoints of interest. The recommendations presented here focus on the tools assessed in this study that were the most discriminatory among GMFCS levels.

Using effect size indices (ESIs) for parametric variables and odds ratios for nonparametric data, the magnitude of differences in scores on the study tools across GMFCS levels were quantified. Binary logistic regression models determined discrimination, and receiver operating characteristic (ROC) curves addressed sensitivity and specificity of the measures.

Based on the study tools’ ability to discriminate (large ESIs) among GMFCS levels, the best tools to measure physical function in ambulatory children with CP are the GMFM Dimension E or GMFM-66, the Parent report PODCI Sports and Physical Function, Global Function, and the FAQ Question One. Tools that performed well and that are good secondary measures are the WeeFIM Self Care and Mobility domains and temporal-spatial gait parameters. The least discriminatory tools were the quality of life and cognition measures; however, these are important in comprehensive assessments of treatment effects.

<table>
<thead>
<tr>
<th>Outcome Tool</th>
<th>GMFCS I Mean (SD)</th>
<th>GMFCS II Mean (SD)</th>
<th>GMFCS III Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMFM D- standing</td>
<td>94 (6)</td>
<td>84 (8)</td>
<td>52 (24)</td>
</tr>
<tr>
<td>GMFM E- walking</td>
<td>93 (8)</td>
<td>75 (16)</td>
<td>32 (20)</td>
</tr>
<tr>
<td>Energy (O₂) cost (mL O₂/kg/m)</td>
<td>0.28 (0.1)</td>
<td>0.38 (0.2)</td>
<td>0.57 (0.3)</td>
</tr>
<tr>
<td>Parent PODCI global function</td>
<td>81 (11)</td>
<td>72 (12)</td>
<td>62 (13)</td>
</tr>
<tr>
<td>Parent PODCI sports</td>
<td>68 (17)</td>
<td>51 (17)</td>
<td>35 (17)</td>
</tr>
</tbody>
</table>
ROC curves were constructed to determine sensitivity and specificity for the most discriminant measures identified by regression models. Based on this analysis, the single most discriminatory measure of physical function, as classified by the GMFCS, is the GMFM-66 score. This is the Rasch-analyzed score of the GMFM using the Gross Motor Activity Estimator program. Rasch analysis equalizes the difference between points on the scale, creating an interval scale, such that the interval of change from 15 to 25 is equivalent to the change from 65 to 75. The GMFM-66 is the best clinical choice for discriminating among GMFCS levels I through III.

A tool's discriminatory ability may be limited by ceiling effects. Ceiling effects suggest that either the participants do not have significant limitations in these areas, or that there is a need for more challenging “upper end” items or questions. Parent and Child reports of PODCI Comfort/Pain and WeeFIM Cognition have large ceiling effects for all three GMFCS levels, suggesting that pain and limitations in cognition are not major issues for these children or that the questions are too easy to identify if such issues are present.

As an example of a tool with limited discriminatory ability, Fig. 2 shows a box plot of the Parent report of PODCI Comfort/Pain scores. Note that there are no upper fences for any GMFCS level, showing ceiling effects. The boxes represent the interquartile ranges for the data, presenting the range of scores obtained by the middle 50% of the participants for each GMFCS level. The size of the boxes shows variability within GMFCS level and degree of overlap between GMFCS levels. For the example shown, ceiling effects and the overlap of interquartile ranges limit the discriminatory ability of this subscale.

**Do Parent and Child Reports Differ on Outcome Tools?**

The PODCI and PedsQL have parent and child report forms. Given the time constraints often present in the clinical setting, it would helpful to know if completion by both parent and child are necessary. Therefore, differences in parent and child responses on the same tools were studied. A direct relationship (as severity increased scores decreased) between outcome scores and GMFCS levels was
seen for both the parent and child reports, despite the differences noted between scores. For 2 of the tools studied, the PODCI and PedsQL, there are both a parent and child report version. Results showed that there were differences between the scores reported by the parent and the child. Children scored themselves higher than their parents for almost all PODCI and PedsQL subscales. As the child’s GMFCS level increased, the differences between parent and child scores increased. Physical subscales of PODCI Sports and Physical Function and PedsQL Physical Functioning showed the greatest differences. Example data are presented in Table 2. This is not the case for able-bodied children, who tend to report the same scores as their parents. The authors speculate that these findings are likely related to perspectives of disability. The child’s perception is one of ability, since the impairment was not acquired after a period of normal development. The child emphasizes what he can do. Parents have the expectation that their child should be able to do everything able-bodied children can. Therefore, the parent’s perspective is more likely one of disability and emphasizes what the child cannot do. For these reasons, it is important to obtain both parent and child perspectives on goals and outcomes.

Do Children with Hemiplegia Differ from Children with Diplegia?

Children with hemiplegia compared with those with diplegia were assessed, since grouping of children with CP with notably different clinical presentations, and presumably different profiles of brain injury, may lead to inaccurate predictions of developmental rate and prognosis. Based on the scores of the studied outcome tools, children with hemiplegia function differently from those with diplegia. Differences in outcome scores between these groups within GMFCS levels were tested and children with hemiplegia consistently scored lower on tools related to upper extremity and school function, and higher on nearly all tools that assessed gait or lower extremity function.

Ceiling effects were shown to be a function of the diagnostic subtype, GMFCS level, and outcome tool. They were mostly higher in GMFCS level I. The ceiling effects reflected the different strengths in each of the groups and were not consistently higher in either diagnostic subtype. For example, the children with hemiplegia had a higher percentage of ceiling effects in the PODCI Transfers and Basic Mobility subscale and WeeFIM Mobility, which shows their reliably higher scores on lower extremity functional assessments. However, the children with diplegia had a higher percentage of ceiling effects on WeeFIM Self Care, which shows their somewhat better upper extremity function. It is important to note that these ceiling effects may have made differences between diagnostic subtypes less evident.

Objective evidence of the distinct differences between children with diplegia and hemiplegia within the same GMFCS classification level exists in various aspects of motor functioning, activity, participation, and quality of life. Using GMFCS level alone

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Outcome Tool</th>
<th>Parent Mean (SD)</th>
<th>Child Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PODCI sports</td>
<td>57 (21)</td>
<td>70 (19)</td>
</tr>
<tr>
<td>PODCI upper extremity</td>
<td>84 (14)</td>
<td>93 (10)</td>
</tr>
<tr>
<td>PODCI transfers</td>
<td>86 (13)</td>
<td>93 (9)</td>
</tr>
<tr>
<td>PedsQL physical functioning</td>
<td>57 (19)</td>
<td>69 (19)</td>
</tr>
<tr>
<td>PedsQL social functioning</td>
<td>57 (19)</td>
<td>65 (22)</td>
</tr>
</tbody>
</table>
as a predictor of motor prognosis may overestimate the lower extremity skills of children with diplegia, and underestimate those of children with hemiplegia.

How Do Scores Change on Outcome Tools for Ambulatory Children with CP Over 1 Year?

Since children with chronic conditions such as CP are assessed repeatedly over time, it is important to know the expected amount of change as a result of natural history. Expected changes can be used by clinicians to assess whether observed changes with time are atypical, thus influencing treatment planning. Change data on the 381 individuals who completed both the baseline and follow-up assessments were analyzed. For the studied tools, small mean changes with large SDs were seen over a 1-year period. The small mean changes were a result of the outcome tool change scores being normally distributed, with some large positive and negative changes. The large variability in change scores is likely because of the heterogeneity of CP. Complete results were published by Oeffinger and colleagues in 2008; sample results are presented in Table 3. In addition, the average change scores and SDs are presented by GMFCS level and surgical or nonsurgical group by Oeffinger and colleagues (2008). Since changes that occur over a year are often small, one should be particularly cognizant of those individuals that present with large changes in function in a short period of time.

Are Outcome Tools Responsive to Change?

Analyses were completed to determine if the tools were responsive to change despite small mean changes over a 1-year period. To assess the responsiveness of the tools, a known change in function was needed. The investigators used a verified change in GMFCS level between assessments as a known change in function. Based on the analysis, the outcome tools were responsive when a change in function occurred large enough to result in a change in GMFCS level. The tools that were the most responsive to change were the functional subscales (GMFM Dimensions D and E; GMFM-66; Parent PODCI, Global Function, Transfers and Sports; gait velocity; and stride length). Large changes in function were reflected appropriately on the outcome tools. As function improved, the outcome scores increased and as function decreased, the outcome scores decreased. Additional research must be conducted to determine the most responsive tools, which can vary based upon GMFCS level and endpoints of interest.

<table>
<thead>
<tr>
<th>Outcome Tool</th>
<th>GMFCS I Mean (SD)</th>
<th>GMFCS II Mean (SD)</th>
<th>GMFCS III Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMFM-66</td>
<td>0.2 (5.4)</td>
<td>0.8 (3.5)</td>
<td>0.5 (3.5)</td>
</tr>
<tr>
<td>Velocity (% normal)</td>
<td>−0.8 (16.5)</td>
<td>0.5 (16.2)</td>
<td>−0.9 (15.2)</td>
</tr>
<tr>
<td>Parent PODCI sports</td>
<td>1.0 (12.4)</td>
<td>2.5 (13.9)</td>
<td>2.1 (13.2)</td>
</tr>
<tr>
<td>Parent PODCI upper extremity</td>
<td>3.9 (10.0)</td>
<td>3.9 (11.6)</td>
<td>1.6 (9.8)</td>
</tr>
<tr>
<td>Parent PedsQL transfers</td>
<td>0.6 (10.8)</td>
<td>2.5 (10.3)</td>
<td>2.3 (12.1)</td>
</tr>
<tr>
<td>Parent PedsQL physical</td>
<td>−0.4 (16.5)</td>
<td>1.8 (16.9)</td>
<td>−0.7 (18.9)</td>
</tr>
<tr>
<td>Parent PedsQL social functioning</td>
<td>2.4 (15.8)</td>
<td>4.1 (17.3)</td>
<td>3.8 (15.2)</td>
</tr>
<tr>
<td>WeeFIM mobility</td>
<td>−0.1 (4.2)</td>
<td>0.7 (8.2)</td>
<td>0.8 (8.8)</td>
</tr>
<tr>
<td>WeeFIM self care</td>
<td>0.8 (7.3)</td>
<td>2.0 (11.0)</td>
<td>3.0 (10.4)</td>
</tr>
</tbody>
</table>

Table 3

Example of mean change scores and SDs for select outcome tools by GMFCS level
**Does Orthopedic Surgery Change Function as Measured by Changes in Scores on the Studied Outcome Measures?**

Lower extremity musculotendinous surgery is a standard treatment for ambulatory children with CP who have joint contractures and bony torsions.36 Ultimately, the objectives of surgical management in CP are to improve function, decrease discomfort, and prevent disabling structural changes.37,38 The assumption is that by improving gait, function in general will improve. Therefore, data from a subset of 75 individuals who underwent lower extremity surgery were evaluated and their changes compared with changes in a matched nonsurgical group.36

The 75 participants who had surgery during the study period were individually matched with 1 of the 294 participants who did not have surgery. For each surgical participant, all nonsurgical participants who exactly matched by gender, GMFCS level and type of involvement (hemiplegic or diplegic) were identified. The surgical procedures included both soft tissue and bone surgery. The nonsurgical group had a standard of care treatment within the study window. This included observation, stretching and strengthening exercises, bracing, and medication management, if necessary.

The mean change scores showed that lower extremity orthopedic surgery is effective in improving gait, but these improvements do not necessarily translate into large changes in outcome tool scores. The greatest changes as a result of lower extremity surgical intervention occurred at the ICF Body Structure and Function level, closest to the level of surgical intervention. The FARG study found that surgery significantly improves gait kinematics as measured by the Gillette Gait Index (GGI), which quantifies the magnitude of deviation from normal gait. The greater the severity level, the greater the magnitude of change in GGI from pre- to postoperative. However, the improvements in gait did not translate into significant changes in measures of Activity and Participation. These data suggest that the outcome tools are not sufficiently responsive to measure the effect gait improvements have on Activity and Participation, or that there are no such changes. The complete analysis and findings are in press (Gorton and colleagues36).

**What is a Minimum Clinically Important Difference in Changes in Outcome Scores?**

Although statistically significant differences are important, it is more relevant for evidence-based practice to define clinically significant changes.31 Reported changes frequently reach statistical significance, but may not be clinically meaningful.39,40 Clinically meaningful infers different connotations to the child, family, and clinician. Minimum clinically important differences (MCIDs) are often subjectively defined, based on training and experience. MCIDs can also be quantified objectively and have been defined and calculated in various ways.16,35,41–43 In this study, MCID is defined as the magnitude of change required for an observable difference in function, and it is quantified using effect sizes. Effect size is a unitless measurement of the number of SDs from the mean (mean/SD). Small effect sizes may be described as imperceptible to the human eye, medium as large enough to be seen in normal observation, and large as grossly observable.6

MCIDs were calculated using the data from the longitudinal component of the study. The details of the analysis were reported by Oeffinger and colleagues31 (2008). MCIDs were established for the GMFM, PODCI, PedsQL, WeeFIM, O2 cost, and temporal-spatial gait parameters for ambulatory children with CP. Change scores exceeding MCIDs for medium effect sizes are considered large enough to be clinically meaningful. Sample results from data previously reported in DMCK61 are shown in Table 4.
TRANSLATION INTO CLINICAL PRACTICE

When data from studies are reported in the literature, the next critical step is the translation of this knowledge into information to help clinicians understand how to use the data in their daily clinical practice.

How Do I Integrate this Information into My Clinical Practice?
Identify outcomes of importance and relevance to your practice. Select tools to match your focus and to best answer your clinical or research questions. Each measure has unique characteristics and examines different parameters. For example, there are measures that quantify physical function, activity, participation, quality of life, or a combination. Similarly, there are tools designed to capture parent report, child report, or clinician ratings. The limitations of each measure must be understood, including ceiling effects, parent versus child report, and whether the population is hemiplegic versus diplegic. Once tools have been selected, you can develop a protocol for administering them routinely in your practice and using the resulting data for evaluation, treatment planning, and outcomes assessment.

Can You Tell How the Child is Doing Compared to Others of Similar Severity Level?
Clinicians should compare the child’s score to the descriptive data presented for ambulatory children with CP by GMFCS level, age, and diagnostic subtype when possible. The descriptive data include means, SDs, 95% confidence intervals, and interquartile ranges for each tool assessed.

By comparing the child’s score to the mean and interquartile range, you can determine the child’s ability relative to average for their peers. For example, if the child’s score falls below the interquartile range, the child is doing worse than 75% of his peers. You can then assess reasons why the child may be performing poorly and focus

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example of minimum clinically important difference (MCID) scores for medium (0.5) and large (0.8) effect sizes on select outcome tools</td>
</tr>
<tr>
<td>Outcome Tool</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>GMFM-66</td>
</tr>
<tr>
<td>O₂ cost (mL O₂/kg/min)</td>
</tr>
<tr>
<td>Velocity (% normal)</td>
</tr>
<tr>
<td>Parent PODCI global function</td>
</tr>
<tr>
<td>Parent PODCI sports</td>
</tr>
<tr>
<td>Parent PedsQL school</td>
</tr>
</tbody>
</table>
the rehabilitation or therapy interventions on those areas. Similarly, if the score is above the interquartile range, the child is performing better than 75% of her peers and interventions may not be needed.

**Can I Predict What the Normal Range of Scores is for a Child with CP Based on Age and GMFCS Level?**

For the Parent report of PODCI, prediction equations have been developed for each of the subscales based on age and GMFCS level. The prediction equations provide expected scores for a child, depending on age and GMFCS level, along with a range in which the predicted score may reside. Formulas exist for the following PODCI subscales: Upper Extremity Function; Transfers and Basic Mobility; Sports and Physical Function; Comfort/Pain; Global Function (composite of the 4 functional assessments); and Happiness.

Using the formulas and substituting for an individual child’s age and GMFCS level, the evaluating clinician can predict the score the child should have on the particular PODCI subscale. By comparing the child’s actual score to the predicted score, clinicians can determine how the child is doing compared with his or her peers with CP.

Below is an example of how to calculate the expected score for the Parent report of PODCI Transfers and Basic Mobility for an 11 year + 3-months old patient with CP in GMFCS level I using the appropriate formula.

**Formula:**

\[
\text{Predicted transfers score} = 78.375 + (7.8460 - G2 \times 8.460 - G3 \times 26.037)
\]

Age, actual age in years with decimal fraction; G2, GMFCS level II (0 = no, 1 = yes); G3, GMFCS level III (0 = no, 1 = yes). GMFCS level I is the default for the equations.

**Formula with substitutions appropriate to the example above:**

\[
\text{Predicted transfers} = 78.375 + (11.25 \times 1.069 - 0 \times 8.460 - 0 \times 26.037)
\]

Predicted Transfers Score = 90.401 with a SEE of ± 11.9

SEE indicates standard error of the estimate and is a constant specific to each PODCI dimension.

The patient’s actual score for this visit was 85. While the patient scored a little less than the expected score of 90, she is still well within the expected range (78–100). Treatment goals may include maintaining function in this area. One can determine areas of focus for treatment by examining the individual questions that make up the summary subscale. For example, questions that are part of the Transfer and Mobility subscale include “During the last week was it easy or hard for your child to: Climb one flight of stairs? Get on and off a toilet or chair? Sit in a regular chair without holding on?” If the patient scores at or above the expected score, appropriate treatment planning would be to maintain or possibly increase function, depending on the patient, family, and clinician goals. Alternatively, the family and clinician may choose to focus the treatment plan on areas in which the child is less functional in comparison to peers, as determined by the outcome scores.

**How Do I Know if a Change Score is Clinically Meaningful?**

By using the same outcome tools over time, you can determine if the child is improving or declining by comparing his or her score changes to expected changes of peers in the same GMFCS level. As previously seen, MCIDs were defined for all of the studied outcome tools and reported by Oeffinger and colleagues. These MCIDs can be used as guidelines for the amount of change that is clinically important. However, because
of the methods used to calculate the MCIDs, they are most appropriately applied to changes in groups or between groups. This can be particularly useful for researchers when trying to determine if a statistical difference is clinically meaningful. On the individual level, it is important to ask the child’s and family’s opinions related to the change in function when changes in outcome scores are measured.

As established by Oeffinger and colleagues, MCID is change that exceeds the value required for a medium effect size. If changes in a group’s outcome tool scores exceed the reported MCID values, a change beyond natural progression is likely to have occurred. If the difference between groups is larger than the MCID threshold, the difference between groups is clinically significant. MCID thresholds can assist clinicians and researchers in advancing interpretations from being based strictly on statistical significance or subjective clinical judgment, to being based on consistent quantitative evidence.

For example, a group of individuals with CP (GMFCS levels I–III) participated in a research study that examined function using velocity (% normal) before and after botulinum toxin-A injections. To reach a MCID at a medium effect size, a change of 5.7% between assessments is needed, and a change of 9.1% is needed to reach a large effect size. The mean at baseline was 75% and at 3-months post-injection was 82%. There was an improvement of 7% which was statistically different and also clinically significant since the change exceeded the MCID for a medium effect size.

Can I Use Outcome Tools to Focus the Treatment Plan?

The findings of the FARG study can be used in developing the treatment plan and goals. Based on the information provided in this article and the data presented in the study publications, you can objectively assess a child’s current functional level and use the scores to focus areas of treatment, thus saving time and improving the cost-effectiveness of your practice. You can then objectively reassess the child’s function over time to determine if a significant change has taken place and adjust treatment prescriptions and goals accordingly.

GENERALIZABILITY AND LIMITATIONS

These findings are appropriate to generalize for children with CP in GMFCS levels I to III and between the ages of 4 and 18 years. Since the majority of the study population was Caucasian (83%) and were English speaking, the study findings may have limited applicability to minority or non-English speaking children with CP.

Several challenges exist with administration of outcome tools in the clinical environment. First, there are a limited number of validated tools available to clinicians that have been studied and have comparison data available for use. Also, implementing a process for routine use of outcome tools that is accepted by staff can be met with resistance. If you are successful in implementing a process, long-term commitment to the process in the clinical setting can be difficult because of limits in clinical resources such as time, personnel, costs associated with purchasing and licensing, and fees for the outcome tools. Inefficient methods for data management and poor accessibility to data lead to little or no use of the collected information. Difficulties can exist in finding accessible comparison data in user-friendly formats for normal and diagnosis-specific populations.

These challenges can be overcome through education of clinicians regarding the usefulness and importance of outcome assessments. Selection of optimal outcome tools that provide meaningful information for all those involved and that can be
completed in a timely manner are critical first steps to a successful outcomes improvement program. Establishing protocols for administration and review of the data collected from the outcome tools, along with a method to assure follow-through with the established protocols, is necessary for success. Implementing changes in practice is difficult, but the difficulty can be overcome more quickly if training and coaching are provided by advanced clinicians, real-time outcomes feedback is provided to the clinician and patient, and synthesized research evidence is easily available at the point of care.8,44 Electronic information systems, such as databases, can provide real-time feedback; synthesized comparative research evidence at the point of care; and reduce time burdens for patients, families, and clinicians for data collection and comparison purposes. These suggestions must be tailored to your specific facility, where a multifaceted approach to clinician behavior change must address various local barriers and must engage clinicians.45 A successfully implemented outcomes improvement program balances the thoroughness of the clinical assessment with the burdens on the patient and family to complete the outcome tools and on the clinician to obtain the results. The outcomes assessed should be meaningful to clinicians and patients and lead to improved evidence-based practice.

SUMMARY

The summary presented here and in the referenced articles provides information on: outcome tools’ discriminatory ability and responsiveness; readily available comparison data on 7 commonly used outcome measures that can be used at the point of care for ambulatory children with CP; prediction equations for the Parent report PODCI by age and GMFCS level; and minimum clinically important difference thresholds by GMFCS level. This information can assist the clinician in selecting the best tools to discriminate among levels of severity and avoid ceiling effects. By documenting the results of care, you can provide objective assessment of changes from pretreatment status and monitor progress toward achievement of treatment goals, which helps to maintain patient motivation. It helps the clinician to know when to adjust or discontinue a program and can also help justify reimbursement.

When using the data presented to compare a child’s score to a population score, outlier scores become evident and may trigger patient reassessment to determine potential causes and courses of action. The scores allow direct comparisons between a specific patient and a matched cohort, assisting clinicians in the creation of comprehensive and individualized management plans. Understanding functional levels of an individual child with CP relative to other children with CP, before and after operative and nonoperative interventions, potentially provides insights into the benefits of specific treatment modalities. Outcome tools should be used in the clinical assessment of children with CP to assess outcomes that are meaningful to patients and families, document results of care, detect an indication for treatment, aid in treatment choice, and evaluate treatments to determine if the goals were met. Use of appropriate outcome tools can lead to best practices and reduced costs in the clinical setting.

ACKNOWLEDGMENTS

This work was funded by Shriners Hospitals for Children (SHC), Clinical Outcomes Study Advisory Board Grant # 9140, “A Cross-sectional and Longitudinal Assessment of Outcome Instruments in Patients with Ambulatory Cerebral Palsy.” The authors acknowledge the participation of the Functional Assessment Research Group study investigators (Mark Abel, MD, and Mark Romness, MD, University of Virginia; Doug Barnes, MD, Elroy Sullivan, PhD, and Judi Linton, BS, SHC-Houston;
Diane Damiano, PhD, Washington University; Diane Nicholson, PhD, SHC-Salt Lake City) and coordinators from each site. The authors would also like to acknowledge the patients and their families for their participation.

REFERENCES