

Validity and clinimetric properties of the Spinal Alignment and Range of Motion Measure in children with cerebral palsy

CHIA-LING CHEN^{1,2} | KATIE P H WU¹ | WEN-YU LIU³ | HSIN-YI KATHY CHENG² | I-HSUAN SHEN⁴ | KEH-CHUNG LIN⁵

1 Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital at Linkou; **2** Graduate Institute of Early Intervention; **3** Department of Physical Therapy; **4** Department of Occupational Therapy, College of Medicine, Chang Gung University, Tao-Yuan; **5** School of Occupational Therapy, College of Medicine, National Taiwan University and Division of Occupational Therapy, Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan.

Correspondence to Dr Chia-Ling Chen, Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital, 5 Fu-Hsing Street Kwei-Shan, Tao-Yuan 333, Taiwan.
E-mail: clingchen@gmail.com

This article is commented on by Bartlett on pages 685–686 of this issue.

PUBLICATION DATA

Accepted for publication 26th February 2013.
Published online 17th April 2013.

ABBREVIATIONS

ADL	Activities of daily living
MCID	Minimal clinically important difference
MDC	Minimal detectable change
ROM	Range of motion
SAROMM	Spinal Alignment and Range of Motion Measure
WeeFIM	Functional Independence Measure for Children

AIM The aim of this study was to assess the validity, responsiveness, and clinimetric properties of the Spinal Alignment and Range of Motion Measure (SAROMM) in children with cerebral palsy (CP).

METHOD Sixty-two children with CP (40 males, 22 females) with a median age of 3 years and 11 months (range 1–6y) and their caregivers participated in this study. Among the children, 56 had spastic CP while six had non-spastic CP; 53 had bilateral CP, while nine had unilateral limb involvement. Thirty-three children were classified as Gross Motor Function Classification System (GMFCS) levels I to III and 23 as levels IV or V. Fifty-six children (90%) received regular rehabilitation by means of regular physical or occupational therapy (50% once or twice per week and 40% more than two times per week) and six children (10%) received irregular rehabilitation (less than once a week). Construct validity was determined by assessing the strength of the correlation between the spinal alignment SAROMM (SAROMM-SA), the range of motion SAROMM (SAROMM-ROM), and the total SAROMM (SAROMM-total), and construct measures, including the 66-item Gross Motor Function Measure (GMFM-66) and Functional Independence Measures for Children (WeeFIM), at baseline and at 6-months follow-up. Responsiveness was examined using effect size. Minimal detectable change (MDC) at the 90% confidence level (MDC90) and minimal clinically important difference (MCID) were analysed.

RESULTS The SAROMM with the GMFM-66 and WeeFIM had fair to good construct validity. The effect size values of all SAROMM scales were 0.24 to 0.48. The MDC90 values and MCID range were 1.43 and 0.47 to 1.67 for the SAROMM-SA, 3.12 and 3.68 to 4.07 for the SAROMM-ROM, and 3.22 and 4.53 to 4.62 for the SAROMM-total.

INTERPRETATION The clinimetric properties of the SAROMM allow clinicians to determine whether a change in SAROMM score represents a clinically meaningful change.

Cerebral palsy (CP) describes a group of permanent disorders of movement and posture that result from non-progressive disturbances in the developing fetal or infant brain.¹ Although the pathophysiology of CP is non-progressive, musculoskeletal structures often deteriorate with advancing age.^{1,2} Tightness of spastic muscles, spinal malalignment, and joint contractures are common and significant CP-related problems in children. Children with severe CP often have more musculoskeletal problems than those with mild CP. For instance, spinal malalignment and limited range of motion (ROM) are strongly correlated with a decline in gross motor function,³ Gross Motor Function Classification System (GMFCS) level,⁴ and activities of daily living (ADL).⁵ Therefore, keeping good spinal

alignment and preventing ROM limitations may contribute to a reduced decline in gross motor capacity³ and ADL⁵ in children and adolescents with CP and to an increase in their participation.

Conventionally, ROM is assessed using a universal goniometer; however, test–retest results vary widely, especially in children with CP.⁶ The Spinal Alignment and Range of Motion Measure (SAROMM), with sufficient intrarater and interrater reliability and construct validity, was developed primarily as a discriminative tool for children and adolescents with CP aged 2 to 18 years.⁴ As an alternative to measuring the ROM of all joints with a goniometer, the SAROMM may be simpler to apply and a more meaningful indicator of whether or not a child has

normal alignment and ROM.⁴ The SAROMM subscales are the spinal alignment subscale (SAROMM-SA) for cervical, thoracic, and lumbar regions, and the scoliosis and range of motion subscale (SAROMM-ROM) for the hip, knee, ankle, and upper extremities.⁴ Currently, the SAROMM is used to describe the pattern of restrictions across different body regions.⁴ However, no study has determined the clinimetric properties (e.g. the minimal detectable change [MDC], minimal clinically important difference [MCID], and responsiveness) of the SAROMM in young children with CP.

The Consensus-based Standards for the Selection of Health-status Measurement Instruments (COSMIN) study reached international consensus on measurement properties for health-related outcomes: measurement error, validity, responsiveness, and interpretability (e.g. MCID).^{7,8} Responsiveness, MDC, and MCID are commonly used to signify an important difference after treatment.⁹ Responsiveness is defined as the ability of an instrument to detect change over time in the construct being measured,^{7,8} and is a measure of longitudinal validity.¹⁰ The MDC, which indicates the smallest amount of change beyond measurement error, reflects a true difference and a statistically reliable change.⁹ The MCID depicts the smallest change between two scores which is considered important from a client's or clinician's perspective.⁹ The concept of an MCID is offered as the new standard for determining the effectiveness of a given treatment and describing patient satisfaction with that treatment.¹¹ Therefore, the MDC and MCID provide clinicians with relevant information for clinical decision-making when interpreting changes in scores posttreatment or at follow-up.

Understanding the clinimetric properties of the SAROMM allows clinicians to determine whether changes posttreatment or at follow-up are significant and important for participants with CP. Differences between pretreatment or baseline scores and follow-up or posttreatment scores reflect not only response to treatment but also the clinical course (typically an increase in restriction over time). Determination of the MDC and MCID values of the SAROMM-SA and SAROMM-ROM is important for clinical decision-making and for ascertaining the clinical course and therapeutic effectiveness in children with CP. Thus, the aim of this study was to assess the construct validity of the SAROMM and its clinimetric properties, including its responsiveness, MDC, and MCID, in children with CP.

METHOD

Participants

Young children with CP from the rehabilitation clinics of three tertiary hospitals in Taiwan (Taipei, Linkou, and Kaohsiung branches of Chang Gung Memorial Hospital) were consecutively recruited to this longitudinal follow-up study. All participants underwent independent examinations by a physiatrist and physical therapist to determine their eligibility. Inclusion criteria were a diagnosis of CP and age of 1 to 6 years. Exclusion criteria were the

What this paper adds

- The SAROMM had fair to good construct validity for young children with CP.
- The SAROMM was responsive to change.
- Minimal detectable change and minimal clinically important change were identified.
- Clinimetric properties of the SAROMM allow clinicians to determine whether a change in SAROMM score represents a 'true' clinical change.

presence of a progressive neurological, genetic, or metabolic disorder, or a severe concurrent illness or disease not typically associated with CP, such as traumatic brain injury or active pneumonia. Sixty-five children were initially recruited, but only 62 enrolled (40 males, 22 females): one had active medical problems and two were lost to follow-up (Table I). Approval from the Institutional Review Board for Human Studies and informed consent from the parents and caregivers were obtained before this study.

Measures

Measures administered were the SAROMM,⁴ the 66-item Gross Motor Function Measure (GMFM-66),¹² and the Functional Independence Measure for Children (WeeFIM).¹³ The SAROMM contains 26 items, four items

Table I: Demographic and clinical characteristics at baseline and follow-up (*n*=62)

Characteristics	Value ^a	
	Baseline	Follow-up
Demographic		
Age, y:mo	3y 9mo [1y 7mo]	
Sex		
Male	40 (65)	
Female	22 (36)	
Clinical		
Limb distribution		
Bilateral	53 (86)	
Unilateral	9 (15)	
Subtypes		
Spastic	56 (90)	
Non-spastic	6 (9)	
GMFCS levels		
I	19 (31)	
II	11 (18)	
III	9 (15)	
IV	13 (21)	
V	10 (16)	
GMFM-66 score	49.8 [21.5]	53.6 [21.7]
SAROMM		
Spinal alignment	2.1 [3.3]	1.7 [2.8]
Range of motion	21.3 [7.4]	20.2 [9.3]
Total	23.4 [9.9]	21.9 [11.4]
WeeFIM		
Self-care	21.3 [14.4]	24.5 [14.9]
Mobility	18.3 [11.2]	20.0 [11.4]
Cognition	17.3 [10.0]	20.1 [10.9]
Total	56.9 [33.8]	64.6 [35.3]

^aValues are expressed as mean [SD] for continuous variables and number (%) for categorical variables. GMFCS, Gross Motor Function Classification System; GMFM-66, 66-item Gross Motor Function Measure; SAROMM, Spinal Alignment and Range of Motion Measure; WeeFIM, Functional Independence Measure for Children.

for spinal alignment, and 11 items for ROM and muscle extensibility, which is tested bilaterally. The manual for the SAROMM can be downloaded from the *CanChild* website (<http://www.canchild.ca/en/measures/saromm.asp>). Each item is scored on a five-point ordinal scale, ranging from 0 (ability to align normally with no passive limitations) to 4 (severe deviations in spinal alignment, limitations in joint ROM, or limitations in muscle extensibility). The SAROMM-SA and SAROMM-ROM subscales are scored by summing the scores for items 1 to 4 (possible range 0–16) and for items 5 to 26 (possible range 0–88) respectively. The SAROMM-total score is obtained by summing the SAROMM-SA and SAROMM-ROM scores (possible range 0–104). The intraclass correlation coefficients (ICCs) for interrater and test–retest reliabilities for all SAROMM subscales and total scores for children and adolescents with CP were all more than 0.80.⁴

The GMFM is recognized as the criterion standard for evaluation of quantitative changes in gross motor function in children with CP. Each item on the GMFM is graded on a four-point scale (0, child unable to initiate the task; 1, child initiates the task; 2, child partially completes the task; and 3, child completes the task). Sixty-six items in the GMFM subset were graded for difficulty using Rasch analysis, with a maximum score of 100.¹² The GMFM-66 score was obtained using Gross Motor Ability Estimator software.¹² The GMFM-66 is best suited for children who can walk and has good validity and reliability.¹²

The WeeFIM, which has excellent reliability,¹³ comprises 18 items on three functional subscales assessing self-care (eight items), mobility (WeeFIM-MO, five items), and cognition (five items). An ordinal rating system, ranging from 1 for total assistance to 7 for complete independence, is used to rate performance.

Procedures

Tests were administered at baseline and at 6-months follow-up by two trained raters (i.e. certified physical therapists). The raters were trained to administer these outcome measures by careful review of written instructions and repeated practice. A senior certified physical therapist assessed rater competence.

Statistical analysis

Descriptive statistics were analysed for baseline characteristics using the Statistical Package for Social Sciences, version 18.0 (SPSS, Inc., Chicago, IL, USA). Continuous data were summarized by the mean (SD) and categorical data by number (percentages).

Construct validity

The construct validity of the SAROMM was assessed by calculating the Pearson's coefficients for the correlation between the SAROMM and construct measures. Construct validity was estimated by correlating the SAROMM with the GMFM-66 and WeeFIM at baseline and follow-up

respectively. A correlation of less than 0.25 was considered low, 0.25 to 0.50 as fair, 0.50 to 0.75 as moderate to good, and more than 0.75 as good to excellent.¹⁴

Responsiveness

The responsiveness of the SAROMM was estimated using the paired *t*-test, which indicates statistically significant changes from baseline to follow-up, and effect size.¹⁵ The effect size is a standardized measure of change obtained by dividing the mean change between the initial and the follow-up measurements by the standard deviation of the initial measurement.¹⁵ The effect size was classified as small (0.2–0.5), moderate (0.5–0.8), or large (>0.8).¹⁵

Estimation of minimum detectable change

The MDC was derived with a confidence level of 90% as follows:

$$\text{MDC}_{90} = 1.65 \times \sqrt{2} \times \text{SEM} = 1.65 \times \text{SD} \times \sqrt{2} (1 - r)$$

where 1.65 is the z-score associated with the desired 90% confidence level, the square root of 2 reflects the variance of two measurements, the SEM is the standard error of measurement, the SD is the pooled standard deviation, and *r* is the ICC. The ICC, one type of test–retest reliability, was estimated using independent data from 12 children with CP who were assessed twice, 2 weeks apart. A two-way mixed model was utilized to estimate the ICC.¹⁶ Changes equal to MDC₉₀ or higher were interpreted as 'true' changes in score, not change fluctuations.

Estimation of minimal clinically important difference

The anchor- and distribution-based approaches were applied to determine the MCID. The anchor-based MCID was calculated as the mean change in SAROMM score, corresponding to children who were defined as exhibiting an MCID (the MCID group), that is, those with a change in WeeFIM-MO score of 1.58 to 7.00 points (4.5–20%). Although no consensus exists in the defined range of change score for the MCID group, studies of paediatric assessment tools have suggested the following: an MCID of 4.3% on the Childhood Health Assessment Questionnaire,¹⁷ of 4.3% to 4.5% on the Pediatric Quality of Life Inventory,¹⁸ of 10% to 11% on the Pediatric Evaluation of Disability Inventory,¹⁹ and of 8% to 19% on the Pediatric Motor Activity Log.²⁰ Therefore, participants in whom a 4.55% to 20% improvement in the WeeFIM-MO between baseline and follow-up was recorded were enrolled in the MCID group.

The distribution-based MCID was estimated using the Cohen effect size benchmark. An effect size of 0.5 (i.e. 0.5 SD of the baseline score) indicated an important change and was used as the MCID threshold in this study. To assess the extent of changes detected by the SAROMM at follow-up, the proportions of participants with change scores exceeding the values of the anchor-based and distribution-based clinically important difference were examined.

As the proportion of patients who exceeded the values increased, the measure's sensitivity increased.

RESULTS

Table I summarizes the participants' demographic and clinical characteristics at baseline. Fifty-six children (90%) received regular rehabilitation by means of regular physical or occupational therapy (50% once or twice per week and 40% more than two times per week) and six children (10%) received irregular rehabilitation (less than once a week). At follow-up, approximately 89%, 66%, and 71% of participants respectively exhibited no or a positive changes in the SAROMM-SA, SAROMM-ROM, and SAROMM-total.

The ICCs for test-retest and interrater reliability were 0.907 and 0.870 respectively, for SAROMM, and 0.997 and 0.987 respectively, for the GMMFM-66. The within-participant SD values for test-retest and interrater reliabilities were 2.50 and 3.44 respectively, for SAROMM, and 1.40 and 2.97 respectively, for the GMMFM-66.

Construct validity

The correlations between all SAROMM scores and GMFM-66 scores were moderate to good ($r=-0.67$ to

-0.81 ; $p<0.01$; Table II) and were fair to good between all SAROMM scores and WeeFIM scores ($r=-0.39$ to -0.73 ; $p<0.01$) at baseline and at follow-up.

Responsiveness

For the SAROMM-total and SAROMM-ROM from baseline to follow-up changes were significant ($t=-7.83$ to -8.34 ; $p<0.001$), and the responsiveness was low (effect size=0.41 to 0.48). The changes in the SAROMM-SA from baseline to follow-up were significant ($t=-0.66$; $p=0.001$), and the SAROMM-SA responsive was low (effect size=0.24).

Estimation of minimal detectable change

The MDC90 values were 1.43 for the SAROMM-SA, 3.12 for the SAROMM-ROM, and 3.22 for the SAROMM-total. Approximately 18% to 34% of respondents exhibited a positive change that exceeded the MDC90 of the SAROMM-SA, SAROMM-ROM, and SAROMM-total (Table III).

Estimation of minimal clinically important difference

The anchor-based MCID equated to a change of 0.47, 4.07, and 4.53 in the SAROMM-SA, SAROMM-ROM, and SAROMM-total respectively (Table III). The distribution-based MCID estimates of the SAROMM-SA, SAROMM-ROM, and SAROMM-total were 1.67, 3.68, and 4.62 respectively. Analytical results indicate that 19% to 23% and 18% to 34% of participants had positive changes that exceeded the anchor- and distribution-based MCIDs of the SAROMM respectively.

DISCUSSION

To the best of our knowledge, this is the first study to investigate the responsiveness and the clinimetric properties of the SAROMM for young children with CP. The SAROMM had fair to good construct validity and was responsive to change from baseline to follow-up. Although the construct validity of the SAROMM has been reported, this study used both cross-sectional (baseline) and longitudinal approaches (follow-up) to further verify its construct validity. The MDC and MCID identified in this study may help clinicians and researchers determine whether a change in the SAROMM score indicates a true or clinically

Table II: Construct validity of the Spinal Alignment and Range of Motion Measure (SAROMM)

SAROMM	GMFM-66	WeeFIM			
		Self-care	Mobility	Cognition	Total
Baseline					
Spinal alignment	-0.64	-0.42	-0.52	-0.39	-0.47
Range of motion	-0.73	-0.56	-0.62	-0.45	-0.58
Total	-0.77	-0.56	-0.64	-0.47	-0.59
Follow-up					
Spinal alignment	-0.67	-0.42	-0.56	-0.48	-0.51
Range of motion	-0.80	-0.63	-0.73	-0.54	-0.67
Total	-0.81	-0.61	-0.73	-0.55	-0.67

All p -values for correlation coefficients were <0.001 . GMFM-66, 66-item Gross Motor Function Measure; WeeFIM, Functional Independence Measure for Children.

Table III: The minimal detectable change and minimal clinically important difference estimates of the Spinal Alignment and Range of Motion Measure (SAROMM)

SAROMM scale	Effect size	MCID estimates					
		MDC90		Anchor-based		Distribution-based	
		Score	(%) ^a	Score	(%) ^b	Score	(%) ^b
Spinal alignment	0.24	1.43	17.7	0.47	19.4	1.67	17.7
Range of motion	0.48	3.12	33.9	4.07	19.4	3.68	33.9
Total	0.41	3.22	33.9	4.53	22.6	4.62	22.6

^aProportion of participants who exceeded the criteria of the minimal detectable change at 90% confidence level (MDC90). ^bProportion of participants who exceeded the criteria of the minimal clinically important difference (MCID).

meaningful effect at posttreatment and at follow-up. When compared with measures of all joints by a goniometer, the SAROMM is a simple, timely, efficient, and valid method to estimate overall spinal alignment and ROM in children with CP. This method is especially suited for use in those who are uncooperative and in young children. Analytical results provide a benchmark for clinical decision-making while managing musculoskeletal difficulties in children with CP.

The study results demonstrate that the SAROMM has fair to good construct validity for children with CP. All SAROMM scores were moderately to well correlated with gross motor function (measured by the GMFM-66) and ADL (measured by the WeeFIM) at baseline and follow-up, which is compatible with the findings of previous studies.^{4,5} Furthermore, all SAROMM scores were more strongly correlated with motor function than ADL scores. The difference in correlations may be related to the SAROMM's design, which includes items that are related to the acquisition or maintenance of basic motor abilities, such as rolling, crawling, sitting, pulling-to-stand, transferring, and walking.⁴ A previous study found that the SAROMM has sufficient content and construct validity,⁴ which is supported by the fact that SAROMM scores were strongly correlated with GMFCS levels.⁴ The GMFM-66 and WeeFIM total scales and three subscales were used as the measures of construct validity because of significant correlations among motor, speech, and social functions in children with CP.²¹ Furthermore, gross motor ability is strongly correlated with the SAROMM scores in adolescents with CP.⁴ These data demonstrate that the SAROMM is a valid method to measure spinal alignment and ROM, and is predictive of future motor function and ADL in a CP population.

The MDC of the SAROMM determined by this study allows researchers and clinicians to determine whether a true change or treatment effect exists. In this study, the MDC90 of the SAROMM-SA, SAROMM-ROM, and SAROMM-total was 1.43, 3.12, and 3.22 points respectively. A previous study reported that the MDC95 of the SAROMM-SA, SAROMM-ROM, and SAROMM-total in children and adolescents with CP was 3, 9, and 9 points respectively, for the test-retest reliability (within 2wks).⁴ The differences in MDC values among studies may be related to participant age (1–6y vs 2–18y), confidence levels (90% vs 95%), and retest time (6mo vs 2wks). Therefore, changes can be interpreted as true and significant for children with CP when an improvement in score of 1.43 points or more on the SAROMM-SA scale, of 3.12 points or more on the SAROMM-ROM scale, and of 3.22 points or more on the SAROMM-total scale is noted.

In this study, the MCID ranges of the SAROMM-SA, SAROMM-ROM, and SAROMM-total were 0.47 to 1.67, 3.68 to 4.07, and 4.53 to 4.62 respectively. This study combined anchor- and distribution-based approaches to define MCID ranges for the SAROMM because both methods have limitations. The MCID values indicate

clinically meaningful differences resulting from a change in clinical course over time or from therapeutic effects.

The intervention efficacy is interpreted as both statistically significant and clinically important when a mean change in a group is in the range of 0.47 to 1.67 on the SAROMM-SP, 3.68 to 4.07 on the SAROMM-ROM, and 4.53 to 4.62 on the SAROMM-total. An individual is likely to experience benefit from a treatment programme when a treatment yields at least 1.67 points of improvement on the SAROMM-SP, 4.07 points on the SAROMM-ROM, and 4.62 points on the SAROMM-total. The MCID data for the SAROMM allow clinicians to determine whether treatment effects are clinically important.

The SAROMM-total and SAROMM-ROM were responsive to change at follow-up. However, the SAROMM-SA was less responsive than the SAROMM-ROM subscale. A low effect size value might be due to the participants' age or motor function severity or to short follow-up duration. An increase in age among individuals in GMFCS levels IV and V was associated with higher SAROMM scores, suggesting that impairment progresses in adolescence whose motor function is at these levels.⁵ In this study, nearly 50% of children in the sample were classified as GMFCS level I or II. These children may be less likely to be responsive to the SAROMM at a young age. The differential responsiveness of the SAROMM subscales may be due to greater differences in ROM in the extremities than in spinal alignment in children with CP, especially those with motor function classified as GMFCS level I or II. In addition, the fact that there are fewer test items in the SAROMM-SA subscale than in the SAROMM-ROM subscale may play a role. In this study, different approaches were used to quantify the responsiveness of the SAROMM at both group and individual levels. The average effects across a group may not be meaningful to individual children.²² The effect size is used to define group-level statistics and the MDC and MCID are used to define individual-level change. Therefore, low effect size values may suggest that longer periods of intervention or follow-up are needed to observe a substantial change in SAROMM in young children with CP. Furthermore, the SAROMM-ROM subscale is more sensitive and responsive to changes than is the SAROMM-SA subscale, and thus it may be used to monitor a child's status over time.

Limitations

The study's design has several limitations, including participant characteristics and outcome measures. Only preschool children were included; older children were excluded. Thus, the study results cannot be generalized to all children with CP. The SAROMM is not intended to provide a detailed ROM measure or an evaluation of spasticity. Goniometry and the Tardieu Scale are needed to measure ROM in selected body regions. Furthermore, outcomes in terms of participation and health-related quality of life were not measured in this study. In this study, the SAROMM was not translated into a Chinese version since

both therapists (raters) are proficient in English. Terminology such as ROM is regularly used in our daily practice. However, a Chinese version of the SARROM may be needed to provide better precision in raters less proficient in English. Future studies may include participation, such as play and school activities, and quality of life as outcome measures.

CONCLUSION

The clinimetric properties of the SAROMM allow clinicians and researchers to determine whether a change score indicates a true or clinically meaningful effect posttreatment and at follow-up. Furthermore, the SAROMM measures are responsive, valid, simple, and efficient to

administer. Therefore, the SAROMM is useful for examining spinal alignment and ROM, monitoring clinical course, detecting changes after intervention, and predicting motor function and ADL in children with CP. Experimental data will help clinicians in decision-making and when setting goals. Future studies based on a larger sample with children of varying ages are warranted to validate the findings of this study.

ACKNOWLEDGEMENTS

The authors would like to thank the National Science Council (NSC 98-2314-B-182-006-MY3) and Chang Gung Memorial Hospital (CMRPG 391671-3), Taiwan, for financial support. Ted Kroy is appreciated for his editorial assistance.

REFERENCES

1. Rosenbaum P, Paneth N, Leviton A, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol* 2007; **49**(Suppl. 109): 8–14.
2. Mutch L, Alberman E, Hagberg B, Kodama K, Perat MV. Cerebral palsy epidemiology: where are we now and where are we going? *Dev Med Child Neurol* 1992; **34**: 547–51.
3. Bartlett DJ, Hanna SE, Avery L, Stevenson RD, Galuppi B. Correlates of decline in gross motor capacity in adolescents with cerebral palsy in Gross Motor Function Classification System levels III to V: an exploratory study. *Dev Med Child Neurol* 2010; **52**: e155–60.
4. Bartlett D, Purdie B. Testing of the spinal alignment and range of motion measure: a discriminative measure of posture and flexibility for children with cerebral palsy. *Dev Med Child Neurol* 2005; **47**: 739–43.
5. Wright M, Bartlett DJ. Distribution of contractures and spinal malalignments in adolescents with cerebral palsy: observations and influences of function, gender and age. *Dev Neurorehabil* 2010; **13**: 46–52.
6. Fosang AL, Galea MP, McCoy AT, Reddihough DS, Story I. Measures of muscle and joint performance in the lower limb of children with cerebral palsy. *Dev Med Child Neurol* 2003; **45**: 664–70.
7. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; **63**: 737–45.
8. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010; **19**: 539–49.
9. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes* 2006; **4**: 54.
10. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther* 1996; **76**: 1109–23.
11. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007; **7**: 541–6.
12. Russell DJ, Rosenbaum PL, Avery L, Lane M. Gross Motor Function Measure (GMFM-66 and GMFM-88): User's Manual. Clinics in Developmental Medicine No. 159. London: Mac Keith Press, 2002.
13. Ottenbacher KJ, Taylor ET, Msall ME, et al. The stability and equivalence reliability of the functional independence measure for children (WeeFIM). *Dev Med Child Neurol* 1996; **38**: 907–16.
14. Colton T. Statistics in Medicine. Boston, MA: Little Brown and Co., 1974.
15. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
16. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–8.
17. Brunner HI, Klein-Gitelman MS, Miller MJ, et al. Minimal clinically important differences of the childhood health assessment questionnaire. *J Rheumatol* 2005; **32**: 150–61.
18. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. *Ambul Pediatr* 2003; **3**: 329–41.
19. Iyer LV, Haley SM, Watkins MP, Dumas HM. Establishing minimal clinically important differences for scores on the pediatric evaluation of disability inventory for inpatient rehabilitation. *Phys Ther* 2003; **83**: 888–98.
20. Lin KC, Chen HF, Chen CL, et al. Validity, responsiveness, minimal detectable change, and minimal clinically important change of the Pediatric Motor Activity Log in children with cerebral palsy. *Res Dev Disabil* 2012; **33**: 570–7.
21. Lee YC, Wu CY, Liaw MY, et al. Developmental profiles of preschool children with spastic diplegic and quadriplegic cerebral palsy. *Kaohsiung J Med Sci* 2010; **26**: 341–9.
22. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol* 2004; **57**: 1008–18.