Clinimetric Properties of Sitting Balance Measures for Children with Cerebral Palsy: A Systematic Review

Benjamin B. Bañas & Edward James R. Gorgon


To link to this article: http://dx.doi.org/10.3109/01942638.2014.881952

Published online: 03 Feb 2014.

Article views: 445

View related articles

View Crossmark data
Clinimetric Properties of Sitting Balance Measures for Children with Cerebral Palsy: A Systematic Review

Benjamin B. Bañas1,* & Edward James R. Gorgon2

1Alternative Learning Resource School, Quezon City, Philippines, 2Department of Physical Therapy, College of Allied Medical Professions, University of the Philippines Manila, Manila, Philippines

ABSTRACT. Assessment of sitting balance in children and youth with cerebral palsy (CP) is critical in order to design appropriate interventions to enhance activities and participation. This systematic review synthesized research evidence on the reliability, validity, responsiveness to change, and clinical utility of sitting balance measures for children and youth with CP. A two-tiered search in August 2012 using nine peer-reviewed electronic databases yielded nine articles with relevant information on seven clinical measures. Four of seven clinical measures: the Pediatric Reach Test (PRT), Sitting Assessment for Children with Neuromotor Dysfunction (SACND), Segmental Assessment of Trunk Control (SATCo), and Trunk Control Measurement Scale (TCMS), demonstrate acceptable overall applicability (at least one study supporting clinical utility, reliability, and validity) and are thus recommended for use in practice. Ongoing research on responsiveness to change, however, is warranted to support validity for outcomes measurement.

KEYWORDS. Assessment, cerebral palsy, children, outcomes measurement, reliability, sitting balance, validity

Sitting is an essential developmental position that is often delayed in children and youth with cerebral palsy (CP). Approximately one-third of children and youth with CP are not able to walk and are either in sitting or lying position throughout their lifespan (Surveillance of CP in Europe, 2002). Poor postural control in sitting in this population often manifests as inadequate postural adjustments and trunk control for balance (Bigongiari et al., 2011; Carlberg & Hadders-Algra, 2005; Heyrman et al., 2013; van der Heide & Hadders-Algra, 2005) that can impact...
negatively on everyday life activities (Lacoste et al., 2009). Postural control in sit-
ing is critical in the development of upper limb function, upright functional skills
and self-care, and cognitive, perceptual, and social skills (Berthental & von Hof-
sten, 1998; Ju et al., 2010). The importance of sitting balance is further underscored
by evidence that independent sitting in children with CP at the age of two can pre-
dict the ability to eventually walk (Wu et al., 2004).

Therapists utilize different interventions, such as adaptive seating devices and
balance activities for example, to improve sitting balance in children and youth with
CP (Harris & Roxborough, 2005). In identifying suitable interventions and quant-
tifying the impact of interventions, therapists rely on sitting balance assessments.
While sophisticated instruments like computerized force platforms allow for more
precise measurement of sitting balance, these are costly and are not commonly
available to clinicians (Bartlett & Birmingham, 2003). Clinical measures provide a
more practical means for clinicians and clinical researchers of assessing sitting bal-
ance and quantifying intervention outcomes. However, standardized assessments
of sitting balance are few in number. Scales like the Gross Motor Function Mea-
sure are designed to evaluate global motor abilities (Sæther & Jørgensen, 2011),
while widely-used functional measures like the Functional Reach Test, Berg Bal-
ance Scale, and Timed Up and Go evaluate standing balance and mobility (Gan
et al., 2008).

It is important to define the parameters of sitting balance for clinical measure-
ment. Balance has been defined as the ability to maintain the center of mass within
the limits of the base of support and is dependent on the requirements of the
task and environment (Huxham et al., 2001; Shumway-Cook & Woollacott, 2011).
Huxham et al. (2001) have proposed a useful framework for comprehensively
assessing balance, which differentiates between postural control and equilibrium
control; identifies different balance control mechanisms (proactive, predictive, and
reactive); and emphasizes task constraints and environmental context in balance
control. This framework aligns well with several studies that indicate postural
control and balance in children and youth with typical development and children
with CP are shaped by elements of the task and environment (Ju et al., 2012; Reilly
et al., 2008; Streepey & Angulo-Kinzler, 2002).

A clinical measure must be valid, reliable, responsive to change, and practical
to administer in order for it to be of value to clinicians (Kimberlin & Winterstein,
2008; Smart, 2006). Such measurement properties are population-specific (Portney
& Watkins, 2009). Reliability relates to the extent to which clinical measurement is
error-free and includes intra-rater, inter-rater, test–retest reliability, internal con-
sistency, and measurement error. Test–retest reliability reflects the stability of a
measurement instrument with repeated administrations, often separated by a time
interval of sufficient length and which may or may not involve a rater (Kimberlin &
Winterstein, 2008; Portney & Watkins, 2009). Intra–rater reliability pertains to the
degree of agreement among data taken by the same rater repeatedly and separated
by brief time intervals (Portney & Watkins, 2009). Validity is the extent to which
a measure addresses what it is intended to measure and includes face, construct,
content, and criterion-related validity (Kimberlin & Winterstein, 2008; Portney &
Watkins, 2009). Responsiveness to change refers to a measure’s ability to detect
change over time (Kimberlin & Winterstein, 2008). Clinical utility pertains to the
usefulness of a measure in clinical practice that includes ease of use, time to complete assessment, examiner training and qualifications, format, and interpretation (Smart, 2006).

Clinical measures of sitting balance whose reliability and validity are supported by research evidence increase the likelihood that clinicians can accurately identify sitting balance problems, describe change in sitting balance over time, and estimate the impact of sitting balance interventions. This highlights the need to comprehensively examine the measurement properties of sitting balance measures for children and youth with CP. This systematic review sought to realize two aims: (1) identify published clinical measures that quantify sitting balance in children and youth with CP and (2) evaluate the overall applicability of the clinical measures in children and youth with CP based on evidence of reliability, validity, responsiveness to change, and clinical utility.

**METHOD**

Data Sources and Searches

A two-phased search was completed using nine electronic databases: PubMed (1966–August 2012); MEDLINE (1950–August 2012); Embase (1974–August 2012); Cumulative Index to Nursing and Allied Health Literature (CINAHL) (1981–August 2012); Web of Science (1900–August 2012); Allied and Complementary Medicine Database (AMED) (1985–August 2012); Science Direct (1995–August 2012); Physiotherapy Evidence Database (PEDro) (1929–August 2012); and Occupational Therapy Systematic Evaluation of Evidence (OTseeker) (1950–August 2012). The first phase of the search was directed toward identifying clinical measures of sitting balance used in research involving children and youth with CP (see Figure 1). Databases that indexed clinical trials mostly, namely PEDro and OTseeker, were utilized in this phase only. Key search terms applied were “balance” or “postural control” or “trunk control” or “postural stability”; “assessment” or “test”; and “sitting” and “cerebral palsy” and “child”, as well as synonyms and relevant medical subject headings (MeSH) terms. Appropriate Boolean terms and symbols were used to narrow the search. Ancestral searching of literature reviews and clinical trials that included sitting balance in children and youth with CP as an outcome was also done to ensure that no relevant clinical measure was missed.

The second phase of the search was aimed at finding research on the clinimetric properties (i.e., validity, reliability, responsiveness or sensitivity to change, and clinical utility) of each clinical measure that had been identified in the first phase of the search (Figure 1). The names of the included clinical measures and specific terms for the clinimetric properties were combined with “child” and “cerebral palsy” as search terms. Common alternative terms, abbreviations, and acronyms for the clinical measures were also incorporated in the search. The reference lists of all included studies were searched to find any additional studies on clinimetric properties.

Study Selection

Included studies (1) were peer-reviewed and published full reports on validity, reliability, responsiveness to change, or clinical utility of a clinical measure of sitting
316  

Bañas and Gorgon

Figure 1. Search strategy for locating clinical measures that assess sitting balance in children with cerebral palsy and supporting research evidence on clinimetric properties.

Phase one

1048 titles and abstracts were identified in the search of 9 databases.

62 fulltext articles on 18 potential clinical measures of sitting balance for children with cerebral palsy were identified.

A total of 10 clinical measures of sitting balance in children with cerebral palsy were identified for phase two of the search.

42 fulltext articles were reviewed.

967 excluded due to duplication or not being relevant to sitting balance in children with cerebral palsy.

18 excluded due to use of only laboratory measures of sitting balance.

1 excluded due to not being reported in the English language.

Fulltext articles were reviewed.

11 clinical measures excluded due to being larger assessments of motor ability, balance, or mobility (<50% of items relevant to sitting balance).

3 potential clinical measures added from ancestral searching.

Phase two

14 fulltext articles on 9 of the 10 clinical measures were identified in the search of 7 databases that incorporated the names of the clinical measures and clinimetric properties.

Fulltext articles were reviewed.

1 clinical measure excluded due to having no study to support clinimetric properties.

4 studies focusing on 2 clinical measures excluded due to not having data on clinimetric properties.

1 study excluded due to inclusion of adults in the sample of children with cerebral palsy.

9 fulltext articles supporting 7 clinical measures of sitting balance were included for critical appraisal and data extraction.

balance; (2) had children and youth with CP (aged 18 years and lower) as participants or part of the sample regardless of severity of disability or Gross Motor Function Classification Scale level (GMFCS; Palisano et al., 1997); and (3) were written in the English language. Balance was defined based on the theoretical framework by Huxham et al. (2001) and covered both postural control (maintenance of the position against the pull of gravity and any movement that may change the center of mass relative to the base of support) and equilibrium control (maintenance of intersegmental stability of the body and its parts despite the forces acting on it). Excluded were studies on clinical measures that included sitting balance as part of a larger assessment of motor ability, balance, or mobility (operationally defined as having less than 50% of items relevant to sitting balance). As well, studies were excluded if the sample was a mixture of adults (individuals older than 18 years) and
Clinimetric Review of Sitting Balance Measures

Data Extraction and Quality Assessment

Data were extracted independently by the first author using the CanChild Outcome Measures Rating Form (Law, 2004) and verified independently by the second author. The CanChild instrument allows for an assessment of the overall applicability of a clinical measure based on its reliability, validity, and clinical utility. Reliability was considered excellent if the coefficient was $\geq 0.80$; adequate if it ranged 0.60–0.79; and poor if it was $< 0.60$ (Law, 2004). Ratings for validity dimensions were excellent, acceptable, and poor, depending on the comprehensiveness and relevance of items or extent of support for the property from available research evidence.

Additional standards supplemented the CanChild Outcome Measures Rating criteria in the interpretation of validity estimates. Correlation coefficients reported for construct validity were considered very high $= 0.91–1.0$, high $= 0.71–0.9$, moderate $= 0.51–0.7$, low $= 0.31–0.5$, and little $= 0.0–0.3$ (Hinkle et al., 1998). Criterion validity was evaluated based on the Pearson’s correlation coefficient: $>0.7 = $ high, $0.5–0.69 = $ moderate, $0.26–0.49 = $ low, and $\leq 0.25 = $ little (Munro, 1997). Overall applicability was rated as excellent (adequate to excellent clinical utility, easily available, and excellent reliability and validity), adequate (adequate to excellent clinical utility, easily available, and adequate to excellent reliability and validity), and poor (poor clinical utility, not easily available, and poor reliability and validity) (Law, 2004).

Methodological quality of individual studies was independently rated by the authors on relevant subscales of the dichotomous scale version of the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN). The COSMIN, developed through an international Delphi study and found to have high inter-rater agreement, can be utilized to evaluate the methodological quality of studies on measurement properties for inclusion in systematic reviews (Mokkink et al., 2010a). Each item on the dichotomous scale requires a “yes” or “no” answer, with a “not applicable” option also available. Using the principle of “worst score counts,” a measurement property is considered poor if a negative answer is given to any of the COSMIN items deemed to be an indicator of major methodological flaw (Terwee et al., 2012). Disagreements in data extraction and quality assessment were discussed by the authors and settled through consensus.

RESULTS

Search Results and Included Studies

In all, seven clinical measures of sitting balance in children and youth with CP and nine studies on the clinimetric properties of the measures were found (Figure 1):
TABLE 1. List of Excluded Clinical Measures of Sitting Balance in Children and Youth with Cerebral Palsy

<table>
<thead>
<tr>
<th>Clinical Measure</th>
<th>Reason for Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alberta Infant Motor Scale (Darrah et al., 1998)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Berg Balance Scale (Kembhavi et al., 2002)</td>
<td>&lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Chailey Levels of Ability (Pountney et al., 1999)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Early Clinical Assessment of Balance for Young Children with Cerebral Palsy</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>(McCoy et al., 2010)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Gross Motor Function Classification Scale (Wood &amp; Rosenbaum, 2000)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Gross Motor Function Measure (Bjornson et al., 1998)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Gross Motor Performance Measure (Boyce et al., 1992; Boyce et al., 1995)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Level of Sitting Ability Scale (Green &amp; Nelham, 1991)</td>
<td>No supporting study on clinimetric properties.</td>
</tr>
<tr>
<td>Modified Schober Measurement of Spinal Extension (Macrae &amp; Wright, 1969)</td>
<td>No supporting study on clinimetric properties.</td>
</tr>
<tr>
<td>Motor Assessment of Infants (Haley et al., 1986)</td>
<td>Broad assessment; &lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Modified Posture Assessment Scale / Posture Assessment Scale (Jonsdottir et al.,</td>
<td>No supporting study on clinimetric properties.</td>
</tr>
<tr>
<td>1997; Bertoti, 1988)</td>
<td></td>
</tr>
<tr>
<td>Pediatric Balance Scale (Franjoine et al., 2003)</td>
<td>&lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>Sitting Assessment Scale (Myhr &amp; von Wendt, 1991; Myhr et al., 1993; Myhr et al.,</td>
<td>&lt;50% of items relevant to sitting balance.</td>
</tr>
<tr>
<td>1995)</td>
<td></td>
</tr>
<tr>
<td>Spinal Alignment and Range of Motion Measure (Bartlett &amp; Purdye, 2005)</td>
<td>&lt;50% of items relevant to sitting balance.</td>
</tr>
</tbody>
</table>

Level of Sitting Scale (LSS), Pediatric Reach Test (PRT), Seated Postural Control Measure (SPCM), Segmental Assessment of Trunk Control (SATCo), Sitting Assessment for Children with Neuromotor Dysfunction (SACND), Trunk Control Measurement Scale (TCMS), and Trunk Impairment Scale (TIS). Eleven measures were excluded because those either were broad assessments of motor ability, balance or mobility, or did not have supporting clinimetric studies (Table 1).

Descriptive data on all included studies are summarized in Table 2. Samples from which data on clinimetrics were derived comprised children and youth with CP mostly of the spastic type (PRT, TCMS, and TIS) or predominantly children and youth with CP mixed with children and youth with other neuromotor or developmental disorders (LSS, SPCM, SATCo, and SACND). Sample size varied from 3 to 114. The mean age varied from 3.8–11.3 years and the studies included children in GMFCS levels I–IV.

**Reliability**

Findings on the reliability of the clinical measures are summarized in Table 3. At least one reliability study (inter-rater) supported each of the measures. Excellent inter-rater reliability estimates were reported for the PRT, SATCo, TCMS, and TIS. Poor to excellent inter-rater reliability estimates were reported for the SPCM and SACND. Poor to adequate inter-rater reliability was reported for the LSS.
<table>
<thead>
<tr>
<th>Clinical Measure</th>
<th>Description</th>
<th>Study</th>
<th>Participants</th>
<th>Clinimetric Property Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Sitting Scale</td>
<td>Sitting classification based on amount of support required to maintain position</td>
<td>(Fife et al., 1991)</td>
<td>40 non-ambulatory children (19 with spastic CP), mean age: 9.06 yr, age range: 1.67–18.5 yr</td>
<td>Inter-rater and test–retest reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Field &amp; Roxborough, 2011)</td>
<td>114 children with ND (10 of 63 in “change” group with CP, 51 of 51 in “stable” group with CP), mean age: 10.8 yr, age range: 1–18 yr</td>
<td>Responsiveness to change</td>
</tr>
<tr>
<td>Pediatric Reach Test</td>
<td>Modification of Functional Reach Test incorporating sitting position and side reaching</td>
<td>(Fife et al., 1991)</td>
<td>114 children with ND (61 with CP), mean age: 10.8 yr, age range: 1–18 yr</td>
<td>Construct validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Field &amp; Roxborough, 2012)</td>
<td>114 children with ND (61 with CP), mean age: 10.8 yr, age range: 1–18 yr</td>
<td>Construct validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Bartlett &amp; Birmingham, 2003)</td>
<td>10 children with CP (9 spastic, 1 hypotonic; GMFCS I, III, IV), mean age: 8.2 yr, age range: 2.6–14.1 yr</td>
<td>Inter-rater and test–retest reliability</td>
</tr>
<tr>
<td>Seated Postural Control</td>
<td>Assessment of sitting for assistive seating device prescription and outcome evaluation in terms of (1) static postural alignment and (2) functional movement</td>
<td>(Fife et al., 1991)</td>
<td>40 non-ambulatory children (19 with spastic CP), mean age: 9.06 yr, age range: 1.67–18.5 yr</td>
<td>Intra-rater and inter-rater reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Field &amp; Roxborough, 2011)</td>
<td>114 children with ND (10 of 63 in “change” group with CP, 51 of 51 in “stable” group with CP), mean age: 10.8 yr, age range: 1–18 yr</td>
<td>Responsiveness to change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Field &amp; Roxborough, 2012)</td>
<td>114 children with ND (61 with CP), mean age: 10.8 yr, age range: 1–18 yr</td>
<td>Construct validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Butler et al., 2010)</td>
<td>24 children with ND (21 with CP: 15 spastic, 6 other types; GMFCS I–V), mean age: 10.3 yr, age range: 1.5–17.1 yr</td>
<td>Concurrent validity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Continued on next page)</td>
</tr>
</tbody>
</table>
TABLE 2. Summary of Studies Supporting Clinical Measures of Sitting Balance in Children and Youth with Cerebral Palsy (Continued)

<table>
<thead>
<tr>
<th>Clinical Measure</th>
<th>Description</th>
<th>Study</th>
<th>Participants</th>
<th>Clinimetric Property Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting Assessment of Children with Neuromotor Dysfunction</td>
<td>Assessment of static and dynamic trunk control</td>
<td>(Reid, 1995)</td>
<td>Reliability: 3 children (2 with spastic CP, 1 with motor delay), age: 10 yr (study 1); 8 children with spastic CP, mean age: 3.8 yr, age range: 2.7–4.7 yr (study 2)</td>
<td>Inter-rater and test–retest reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Construct validity: 6 children with spastic CP, mean age: 6 yr, age range: 4–8 yr (study 1); 8 children with spastic CP, mean age: 3.8 yr, age range: 2.7–4.7 yr (study 2)</td>
<td>Internal consistency Content validity Construct validity</td>
</tr>
<tr>
<td>Trunk Control Measurement Scale</td>
<td>Assessment of static and dynamic trunk control</td>
<td>(Heyrman et al., 2011)</td>
<td>26 children with spastic CP (GMFCS I–III), mean age: 11.3 yr, age range: 8.3–15.3 yr</td>
<td>Inter-rater and test–retest reliability</td>
</tr>
<tr>
<td>Trunk Impairment Scale</td>
<td>Assessment of static and dynamic trunk control</td>
<td>(Sæther and Jørgensen, 2011)</td>
<td>20 children with CP (18 spastic, 2 dyskinetic; GMFCS I–IV), mean age range across motor levels: 8.0–9.8 yr, age range: 5–12 yr</td>
<td>Inter-rater and test–retest reliability Measurement error</td>
</tr>
</tbody>
</table>

Note. CP = cerebral palsy; ND = neuromotor disorder; yr = years.
Test–retest reliability estimates were published for six measures: excellent for the SACND, TCMS, and TIS; poor to excellent for the PRT; and poor for the LSS and SPCM. The SATCo demonstrated excellent intra–rater reliability. Evidence of internal consistency of items was reported for the TCMS and SACND. Measurement error has been reported for the TCMS and TIS.

**Validity and Responsiveness to Change**

Validity and responsiveness to change of the clinical measures are detailed in Table 4. Six of seven measures were supported by at least one study on validity, but no study investigated validity of the TIS in children and youth with CP. Content validity was investigated for the SPCM and SACND only using expert evaluations on the comprehensiveness of items (excellent on the CanChild Outcome Measures Rating criteria). Criterion validity was supported for the SATCo only with high correlations with scores on the sitting subscales of more established motor function measures (Gross Motor Function Measure [GMFM] and Alberta Infant Motor Scale) (adequate on the CanChild Outcome Measures Rating criteria). One study on construct validity was each found for the LSS, PRT, SACND, and TCMS. Construct validity of three of the measures was examined by correlation analysis. The magnitude of the correlation between scores on the TCMS and GMFM was high (adequate on the CanChild Outcome Measures Rating criteria), moderate between scores on the PRT and GMFCS (adequate on the CanChild Outcome Measures Rating criteria), and low between scores for the LSS and external postural support in sitting (poor on the CanChild Outcome Measures Rating criteria). One study reported responsiveness for the LSS and another study reported responsiveness of the SPCM (Function subscale). Correlations with Global Change Scale (GCS) were weak and fair-to-moderate (poor on the CanChild Outcome Measures Rating criteria).

**Clinical Utility**

Data on key dimensions of clinical utility are described in Table 5. Clarity of instructions was excellent for all the measures, defined on the CanChild Outcome Measures Rating criteria as being clear, comprehensive, concise, and available. Most utilized ordinal scales, with only the PRT and SATCo involving an interval measure and nominal scale respectively. All measures required active physical participation from the client through performance of sitting or sitting in combination with specific movements or tasks. Time for completion of the procedures ranged from 5–10 to 20 min, although this was unclear for the SATCo, TCMS, and TIS. Formal training for assessors was not explicitly required for any of the measures but all the studies estimated measurement properties employing trained assessors. Cost was not reported in any of the studies, although procedures and instruments were available in the research reports. The LSS, PRT, TCMS, and TIS did not involve equipment and materials apart from those that could be accessed easily in most clinical settings, while the SPCM, SACND, and SATCo identified specific materials with specific dimensions (e.g., small toys attached to a 17-inch by 22-inch plexiglass board for the SACND and a strapping system for the SATCo).
<table>
<thead>
<tr>
<th>Clinical Measure</th>
<th>Intra-rater</th>
<th>Inter-rater</th>
<th>Test–retest</th>
<th>Internal Consistency</th>
<th>Measurement Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level of Sitting Scale</strong></td>
<td>Not reported.</td>
<td>$k = 0.58–0.62%$ agreement = 68.4–68.6 (Fife et al., 1991).</td>
<td>$k = 0.54–0.55%$ agreement = 63.6–64.7 (Fife et al., 1991).</td>
<td>Not reported.</td>
<td>Not reported.</td>
</tr>
<tr>
<td><strong>Pediatric Reach Test</strong></td>
<td>Not reported.</td>
<td>Sitting section ICC = 0.84</td>
<td>Combined Standing and Sitting sections ICC = 0.54–0.88 (Bartlett &amp; Birmingham, 2003).</td>
<td>Not reported.</td>
<td>Not reported.</td>
</tr>
<tr>
<td><strong>Seated Postural Control Measure</strong></td>
<td>Not reported.</td>
<td>Without seating system Alignment subscale $k =$ 0.41–0.47,% agreement = 67.1–70.0</td>
<td>Without seating system Alignment subscale $k =$ 0.36–0.41,% agreement = 63.4–67.0</td>
<td>Not reported.</td>
<td>Not reported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Function subscale $k =$ 0.79–0.87,% agreement = 88.4–93.4 (Fife et al., 1991).</td>
<td>Function subscale $k =$ 0.28–0.29,% agreement = 49.0–49.9 (Fife et al., 1991).</td>
<td>Not reported.</td>
<td>Not reported.</td>
</tr>
<tr>
<td><strong>Segmental Assessment of Trunk Control</strong></td>
<td>Across aspects of control, ICC = 0.98</td>
<td>Static control ICC = 0.8</td>
<td>Not reported.</td>
<td>Not reported.</td>
<td>Not reported.</td>
</tr>
<tr>
<td></td>
<td>(Butler et al., 2010).</td>
<td>Active control ICC = 0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reactive control ICC = 0.8 (Butler et al., 2010).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sitting Assessment for Children with Neuromotor Dysfunction</strong></td>
<td>Not reported.</td>
<td>Rest and Reach modules% agreement = 67–92</td>
<td>Rest and Reach modules ICC &gt; 0.99</td>
<td>Rest module Cronbach $\alpha =$ 0.47–0.49</td>
<td>Not reported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rest module subscales $k =$ 0.35–1.0,% agreement = 72–100</td>
<td>Rest module subscales $k =$ 0.87–1.0</td>
<td>Reach module Cronbach $\alpha =$ 0.71–0.78 (Reid, 1995).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reach module subscales $k =$ 0.55–1.0,% agreement = 72–94 (Reid, 1995).</td>
<td>Reach module subscales $k =$ 0.91–1.0 (Reid et al., 1996).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rest and Reach modules ICC &gt; 0.99</td>
<td>Rest module</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rest module subscales $k =$ 0.91–1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reach module subscales $k =$ 0.96–1.0 (Reid et al., 1996).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trunk Control Measurement Scale</td>
<td>Not reported.</td>
<td>ICC = 0.98 Static sitting balance subscale (ICC = 0.98)</td>
<td>ICC = 0.97 Selective movement control subscale (ICC = 0.94)</td>
<td>ICC = 0.97 Dynamic reaching subscale (ICC = 0.99) (Heyrman et al., 2011).</td>
<td>Cronbach $\alpha = 0.94$ Static sitting balance subscale ($\alpha = 0.82$) Selective movement control subscale ($\alpha = 0.89$) Dynamic reaching subscale ($\alpha = 0.9$) (Heyrman et al., 2011).</td>
</tr>
<tr>
<td>Trunk Impairment Scale</td>
<td>Not reported.</td>
<td>ICC = 0.97–1.0 Static sitting balance subscale items ($k = 0.78–1.0$, % agreement = 88–100) Dynamic reaching subscale (ICC = 0.97–0.99) and subscale items ($k = 0.66–1.0$) Coordination subscale items ($k = 0.6–0.86$, % agreement = 88) (Sæther &amp; Jørgensen, 2011).</td>
<td>ICC = 0.97–0.99 Static sitting balance subscale items ($k = 0.47–1.0$, % agreement = 88–100) Dynamic reaching subscale (ICC = 0.94–0.99) and subscale items ($k = 0.57–1.0$) Coordination subscale items ($k = 0.7–1.0$) (Sæther &amp; Jørgensen, 2011).</td>
<td>Not reported.</td>
<td>SEM = 0.94–1.7 (inter-rater) and 0.73–1.70 (test-retest) (Sæther &amp; Jørgensen, 2011).</td>
</tr>
<tr>
<td>Clinical Measure</td>
<td>Content</td>
<td>Construct</td>
<td>Criterion</td>
<td>Responsiveness</td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Level of Sitting Scale</td>
<td>Not reported.</td>
<td>Negative correlation between LSS scores and amount of external postural support used in sitting (Spearman's ( \rho = -0.42 )) (Field &amp; Roxborough, 2012).</td>
<td>Not reported.</td>
<td>Positive correlations between LSS change scores and GCS by parents and therapists (Pearson's ( r = 0.16-0.31 )) (Field &amp; Roxborough, 2011).</td>
<td></td>
</tr>
<tr>
<td>Pediatric Reach Test</td>
<td>Not reported.</td>
<td>Negative correlation between PRT sitting subscale and GMFCS (Spearman's ( \rho = -0.69 )) (Bartlett &amp; Birmingham, 2003).</td>
<td>Not reported.</td>
<td>Not reported.</td>
<td></td>
</tr>
<tr>
<td>Seated Postural Control Measure</td>
<td>Items generated through consultations with seating-experienced therapists and evaluated by 7 external seating experts for content validity and clinical utility (Fife et al., 1991).</td>
<td>Not reported.</td>
<td>Not reported.</td>
<td>No significant correlation between SPCM Alignment subscale and GCS scores. Positive correlations between SPCM Function subscale and GCS scores (Pearson's ( r = 0.26-0.28 )) (Field &amp; Roxborough, 2011).</td>
<td></td>
</tr>
<tr>
<td>Segmental Assessment of Trunk Control</td>
<td>Not reported.</td>
<td>Not reported.</td>
<td>Positive correlation between SATCo and GMFM Dimension B (sitting) (Pearson's ( r ) for static control = 0.83, active control = 0.77, reactive control = 0.73)</td>
<td>Positive correlation between SATCo and AIMS (sitting) (Pearson's ( r ) for static control = 0.88, active control = 0.86, reactive control = 0.87) (Butler et al., 2010).</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>Description</td>
<td>Results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sitting Assessment for Children with Neuromotor Dysfunction</td>
<td>Items and scoring developed from 4 key constructs derived from literature review. Agreement of 13 pediatric occupational therapists on importance of concept = 67%, clarity of wording 70-73%, and reflection of operational definition on scoring = 66-69% (Reid, 1995).</td>
<td>No difference between Rest and Reach subscales on 14 of 32 items failing to support projected difference between static and dynamic sitting postural control (Reid, 1995). Better seated postural control (lower SACND Rest subscale scores) in children given saddle seat versus flat bench $p &lt; 0.01$ (Reid, 1995). Discriminative between children with and without CP $p &lt; 0.01$ (mean Rest subscale scores = 12.7 and 6.9, respectively, $p = 0.02$, and mean Reach subscale scores = 13.8 and 6.0, respectively) (Reid, 1995).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trunk Control Measurement Scale</td>
<td>Not reported.</td>
<td>Positive correlation between TCMS and GMFM total scores (Spearman’s $\rho = 0.88$) (Heyrman et al., 2011)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trunk Impairment Scale</td>
<td>Not reported.</td>
<td>Not reported.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: AIMS = Alberta Infant Motor Scales; CP = cerebral palsy; GCS = Global Change Scale; GMFCS = Gross Motor Function Classification System; GMFM = Gross Motor Function Measure; LSS = Level of Sitting Scale; SACND = Sitting Assessment for Children with Neuromotor Dysfunction; SATCo = Segmental Assessment of Trunk Control; SPCM = Seated Postural Control Measure; TD = typical development; yr = years.
### TABLE 5. Clinical Utility of Sitting Balance Measures for Children and Youth with Cerebral Palsy

<table>
<thead>
<tr>
<th>Clinical Measure</th>
<th>Clarity of Instructions</th>
<th>Format of Administration</th>
<th>Administration Time</th>
<th>Assessor Qualifications</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Sitting Scale</td>
<td>Excellent.</td>
<td>Performance of sitting by client rated on 8-point ordinal scale, no special equipment required.</td>
<td>5–10 min</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
<tr>
<td>Pediatric Reach Test</td>
<td>Excellent.</td>
<td>Performance of reaching in three directions by client measured as distance reached (in centimeters), no special equipment required.</td>
<td>15 min</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
<tr>
<td>Seated Postural Control Measure</td>
<td>Excellent.</td>
<td>Performance of sitting (Alignment subscale) and various seated functional tasks (Function subscale) by client rated on 4-point ordinal scales, materials with specific dimensions required.</td>
<td>20 min</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
<tr>
<td>Segmental Assessment of Trunk Control</td>
<td>Excellent.</td>
<td>Performance of sitting and lifting the arms by client rated on nominal scale, strapping system required but video equipment optional, assistant required to apply nudge.</td>
<td>Not reported.</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
<tr>
<td>Sitting Assessment Test for Children with Neuromotor Dysfunction</td>
<td>Excellent.</td>
<td>Videotaped performance of sitting (Rest module) and sitting while pointing or reaching (Reach module) by client rated on 4-point ordinal scale, plexiglass board with toy attachments required.</td>
<td>10 min (two 5-min sessions)</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
<tr>
<td>Trunk Control Measurement Scale</td>
<td>Excellent.</td>
<td>Performance of sitting and various movements and tasks in sitting by client rated on 2-, 3-, or 4-point ordinal scales, no special equipment required.</td>
<td>Not reported.</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
<tr>
<td>Trunk Impairment Scale</td>
<td>Excellent.</td>
<td>Performance of sitting and various movements in sitting by client rated on 2-, 3-, or 4-point ordinal scales, no special equipment required.</td>
<td>10 min</td>
<td>No formal training requirement specified for average users but study used trained assessors.</td>
<td>Not reported.</td>
</tr>
</tbody>
</table>
Quality Assessment of Included Studies

All studies on reliability and validity rated favorably on most key indicators of methodological quality, but garnered poor overall ratings on the COSMIN due to a “poor” rating on at least one item which resulted in a “worst score count”. A consistent limitation in all these studies was the lack of an appropriate sample size, with the exception of the study by Fife et al. (LSS and SPCM). In addition, two reliability studies did not report subscale item kappa coefficients for ordinal (TCMS) and nominal (SATCo) measures, and relied on the ICC in estimating reliability. In two construct validity studies (PRT and TCMS), lack of an explicit a priori hypothesis about the relationship between the sitting balance measure and another measure representing the construct might have led to some bias. The study on responsiveness to change included the LSS and SPCM and satisfied key methodological quality indicators on the COSMIN.

DISCUSSION

This is the first systematic review to focus on sitting balance measures for children and youth with CP. A comprehensive literature search identified seven clinical measures of sitting balance that have published data on psychometric properties. The seven measures address postural control and equilibrium control, and predictive and proactive sitting balance mechanisms. Only the SATCo, which comprised items mostly on static sitting balance, also focused on the reactive sitting balance mechanism. This indicates that it might be of value to use the SATCo in combination with another, more dynamic sitting balance measure such as the PRT to improve the comprehensiveness of assessment of the sitting balance control mechanisms. However, all clinical measures involved closed task conditions only and assessments were carried out in simple and stable environmental conditions only. For example, the measures commonly involve maintaining steady sitting or reaching toward a stationary target or object. This finding underscores a key limitation of sitting balance measures for CP. Balance assessments should incorporate different biomechanical and information processing demands reflective of “real-world” task and environment contexts (Huxham et al., 2001). Thus, development of clinical measures that emphasize contextual factors in functions that require sitting balance is recommended. For instance, donning and doffing a shirt and catching and throwing a ball while at play in sitting represent meaningful contexts in which sitting balance can be further assessed.

Level of Sitting Scale

Standardization studies for the LSS indicate poor reliability overall, attributed to the need to improve item definitions and administration guidelines, and assessor training (Fife et al., 1991). Another study reported generally satisfactory inter-rater and test–retest reliability estimates for the LSS (Fife et al., 1993), but it was excluded due to its use of adults aged up to about 28 years in the sample. Fife et al. (1993) employed “well-trained raters,” although they did not describe the extent of assessor training. Only one study has investigated the validity of the LSS and generally poor construct validity was gleaned from the low correlation coefficient
(Field & Roxborough, 2012). The study on responsiveness to change (Field & Roxborough, 2011) was methodologically sound but showed low positive correlations between LSS and GCS change scores. The LSS appears practical to administer but lacks overall applicability due to the absence of evidence for reliability and validity.

**Pediatric Reach Test**

Overall applicability of the PRT is adequate. It demonstrated generally excellent reliability, adequate validity, and feasibility for clinical use (Bartlett & Birmingham, 2003). Responsiveness to change of the PRT, however, has not been studied. This poses uncertainty regarding its value in monitoring clinically meaningful change in sitting balance.

**Seated Postural Control Measure**

Standardization studies for reliability and responsiveness to change for the SPCM and LSS have been conducted together with parallel findings for the two measures (Fife et al., 1991; Field & Roxborough, 2011). Reliability of the SPCM was generally poor, especially for the Alignment subscale, with the same underlying reasons as identified for the LSS (Fife et al., 1991). Responsiveness to change was not supported, with weak correlations between the Function subscale and GCS scores (Field & Roxborough, 2011). Clinical utility of the SPCM was acceptable for most dimensions, except for the required feeder seat with a modified wheeled base in the original study (Fife et al., 1991). Further standardization of the items of the SPCM is warranted before it can be recommended for use in practice.

**Segmental Assessment of Trunk Control**

The SATCo possesses adequate overall applicability, with one study reporting excellent reliability and high agreement with two gold standard measures (Butler et al., 2010). In addition to addressing reactive sitting balance control, it is also one of only two measures that can be used on clients with more severe limitations in sitting balance (GMFCS levels IV-V) (Butler et al., 2010). Apart from the required strapping system, most clinical utility dimensions of the SATCo indicate its usefulness in most practice settings. As with the PRT, its responsiveness to change has not been determined.

**Sitting Assessment for Children with Neuromotor Dysfunction**

The SACND has shown adequate evidence for reliability and validity, although responsiveness to change has not been addressed. Following modifications in wording and scoring criteria illustrations, excellent inter-rater and test–retest reliability estimates were established for all items in the two modules of the SACND (Reid et al., 1996). Such modifications were justified by poor reliability achieved by the Balance item and Rest module in the original study (Reid, 1995). Administration is clinically feasible overall, though a plexiglass board with specified dimensions needs to be available.
Trunk Control Measurement Scale

Given excellent estimates in multiple reliability dimensions and very high positive correlation with the GMFM in one study (Heyrman et al., 2011), overall applicability of the TCMS is adequate. Feasibility is supported for most of the clinical utility dimensions. Administration however may be limited to children and youth with CP who can comprehend and follow instructions. As with the PRT, SATCo, and SACND, responsiveness to change has not been explored.

Trunk Impairment Scale

Excellent reliability and feasibility of administration described in one study for the TIS (Sæther & Jørgensen, 2011) suggest its potential applicability. The absence of validity research to support its use on children and youth with CP, however, diminishes overall applicability. Proper validation is necessary to allow clinicians to understand well what information clinical measures provide them; help them in selecting appropriate measures for their specific purposes; and provide them a basis for interpreting and relating measurements with patient care and clinical service toward effective decision making (Kimberlin & Winterstein, 2008; Tyson & Connell, 2009). This finding emphasizes the need for further research to establish the validity of the TIS for the intended client group.

In view of the CanChild Outcome Measures Rating criteria (Law, 2004), the PRT, SACND, SATCo, and TCMS can be considered to have adequate overall applicability and may be recommended for use in clinical practice. Before a measure can be considered useful to clinicians and researchers, clinimetric properties must be well established (Scholtes et al., 2011). Ongoing development of these sitting balance measures, however, needs to be carried out. Apart from the SACND (Reid, 1995), none of these measures are supported in more than one validity dimension. Future research should also examine the instruments’ responsiveness to clinically important change. Absence of evidence on responsiveness to change has important negative ramifications on longitudinal assessment in clinical practice and outcome research. The impact of interventions over time may not be ascertained if the clinical tool has no known responsiveness or is prone to instability, or floor or ceiling effects (Kimberlin & Winterstein, 2008; Stratford et al., 1996).

The seven measures demonstrated feasibility for clinical use. The availability of the tools, procedures, and materials list in research articles and on some proponents’ websites allows insight on the acceptability of cost for many of the measures. What is unclear though is the degree of experience and training required to reliably administer the clinical measures in practice given the insufficiency of information provided in some of the studies. Also, samples in the studies had comprised principally CP of the spastic type and motor functions at GMFCS levels I-III, and certain measures (e.g., the PRT and TCMS) appear to be feasible only for clients who are able to follow verbal instructions. This might influence the generalizability of assessment given the spectrum of postural dysfunctions and functional abilities observed in children and youth with CP (Chen et al., 2010; van der Heide & Hadders-Algra, 2005). This highlights the need to investigate the clinimetric properties of sitting balance measures in light of sub-groups within the CP population in terms of severity of motor and cognitive disability.
The COSMIN complemented the CanChild Outcome Measures Rating criteria in that it provided further basis for making an assessment of standardization for reliability, validity, and responsiveness to change. In using the COSMIN, specific gaps in the existing literature on sitting balance measures were identified. For example, many of the studies did not meet the required number of participants (Bartlett & Birmingham, 2003; Butler et al., 2010; Heyrman et al., 2011; Reid et al., 1996; Reid, 1995; Sæther & Jørgensen, 2011). Studies with an insufficient sample size would be less likely to capture the diversity of sitting balance problems in children and youth with CP. Thus, the external validity of the findings of the studies might be diminished even when favorable reliability estimates had been obtained. Some of the studies were limited by inadequate reporting of either hypotheses (Bartlett & Birmingham, 2003; Heyrman et al., 2011) or statistical results (Heyrman et al., 2011; Butler et al., 2010) which could limit the ability to determine an accurate estimate of the validity or reliability. Without a priori hypotheses in construct validity testing, risk of bias can increase when low correlations found might be justified with alternative explanations instead of questioning the validity of an instrument (Mokkink et al., 2010b). In reliability studies, categorical data (i.e., those derived from the SATCo and TCMS) are appropriately reported using kappa statistics (Mokkink et al., 2010b). Calculating intraclass correlation coefficients for such data, as was done in the studies by Heyrman et al. and Butler et al., might obscure the true reliability estimates. This underscores the need for additional research that would address such limitations in order to establish robust support for the psychometric properties of the measures. Future methodological studies may find it useful to be guided by the COSMIN criteria in the design of more rigorous research (Mokkink et al., 2010a; Mokkink et al., 2010b).

The exclusion of unpublished or grey literature might have had a negative impact on the comprehensiveness of the search. Finally, similar to the experience in a recent clinimetric review of measures for children with disability (Adair et al., 2012), the use of the COSMIN in this study resulted in a floor effect in the evaluation of the reliability and validity studies. All studies garnered a “poor” rating primarily on the basis of an inadequate sample size, although the reliability studies supporting five of the seven measures employed samples that were near the COSMIN cutoff of 30 participants. This was despite having features that enhanced methodological quality in other dimensions of the study design. The small sample sizes may be related to the preliminary nature of most of the reliability studies that have not been amply followed up to date. Although a key strength of the COSMIN is its having known psychometric properties, care must be taken when interpreting study results in light of its floor effect (Adair et al., 2012).

**CONCLUSION**

This systematic review has provided comprehensive information on the clinimetric properties and clinical utility of seven instruments for assessing sitting balance in children and youth with CP. Four of these clinical measures, the PRT, SACND, SATCo, and TCMS, have acceptable overall applicability and therefore are recommended for use in clinical practice. Although there is inadequate evidence for validity of the TIS, it appears to be clinically feasible and highly reliable, and
therefore might warrant being investigated further for clinimetric properties. Further research is needed to determine responsiveness to change and thus recommendations for use as outcomes measures.

**ACKNOWLEDGMENTS**

The authors thank the following for their specific contributions to this study: Esmerita Rotor for providing the initial idea for the review question; Lenin Grajo for providing support in accessing key literature; and Esmerita Rotor and Ma. Concepcion Cabatan for providing critical comments in the early stage of manuscript preparation.

**Declaration of Interest:** The authors report no declarations of interest. The authors alone are responsible for the content and writing of this article.

**ABOUT THE AUTHORS**

Benjamin B Bañas received the Master of Rehabilitation Science (MRS) and Bachelor of Science in Physical Therapy (BSPT) degrees, and is a licensed physical therapist in the Philippines (PTRP). He is a Pediatric Physical Therapist with the Alternative Learning Resource School, Philippines and is also involved with Goal-Oriented Activities for Learning and Skill Development. Edward James R Gorgon received the Master of Physiotherapy (MPhysio) and Bachelor of Science in Physical Therapy (BSPT) degrees, and is a licensed physical therapist in the Philippines (PTRP). He holds the Chair and is an Assistant Professor at the Department of Physical Therapy, College of Allied Medical Professions, University of the Philippines Manila.

**REFERENCES**


