Outcome assessments in children with cerebral palsy, Part II: discriminatory ability of outcome tools

Anita M Bagley* PhD, Shriners Hospital for Children (SHC), Sacramento, CA;
George Gorton BS, SHC, Springfield, MA;
Donna Oeffinger PhD, SHC, Lexington, KY;
Douglas Barnes MD;
Janine Calmes MS PT, SHC, Houston, TX;
Diane Nicholson PhD PT, SHC, Salt Lake City, UT;
Diane Damiano PhD PT, Washington University, St Louis, MO;
Mark Abel MD, University of Virginia, Charlottesville, VA;
Richard Kryscio PhD, University of Kentucky, Lexington, KY;
Sarah Rogers MPH;
Chester Tylkowski MD, SHC, Lexington, KY, USA.

**Correspondence to first author at* Shriners Hospital for Children, Northern California Motion Analysis Laboratory, 2425 Stockton Boulevard, Sacramento, CA 95817, USA.
E-mail: abagley@shrinenet.org.

Discriminatory ability of several pediatric outcome tools was assessed relative to Gross Motor Function Classification System (GMFCS) level in patients with cerebral palsy. Five hundred and sixty-two patients (400 with diplegia, 162 with hemiplegia; 339 males, 223 females; age range 4–18y, mean 11y 1mo [SD 3y 7mo]), classified as GMFCS Levels I to III, participated in this prospective multicenter, cross-sectional study. All tools were completed by parents and participants when appropriate. Effect size indices (ESIs) for parametric variables and odds ratios for non-parametric data quantified the magnitude of differences across GMFCS levels. Binary logistic regression models determined discrimination, and receiver operating characteristic curves addressed sensitivity and specificity. Between Levels I and II, the most discriminatory tools were Gross Motor Function Measure (GMFM-66), velocity, and WeeFIM Mobility. Between Levels II and III, the most discriminatory tools were GMFM Dimension E, Pediatric Functional Independence Measure (WeeFIM) Self-Care and Mobility, cadence, and Gillette Functional Assessment Questionnaire Question 1. Large ESIs were noted for Parent and Child reports of Pediatric Outcomes Data Collection Instrument (PODCI) Sports & Physical Function, Parent report of PODCI Global Function, GMFM Dimension E, and GMFM-66 across all GMFCS level comparisons. The least discriminatory tools were the Quality of Life and cognition measures; however, these are important in comprehensive assessments of treatment effects.

See end of paper for list of abbreviations.

Although others have investigated relations between outcome tools or the relation of an outcome score to function in children with cerebral palsy (CP), few have investigated how well the tools discriminate among levels of severity based on the Gross Motor Function Classification System (GMFCS). The aim of this study was to evaluate the discriminatory ability of outcome tools commonly used in pediatric orthopedics. This work is Part II of a prospective, cross-sectional study of ambulatory patients with CP (Oeffinger et al. Part I, p 172).

The authors published a retrospective study that demonstrated relations between GMFCS level and the Pediatric Outcomes Data Collection Instrument (PODCI), Gross Motor Function Measure (GMFM) Dimensions D (standing) and E (walking, running, jumping), energy cost during walking (O_2 cost), and temporal–spatial gait parameters.[1] In that study, tools were not collected concurrently or on all participants, which prevented comparison among tools. A prospective follow-up study (Oeffinger et al.), including more outcome tools collected during a single session, confirmed the earlier findings.

The specific aim of Part II of this work was to identify outcome tools that discriminate among GMFCS Levels I, II, and III. It was hypothesized that among study outcome tools some would be more discriminating than others, as demonstrated by greater differences among GMFCS levels and higher ratings on measures of sensitivity and specificity. This information is critical to the selection of appropriate outcome tools to be used in future evaluations of treatment effects in this population.

Method

This was a 3-year prospective multicenter study of ambulatory participants with CP described in Oeffinger et al. The 562 participants in this analysis include 240 classified as GMFCS Level I, 196 as GMFCS Level II, and 126 as GMFCS Level III; 339 males, 223 females; 400 with diplegia, 162 with hemiplegia (age range 4–18y, mean 11y 1mo [SD 3y 7mo]). The outcome tools evaluated were the Gillette Functional Assessment Questionnaire (FAQ),[2] the Gross Motor Function Measure (GMFM),[3] the Pediatric Quality of Life Inventory (PedsQL),[4] the Pediatric Outcomes Data Collection Instrument (PODCI),[5] the Pediatric Functional Independence Measure (WeeFIM),[6] temporal–spatial gait parameters, and O_2 cost. For a full description of study participants, tools, and data collection refer to Oeffinger et al.

DATA ANALYSIS

Analyses of variance (ANOVAs) were performed with GMFCS level as the independent variable and the scores for each tool subscale as the dependent variables.

The effect size index (ESI) of each parametric outcome score was calculated based on the ANOVAs. The ESI is a standardized value that quantifies the magnitude of difference between groups as the ratio of the difference of the means relative to the estimate of the common standard deviation. In this study, the ESI shows the magnitude of the difference between groups in number of SDs for each pair of GMFCS levels. Effect sizes of 0.2 are generally considered small, 0.5 are considered medium, and 0.8 are considered large.[7]

The odds ratio (OR) of each non-parametric outcome score was calculated. The OR expresses the odds of a particular outcome when a condition is present to the odds of that outcome when the condition is absent. In this study, the OR is a numeric value that represents the number of times more likely the respondent will belong to one GMFCS level compared with

another based on the 'yes' response to a subscale item.[8] An OR of 1.0 signifies an equal likelihood of belonging to either group. An OR of 0.5 or 2.0 reflects two times the likelihood of belonging to one group compared with the other.

Two binary logistic regression models with stepwise selection were fitted to two different subsets of data: discriminating GMFCS Level I versus II, and discriminating GMFCS Level II versus III. In each case, the significance level for variable entry was set to 0.15, and the significance level to stay was set to 0.05. Measures from the child tools of PODCI and PedsQL were not included in the regression analyses owing to a significant amount of missing data from children below the age range of these tools. O_2 cost was also not included owing to limited sample size for this assessment. In addition, Parent report of PODCI Expectation and Satisfaction scores were not included owing to limited validity of these domains. Receiver operating characteristic (ROC) curves were then constructed to determine sensitivity and specificity for the most discriminant measures identified by the regression models.

Table I: Effect size index (ESI) derived from analysis of variance (ANOVA) for each outcome tool subscale measured on an interval level (values in bold type are defined as large effect sizes, ≥0.8) by Gross Motor Function Classification System (GMFCS) level comparisons

Outcome tool subscale	*GMFCS level*		
	I vs II ESI	*I vs III ESI*	*II vs III ESI*
PODCI (Parent report)			
Global Function	**0.8**	**1.6**	**0.8**
Upper Extremity Function	0.5	**0.8**	0.4
Transfers & Basic Mobility	0.7	**2.1**	**1.4**
Sports & Physical Function	**1.0**	**1.9**	**0.9**
Pain/Comfort	0.2	0.2	0.0
Happiness	0.1	0.0	0.1
PODCI (Child report)			
Global Function	0.7	**1.3**	0.6
Upper Extremity Function	0.4	0.7	0.3
Transfers & Basic Mobility	0.6	**1.8**	**1.2**
Sports & Physical Function	**0.8**	**1.7**	**0.9**
Pain/Comfort	0.3	0.2	0.1
Happiness	0.2	0.1	0.1
PedsQL (Parent report)			
Physical Functioning	0.7	**0.9**	0.2
Emotional Functioning	0.1	0.1	0.1
Social Functioning	0.3	0.3	0.1
School Functioning	0.2	0.3	0.1
PedsQL (Child report)			
Physical Functioning	0.5	0.7	0.2
Emotional Functioning	0.1	0.1	0.0
Social Functioning	0.3	0.1	0.2
School Functioning	0.2	0.1	0.1
WeeFIM			
Self-Care	0.5	**1.0**	0.5
Mobility	0.5	**1.7**	**1.2**
Cognition	0.3	0.3	0.0
GMFM			
Dimension D (standing)	**0.8**	**3.3**	**2.5**
Dimension E (walking, running, jumping)	**1.3**	**4.4**	**3.1**
GMFM-66	**1.5**	**3.3**	**1.8**
Gait analysis			
O_2 cost measured	0.7	**1.9**	**1.2**
Cadence (% normal)	0.3	**1.5**	**1.3**
Stride length (% normal)	0.6	**1.7**	**1.1**
Velocity (% normal)	0.7	**2.1**	**1.4**
Gillette FAQ			
Question 1	0.7	**1.7**	**1.0**

PODCI, Pediatric Outcomes Data Collection Instrument;[5] PedsQL, Pediatric Quality of Life Inventory;[4] WeeFIM, Pediatric Functional Independence Measure;[6] GMFM, Gross Motor Function Measure;[3] Gillette FAQ, Functional Assessment Questionnaire.[2]

Results

PARAMETRIC RESULTS

The ESIs for comparison of parametric scores among GMFCS levels are shown in Table I. Subscales with large effect sizes (≥0.8) across all GMFCS level comparisons were the PODCI Sports and Physical Function by Parent and Child report, PODCI Global Function by Parent report, GMFM Dimensions D and E, and the GMFM as scored by the Gross Motor Activity Estimator software (GMFM-66[3]). Figure 1 shows a box plot of the GMFM-66 demonstrating good separation of the means among all three GMFCS levels.

Subscales with at least medium effect sizes (≥0.5) across all GMFCS level comparisons were the PODCI Transfers & Basic Mobility by Parent and Child report, PODCI Global Function by Child report, WeeFIM Self-Care and Mobility, O_2 cost, stride length, velocity, and FAQ Question 1. Specifically, these tools demonstrated large ESIs between GMFCS Levels I and III, medium to large ESIs between Levels II and III, and medium ESIs between Levels I and II.

The tools with the smallest effect sizes, ranging from 0 to 0.3, for all GMFCS levels were the Parent and Child reports of

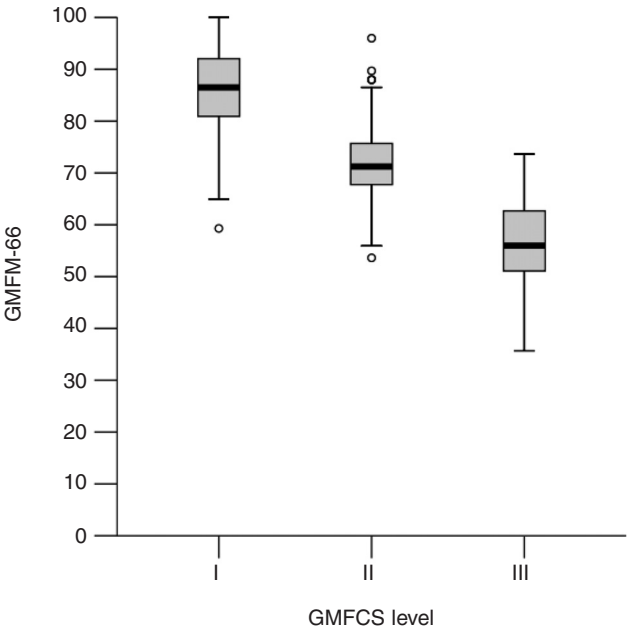

Figure 1: *Box plot of Gross Motor Function Measure (GMFM-66) shows minimal ceiling effects among Gross Motor Function Classification System (GMFCS) levels, with excellent discrimination among all levels.*

PODCI Pain/Comfort and Happiness, Parent and Child reports of PedsQL Emotional Functioning, Social Functioning, and School Functioning, and WeeFIM Cognition.

NON-PARAMETRIC RESULTS

The ORs for comparison of non-parametric scores among GMFCS levels are shown in Table II. For FAQ Question 2, a 'yes' response indicates limitations of function; therefore a low OR (≤0.5) indicates a higher likelihood that a participant will be in the lower functioning GMFCS level. For FAQ Question 3, a 'yes' response indicates ability to perform a skill, not limitations; therefore, a high OR (≥2.0) indicates a higher likelihood that a participant will be in the higher functioning GMFCS level.

For FAQ Question 2, the lowest OR for GMFCS Levels I versus II, and I versus III were noted for concerns about safety and balance. Participants in GMFCS Level I had the least likelihood of identifying safety or balance concerns that limited their walking ability. For FAQ Question 3, the tasks with the highest OR among all GMFCS levels were: walk up and down curb independently; runs; get on/off bus by him- or herself; jumps off single step; rides 2-wheel bicycle; and rides an escalator without help. Three FAQ Question 3 tasks with high OR between GMFCS Levels I and III, and Levels II and III, were: walk carrying an object; walk carrying a fragile object; and walk up and down stairs with railing.

REGRESSION MODELS AND ROC RESULTS

ROC curves showing sensitivity and specificity to discriminate between GMFCS Level I versus II and the regression equations for both GMFCS Level I versus II, and II versus III are shown in Figure 2. Table III illustrates the effect of incrementally adding each discriminant tool subscale to the regression model on area under the curve (AUC), sensitivity, and specificity.

Based on stepwise regression analysis, GMFM-66, velocity, and WeeFIM Mobility together best discriminate GMFCS Level I from II (Table III). All three have a positive effect in predicting GMFCS Level I from II. The chance that a child is GMFCS Level I increases by 24%, 3%, and 5% when his or her GMFM-66, velocity, and WeeFIM Mobility scores increase by 1 respectively. The AUC is 91.3%, with sensitivity 86.6%, and specificity 82.8 %.

Similarly, stepwise regression selected GMFM Dimension E, WeeFIM Self-Care, cadence, FAQ Question 1, and WeeFIM Mobility as best discriminating GMFCS Level II from III (Table III). The discriminant rule selected by stepwise logistic regression for GMFCS Level I versus II is 0.0528 × WeeFIM Mobility + 0.2134 × GMFM-66+0.0244 × Velocity-23.9123=0. A patient is in GMFCS Level I if >0, and in Level II if <0. The chance that a child is GMFCS Level II increases by 10, 4, 53, and 7%, when his or her GMFM Dimension E, cadence, FAQ Question 1, WeeFIM Mobility scores increase by 1 respectively. WeeFIM Self-Care has a negative effect so that the chance that a child is GMFCS Level II decreases by 8% when his or her WeeFIM Self-Care score increases by 1. The AUC is 96.4%, with sensitivity 90.5%, and specificity 87.3%.

On further analysis, if only one outcome tool was selected to discriminate GMFCS Levels I from II and II from III, the

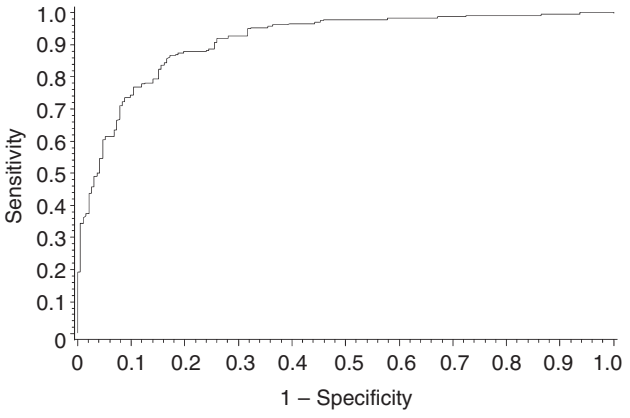

Figure 2: *The Receiver Operating Characteristic curve predicting Gross Motor Function Classification System Level I from II using Gross Motor Function Measure , Pediatric Functional Independence Measure Mobility, and velocity demonstrates high sensitivity and high specificity.*

Table II: Odds ratios (ORs) for each task of the Gillette Functional Assessment Questionnaire (FAQ) Question 2 and Question 3, by Gross Motor Function Classification System (GMFCS) level comparisons (values in bold type are defined as ORs, ≥2.0 or ≤0.5)

Outcome tool subscale	*ORs*		
	GMFCS I vs II	*GMFCS I vs III*	*GMFCS II vs III*
FAQ Question 2			
1) Pain	0.7	1.2	1.9
2) Weakness	0.6	0.6	1.0
3) Safety concerns	**0.3**	**0.3**	0.8
4) Balance	**0.3**	**0.3**	0.9
5) Mental ability	**0.5**	0.7	1.4
6) Endurance	0.7	0.8	1.2
7) Other	**0.4**	**0.2**	0.6
FAQ Question 3			
1) Walk carrying an object	0.9	**5.4**	**6.2**
2) Walk carrying a fragile object	1.5	**10.8**	**7.2**
3) Walk up and down stairs with railing	0.5	**5.3**	**11.7**
4) Walk up and down stairs without railing	**4.9**	**8.6**	1.8
5) Walk up and down curb independently	**3.7**	**7.7**	**2.1**
6) Runs	**2.7**	**10.2**	**3.8**
7) Runs corners well with good control	**3.5**	**4.8**	1.4
8) Can take steps backwards	1.8	**2.9**	1.6
9) Can maneuver in tight areas	1.9	**2.2**	1.2
10) Get on and off a bus by him/herself	**3.4**	**11.4**	**3.4**
11) Jumps rope	**3.4**	[a]	[a]
12) Jumps off a single step	**4.1**	**20.8**	**5.0**
13) Hop on right foot	**4.2**	**5.3**	1.3
14) Hop on left foot	**3.7**	**5.5**	1.5
15) Step over an object, right foot first	1.3	1.9	1.5
16) Step over an object, left foot first	1.1	1.8	1.7
17) Kick a ball with right foot	1.4	1.5	1.1
18) Kick a ball with left foot	1.6	1.4	0.9
19) Ride 2-wheel bicycle	**5.1**	**21.0**	**4.1**
20) Ride 3-wheel bicycle	**2.2**	**3.6**	1.7
21) Ice skate or roller skate	**3.4**	**3.2**	0.9
22) Ride an escalator, without help	**2.9**	**11.1**	**3.9**

[a]Insufficient number of affirmative responses.

GMFM-66 provides excellent discrimination with an AUC of 0.903 and 0.935. The discriminant rule selected by stepwise logistic regression for GMFCS Level II versus III is 0.0632 × WeeFIM Mobility × 0.0804 × WeeFIM Self-Care + 0.0908 × GMFM E + 0.0406 × cadence + 0.4281 × Gillette Functional Assessment Questionnaire Question 1 × 14.97720=0. A patient is in GMFCS Level II if >0, and in Level III if <0.

CEILING EFFECTS

The percentages of participants scoring at the maximum for each subscale by GMFCS level are shown in Table IV. Tool subscales identified as having more than 20% of the participants within a GMFCS level who scored a maximum value were defined as having ceiling effects.

Subscales with ceiling effects for GMFCS Levels I, II, and III were PODCI Pain/Comfort by both Parent and Child report and WeeFIM Cognition. Ceiling effects were present for GMFCS Levels I and II for Child reports of PODCI Upper Extremity Function and Transfers and Basic Mobility, and WeeFIM Self-Care. WeeFIM Mobility, GMFM dimension D, and FAQ Question 1 demonstrated ceiling effects only in GMFCS Level I.

Among the subscales that differentiated among all GMFCS levels, PODCI Sports and Physical Function and Global Function by both Child and Parent report, and GMFM Dimension E and GMFM-66 demonstrated no ceiling effects. Among the subscales included in the discriminant analyses, WeeFIM Mobility and Self-Care demonstrated ceiling effects.

Discussion

In this multicenter prospective cross-sectional study, the GMFCS was used to classify levels of severity in 562 patients with CP. Scores measuring physical function were able to differentiate among GMFCS levels, whereas those measuring cognition and Quality of Life were not able to differentiate among GMFCS levels.

For measures of Activities and Participation, either an observed test of function or a questionnaire may be effective in capturing this International Classification of Functioning, Disability and Health (ICF) component. Consistent with Oeffinger et al.,[1] GMFM Dimensions D and E, GMFM-66, Parent and Child reports of PODCI Sports and Physical Function, Parent report of PODCI Global Function, and six of the 22 skills in FAQ Question 3 could differentiate among all GMFCS levels. GMFM Dimension E demonstrated the largest ESI. The relation between the GMFCS and GMFM is expected since the GMFCS is based in part on items of the GMFM.[9] The FAQ Question 3 skills that differentiated between Levels I and II, and Levels I and III, are high-level functional skills (e.g. hopping and skating) that require little or no impairment in balance or motor control. The items that differentiated between Levels I and III, and Levels II and III, require the use of hands and because Level III individuals use assistive devices their ability to perform these skills is limited.

Within the ICF component of Body Functions & Structures, O_2 cost could differentiate among all three GMFCS levels with at least a medium effect size. Items in FAQ Question 2 can differentiate GMFCS Level I from Levels II and III based on safety and balance issues. Additional measurements may be needed to define fully this ICF component.

The subscales of PedsQL Emotional, Social, and School Functioning, and PODCI subscales of Pain/Comfort and Happiness, focus on Quality of Life and were unable to differentiate among GMFCS levels. Using the GMFCS may have favored the functional domains of the outcome tools, rather than the Quality of Life elements, or these results could truly reflect no difference among the GMFCS levels in Quality of Life.

Logistic regression was able to identify the combination of factors that could best discriminate between GMFCS Levels I and II, and between Levels II and III, with high AUC and levels of sensitivity and specificity. The factors that best discriminate are measures of mobility. More than 90% of the AUC was determined by the GMFM-66 (Level I vs II) or GMFM Dimension E (Level II vs III) with minor increases from additional subscales. For example, adding velocity to GMFM-66 resulted in a 0.4% increase in AUC, and adding WeeFIM Self-Care to GMFM Dimension E resulted in a 0.6% increase AUC (Table III).

Motor control curves from the GMFM[10] demonstrate larger differences between Levels II and III than between Levels I and II once development reaches a stable point. The GMFM-66 equalizes the difference between Levels I and II, and II and III. Therefore, the interval GMFM-66 discriminates better between Levels I and II whereas the 'unconverted' measure of GMFM Dimension E discriminates better between Levels II and III. This is consistent with the ESI analysis. Using the GMFM-66 instead of GMFM Dimension E results in only a 0.5% decrease in the AUC for discrimination between Levels II and III. Therefore, the GMFM-66 alone is able to effectively discriminate between both GMFCS Levels I and II, and II and III.

Table III: Best discriminant analysis for Gross Motor Function Classification System (GMFCS) Level I versus II, and II versus III

Discriminating GMFCS levels	*Discriminant scores*	*Standardized coefficient*	*p value*	*OR (95% CI)*	*Fitted n*	*AUC*	*Sensitivity (%)*	*Specificity (%)*
I vs II	GMFM-66	1.2	<0.0001	1.2 (1.2–1.3)	434	0.903	85.8	79.9
(*n*=437)	Velocity	0.2	0.009	1.0 (1.0–1.0)	434	0.907	85.0	83.0
	WeeFIM Mobility	0.2	0.04	1.1 (1.0–1.1)	430	0.913	86.6	82.8
II vs III	GMFM Dimension E	1.4	<0.0001	1.1 (1.1–1.1)	321	0.940	93.3	82.2
(*n*=321)	WeeFIM Self-Care	–0.8	<0.0001	0.9 (0.9–1.0)	309	0.946	90.2	87.9
	Cadence	0.4	0.008	1.0 (1.0–1.1)	308	0.950	87.6	87.8
	FAQ Question 1	0.3	0.02	1.5 (1.1–2.2)	299	0.962	88.9	88.2
	WeeFIM Mobility	0.4	0.045	1.1 (1.0–1.1)	299	0.964	90.5	87.3

GMFM, Gross Motor Function Measure; WeeFIM, Pediatric Functional Independence Measure; FAQ, Gillette Functional Assessment Questionnaire; CI, confidence interval; AUC, area under the curve.

Discriminatory ability may be limited by ceiling effects. Tools with high ceiling effects suggest either that participants may not have significant limitations in these areas, or that there may be a need for a more discriminant subscale. As an example, the box plot of the Child report of PODCI Upper Extremity Function scores (Fig. 3) shows no upper fence for either GMFCS Levels I or II. The ceiling effects and interquartile overlap in this subscale limit its discriminatory ability. Parent and Child reports of PODCI Pain/Comfort and WeeFIM Cognition had high ceiling effects for all three GMFCS levels, suggesting that pain and limitations in cognition are not major issues in this population.

This study provides information on the discriminatory ability of outcome tools that can help clinicians select measures to monitor patient treatment goals. Best practice and reduced cost may be obtained by administering tools that discriminate among severity and avoid ceiling effects. These results may also be applied in the research setting by identifying which tools are the best functional discriminators based on sensitivity and specificity. Also, in developing a study, a power analysis based on a highly discriminatory tool will minimize the required patient sample size. The raw data reported in Parts I and II of this work provide the means, variances, and effect sizes necessary to conduct a priori power analyses.

Primary limitations of this study have been discussed in Part I (Oeffinger et al.) and were addressed in the study design to minimize their effects on results and conclusions. The discriminant analysis was limited in part by the broad age range in this study. Several participants were too young to complete the child versions of the PODCI or PedsQL so these data were not included in the regression analyses.

Table IV: Percentage of participants reaching highest score (ceiling) for each outcome tool subscale, by Gross Motor Function Classification System (GMFCS) level (values in bold type ≥20%)

Outcome tool subscale	*GMFCS level*		
	I	*II*	*III*
	% ceiling	*% ceiling*	*% ceiling*
PODCI (Parent report)	(*n*=240)	(*n*=196)	(*n*=123)
Global Function	1	0	0
Upper Extremity Function	13	4	6
Transfers & Basic Mobility	19	7	1
Sports & Physical Function	1	0	0
Pain/Comfort	**38**	**31**	**28**
Happiness	16 (*n*=234)	9 (*n*=193)	16 (*n*=121)
PODCI (Child report)	(*n*=117)	(*n*=80)	(*n*=51)
Global Function	4	0	0
Upper Extremity Function	**53**	**39**	18
Transfers & Basic Mobility	**58**	**25**	4
Sports & Physical Function	7	1	0
Pain/Comfort	**41**	**30**	**45**
Happiness	**24**	19	**25**
PedsQL (Parent report)	(*n*=239)	(*n*=194)	(*n*=124)
Physical Functioning	2	1	0
Emotional Functioning	3	2	4
Social Functioning	4	2	2
School Functioning	3	2	2
PedsQL (Child report)	(*n*=224)	(*n*=166)	(*n*=105)
Physical Functioning	7	2	1
Emotional Functioning	8	10	10
Social Functioning	10	3	3
School Functioning	4	4	6
WeeFIM	(*n*=237)	(*n*=196)	(*n*=126)
Self-Care	**43**	**20**	12
Mobility	**40**	8	4
Cognition	**56**	**41**	**44**
GMFM	(*n*=239)	(*n*=195)	(*n*=119)
Dimension D	23	2	0
Dimension E	13	0	0
GMFM-66	13	0	0
Gillette FAQ	(*n*=238)	(*n*=190)	(*n*=119)
Question 1	**45**	15	2

PODCI, Pediatric Outcomes Data Collection Instrument;[5] PedsQL, Pediatric Quality of Life Inventory;[4] WeeFIM, Pediatric Functional Independence Measure;[6] GMFM, Gross Motor Function Measure;[3] Gillette FAQ, Functional Assessment Questionnaire.[2]

Conclusions

This multicenter prospective study examined the discriminatory ability of a variety of outcome tools used in the treatment of patients with CP. Discrimination was evaluated based on classification of the participants into GMFCS Levels I to III. The data provide a comprehensive analysis of how the tools studied, which represent ICF Activities & Participation, ICF Body Functions & Structures, and Quality of Life elements, discriminate within a large ambulatory CP population.

Between Levels I and II, the most discriminatory tools were the GMFM-66, velocity, and WeeFIM Mobility. Between Levels II and III, the most discriminatory tools were GMFM

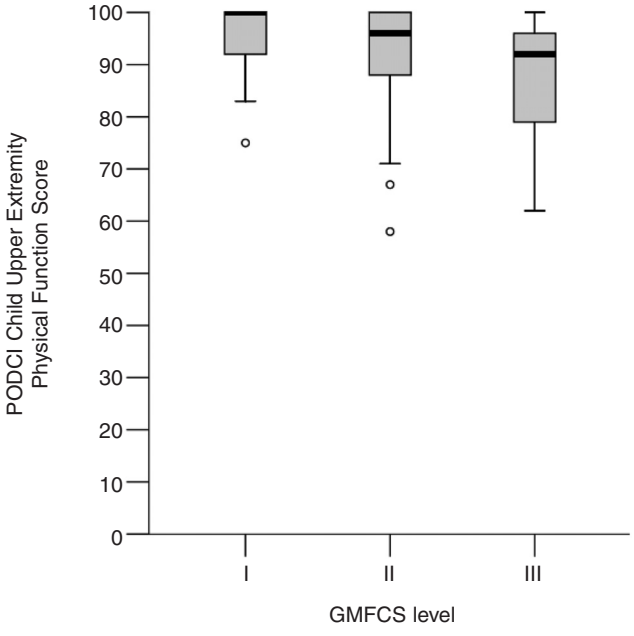

Figure 3: *Box plot of Pediatric Outcomes Data Collection Instrument (PODCI) Child report of Upper Extremity Function shows significant overlap among scores for three Gross Motor Function Classification System (GMFCS) levels, significant ceiling effects for GMFCS Levels I and II, and only moderate discriminatory ability between Levels I and III.*

Dimension E, WeeFIM Self-Care, cadence, FAQ Question 1, and WeeFIM Mobility. Large ESIs were noted for Parent and Child reports of PODCI Sports and Physical Function, Parent report of PODCI Global Function, GMFM Dimensions D and E, and GMFM-66 across all GMFCS level comparisons. In general, PODCI Child report and WeeFIM had the greatest percentage of ceiling effects. The least discriminatory tools were the Quality of Life and cognition measures. Cognition and Quality of Life issues need to be considered in a comprehensive assessment of treatment effects, but should not be expected to discriminate based on level of functional impairment.

The results of this work set the foundation for long-term studies to measure the effects of interventions on ambulatory patients with CP. More work is needed to expand the number of tools analyzed for discriminatory ability and ceiling and floor effects. Additional studies are needed to identify which tools are redundant, which can show differences in functional profiles (e.g. spastic vs dyskinetic), and which tools will be responsive to change over time.

Accepted for publication 11th September 2006.

Acknowledgments

This work was funded by Shriners Hospitals for Children, 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 (Chester Tylkowski MD, Donna Oeffinger PhD, Rebecca Davis BS, Sarah Rogers MPH, Richard Kryscio PhD, George Gorton BS, Janine Calmes MS PT, Douglas Barnes MD, Richard Haynes MD, Diane Nicholson PhD PT, Diane Damiano PhD PT, Mark Abel MD, and Anita Bagley PhD) and the contributions of the study coordinators at each site. We also acknowledge the patients and their families for their participation.

References

1. Oeffinger D, Tylkowski C, Rayens M, Davis R, Gorton G, D'Astous J, Nicholson D, Damiano D, Abel M, Bagley A, Luan J. (2004) Gross Motor Function Classification System and outcome tools for assessing ambulatory cerebral palsy: a multicenter study. *Dev Med Child Neurol* **46:** 311–319.
2. Novacheck TF, Stout JL, Tervo R. (2000) Reliability and validity of the Gillette Functional Assessment Questionnaire as an outcome measure in children with walking disabilities. *J Pediatr Orthop* **20:** 75–81.
3. Russell D, Rosenbaum P, Avery L, Lane M. (2002) *Gross Motor Function Measure (GMFM-66 & GMFM-88) User's Manual. Clinics in Developmental Medicine No. 159.* London: Mac Keith Press.
4. Varni J, Seid M, Rode C. (1999) The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care* **37:** 126–139.
5. Daltroy LH, Liang MH, Fossel AH, Goldberg MJ. (1998) The POSNA pediatric musculoskeletal functional health questionnaire: report on reliability, validity, and sensitivity to change. Pediatric Outcomes Instrument Development Group. Pediatric Orthopaedic Society of North America. *J Pediatr Orthop* **18:** 561–571.
6. McCabe M, Granger C. (1990) Content validity of a pediatric functional independence measure. *Appl Nurs Res* **3:** 120–122.
7. Portney L, Watkins, MP. (2000) *Foundations of Clinical Research: Applications to Practice.* 2nd edn. Upper Sadle River: Prentice Hall Health.
8. Piantadosi S. (1997) *Clinical Trials – A Methodological Perspective.* New York: John Wiley & Sons.
9. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. (1997) Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol* **39:** 214–223.
10. Rosenbaum PL, Walter SD, Hanna SE, Palisano RJ, Russell DJ, Raina P, Wood E, Bartlett DJ, Galuppi BE. (2002) Prognosis for gross motor function in cerebral palsy: creation of motor development curves. *J Am Med Assoc* **288:** 1357–1363.

List of abbreviations

AUC	Area under the curve
ESI	Effect size index
FAQ	Gillette Functional Assessment Questionnaire
GMFM-66	Gross Motor Function Measure Score calculated using the Gross Motor Activity Estimator Program
ICF	International Classification of Functioning, Disability and Health
O_2	Energy (oxygen) cost
OR	Odds ratio
PedsQL	Pediatric Quality of Life Inventory
PODCI	Pediatric Outcomes Data Collection Instrument
ROC	Receiver operating characteristic
WeeFIM	Pediatric Functional Independence Measure