Accepted Manuscript

Reproducibility of the Kids Balance Evaluation Systems Test (Kids-BESTest) and the Kids-Mini-BESTest for children with cerebral palsy

Rosalee Dewar, BPhty (Hons), Andrew P. Claus, PhD, Kylie Tucker, PhD, Robert S. Ware, PhD, Leanne M. Johnston, PhD

PII: S0003-9993(19)30002-4

DOI: https://doi.org/10.1016/j.apmr.2018.12.021

Reference: YAPMR 57467

To appear in: *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION*

Received Date: 19 October 2018

Revised Date: 28 November 2018

Accepted Date: 13 December 2018

Please cite this article as: Dewar R, Claus AP, Tucker K, Ware RS, Johnston LM, Reproducibility of the Kids Balance Evaluation Systems Test (Kids-BESTest) and the Kids-Mini-BESTest for children with cerebral palsy, *ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION* (2019), doi: https://doi.org/10.1016/j.apmr.2018.12.021.

Reproducibility of the Kids Balance Evaluation Systems Test (Kids-BESTest) and the Kids-Mini-BESTest for children with cerebral palsy

Rosalee Dewar BPhty (Hons)[1], Andrew P Claus PhD[1], Kylie Tucker PhD[2], Robert S Ware PhD[3,4], Leanne M Johnston PhD[1]

[1] The University of Queensland, School of Health and Rehabilitation Sciences, Physiotherapy, Brisbane, Australia, [2] The University of Queensland, School of Biomedical Sciences, Brisbane, Australia, [3] Griffith University, Menzies Health Institute Queensland, Australia [4]The University of Queensland, Queensland Centre for Intellectual and Developmental Disability, Brisbane, Australia.

Presentations: This work has been presented at the Australasian Academy of Cerebral Palsy and Developmental Medicine. Auckland, New Zealand. March 21– March 24, 2018.

References: 38

Tables: 4

Figures: 1

Appendix: 0

Competing interests: None

Source(s) of support: We acknowledge funding support awarded by the Research Foundation, Cerebral Palsy Alliance (PG4114) through the Children's Motor Control Research Collaboration.

Acknowledgements: We would like to acknowledge the Children's Motor Control Research Collaboration and The University of Queensland, Laboratory for Motor Control and Pain Research, for support to conduct the study. We would like to thank Anne Kelly BSC (Physiotherapy) MPHTY (Paediatrics) APAM titled paediatric physiotherapist, for her assistance as the second rater for the Kids-BESTest. We would also like to thank the staff of

Choice Passion Life, the Queensland Cerebral Palsy Register and most importantly, we would like to thank the children and parents who participated in this study.

Correspondence: Rosalee Dewar, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane QLD 4072, Australia. Phone: +61 7 3365 2791.

Email: rosalee.sheather@uq.net.au

Reproducibility of the Kids Balance Evaluation Systems Test (Kids-BESTest) and the Kids-Mini-BESTest for children with cerebral palsy

ABSTRACT

Objective: To evaluate the reproducibility, including reliability and agreement, of the Kids Balance Evaluation Systems Test (Kids-BESTest) and short-form Kids-Mini-BESTest for measuring postural control in school-aged children with cerebral palsy.

Design: Psychometric study of intra-rater, inter-rater and test-retest reliability and agreement

Setting: Clinical laboratory and home.

Participants: Convenience sample of 18 children aged 8 to 17 years with ambulant cerebral palsy (Gross Motor Function Classification System I-II) with spastic or ataxic motor type.

Intervention: Not applicable.

Main Outcome measures: Postural control was assessed using the Kids-BESTest and the short-form Kids-Mini-BESTest. An experienced physiotherapist assessed all children in real-time and the testing session was videoed. The same physiotherapist viewed and scored the video twice, at least two weeks apart, to assess intra-rater reproducibility. Another experienced physiotherapist scored the same video to determine inter-rater reproducibility. Thirteen children returned for a repeat assessment with the first physiotherapist within 6 weeks and their test-retest performance was rated in real time and with video.

Results: Excellent reliability was observed for both the Kids-BESTest (ICC 0.96 to 0.99) and Kids-Mini-BESTest (ICC 0.79 to 0.98). The Smallest Detectable Change was good to excellent for all Kids-BESTest agreement analyses (5% to 9%), but poor to good for Kids-Mini-BESTest analyses (9% to 16%).

Conclusion: The Kids-BESTest shows an excellent ability to discriminate postural control abilities of school-aged children with cerebral palsy and it has a low Smallest Detectable Change, suitable for use as a pre-post intervention outcome measure. Although the Kids-Mini-BESTest is 5-10 min shorter to administer, it has poorer reproducibility and focuses only on falls-related balance, which excludes two domains of postural control.

Key words: Postural Balance; Cerebral Palsy; Reproducibility; Kids-BESTest; Kids-Mini-BESTest

ABBREVIATIONS

CP Cerebral Palsy

BESTest Balance Evaluation Systems Test

Cerebral palsy (CP) is disorder of movement and posture, caused by a permanent disturbance in the fetal or infant brain [1]. Considerable research has explored classification and assessment of movement disorders experienced by children with CP [2-7] however there is limited research on assessment of postural control dysfunction [8]. Postural control is the ability to control the body's position in space for balance and orientation [8]. It can be understood using Shumway-Cook and Woollcott's *Systems Approach*, which explains how multiple body systems contribute to postural control [9]. Children with CP have been reported to display postural control deficits across all *Systems Approach* components, including: *Musculoskeletal components* (poor muscle strength and joint range of motion) [10], *Sensory systems and sensory strategies* (poor vestibular, vision ,proprioception function and how they are integrated) [11, 12], *Anticipatory mechanisms* (dysfunctional feed forward postural adjustments) [13-17],*Adaptive mechanisms* and *Neuromuscular synergies* (poor ankle, hip and stepping strategies and feedback postural reactions) [13, 18-21] and *Internal representations* of limits of stability (reaching in sitting and standing) [22]. Despite children showing deficits across the systems of postural control currently, there is no comprehensive systems-based postural control assessment with published psychometrics for children with CP to aid clinicians to develop targeted intervention programs [8].

Optimal postural control assessment for children with CP requires examination of performance across all systems to profile deficits and allow development of targeted rehabilitation programs. However, comprehensive postural control assessment does not seem to be occurring in practice for children with CP. A recent Delphi study revealed that researchers and clinicians utilise mostly unidimensional tests [8], for example single-item tests such as timed single-leg stance for *Anticipatory Mechanisms* [23]. Or, single-aspect tests, such as reactionary posture and balance responses (*Adaptive Mechanisms)* examined in the Neuro-Sensory Motor Developmental Assessment [24]. The Delphi revealed that the main limitation to comprehensive assessment was a lack of multi-dimensional paediatric clinical tools. A recent systematic review of postural control assessments for children with CP [25] reported only two assessments that assessed more than one

postural control system: the Berg Balance Scale, and its companion paediatric version, the Pediatric Balance Scale [26]. Even so, both versions evaluate only 3 of 7 *Systems Approach* components: *Anticipatory mechanisms*, *Sensory systems* and *Internal representations*.

In response to this practice gap, data was recently published for reproducibility of the Kids-BESTest in typically developing children [27], which is a comprehensive postural control assessment modified for children from the adult BESTest [28]. The Kids-BESTest assesses all *Systems Approach* components [27], through 36 tasks (27 items) divided into six domains (see Figure 1) : *Biomechanical constraints* (5 tasks, 0-15 points) ; *Stability limits/verticality* (7 tasks, 0-21 points); *Reactive postural responses* (6 tasks, 0-18 points)*; Anticipatory postural adjustments* (6 tasks, 0-18 points); *Sensory orientation* (5 tasks, 0-15points) and *Stability in gait* (7 tasks, 0-21 points). A short-form of the Kids-BESTest is called the Kids-Mini-BESTest, which contains 17 tasks (14 items) divided into four domains: *Anticipatory postural adjustments* (4 tasks, 0-6 points)*; Reactive postural responses* (4 tasks, 0-6 points); *Sensory orientation* (3 tasks, 0-6 points) and *Stability in gait* (6 tasks, 0-10 points).

A recent study showed that the Kids-BESTest is a feasible and reproducible tool for typically developing children [27]. Reproducibility was good to excellent for the Kids-BESTest and fair to excellent for the short-form Kids-Mini-BESTest [27]. Both test versions could discriminate postural control abilities, but this was better for the Kids-BESTest. Both versions demonstrated that they could be sensitive to detect a change in postural control function over days [27]. Specific research is now needed with children with CP and other motor disorders to determine validity, reproducibility and clinical utility for specific clinical practice applications. The aim of this study is therefore to evaluate the reproducibility of the Kids-BESTest and the Kids-Mini-BESTest when assessing postural control in school-aged children with CP.

METHOD

Study Design and Participants

Intra-rater, inter-rater and test-retest reproducibility of the Kids-BESTest and Kids-Mini-BESTest were examined with school-aged children with CP. Ethical approval was obtained from appropriate Human Research Ethics Committees.

Children were eligible for inclusion if they (i) had CP and were (ii) aged between 8-18 years, (iii) ambulant (GMFCS I-III), and (iv) able to follow child-friendly test instructions. Children were excluded if they had a history of: (i) spasticity management (e.g. chemodenervation) within three months, (ii) orthopaedic or neurological surgery within 12 months, (iii) intellectual or behavioural difficulties limiting full participation in assessment, (iv) uncontrolled seizures or (v) co-morbidities interfering in physical functioning e.g. autism. Potential participants were identified from (i) databases of a state-wide CP service and CP register (ii) staff referrals from the CP service, or (iii) parent referrals in response to community advertisements. Prior to involvement children and guardians were provided with written and verbal study information. All guardians signed consent forms and all children signed assent forms.

Outcome measures

Postural control of children with CP was assessed using the Kids-BESTest according to the protocol published by Dewar et al 2017 [27]. Each of the 36 tasks in the Kids-BESTest was scored from 3 (best performance) to 0 (worst performance) to generate six *Domain scores*, and a *Total Score* ranging from 0 to 108. The tool takes approximately 30 minutes to administer. The subset of Kids-Mini-BESTest items was then re-scored using the Kids-Mini-BESTest scoresheet [27].The Kids-Mini-BESTest, contains a subset of 14 tasks (14 items) evaluating four *Systems Approach* domains. The subset is designed to quickly identify individuals at risk of falls [29]. It takes 15 minutes to administer and items are scored on a reduced scale from to 2 (best performance) to 0 (worse performance) with a maximum of 28 points. For both test versions, item scores were summed to produce a *Total Score*

and *Domain scores* (Kids-BESTest = 6; Kids-Mini-BESTest = 4). Performance was scored by two paediatric physiotherapists (Examiner 1 - the first author; and Examiner 2 - an independent examiner) each with 20 years of experience with children with CP. To promote consistency, both examiners completed administration and scoring training via the BESTest website [30] as well as training on the paediatric modifications using the Kids-BESTest protocol [27].

Procedure

Reproducibility was examined under four conditions: (1) Test-retest real time; (2) Test-retest video; (3) Intra-rater video and (4) Inter-rater video. To achieve this, children were assessed in real time and all assessments were videoed concurrently using the published Kids-BESTest video recording protocol [27]. Real time assessments were completed on Day 1 (n=18) and Day 2 (n=13) by Examiner 1. The interval between real-time assessments was 1 to 42 days. Video-based assessments were performed retrospectively after all real-time assessments were completed. Test-retest reproducibility was evaluated from Day 1 and Day 2 performance in real time and via video by Examiner 1. Intra-rater reproducibility was assessed with Day 1 video by Examiner 1. Inter-rater reproducibility was assessed with Day 1 video by Examiner 1 and separately by the independent Examiner 2. In each case, reproducibility was evaluated for the Total Score as well as all Domains of the Kids-BESTest (6 domains) and the Mini-BESTest (4 domains). To enhance family centred care, families were given the option for their child to participate at the university, their local CP clinic, or their home, whichever they felt would be optimal for their child. In each case, assessments were conducted in an open room space with standardized equipment and floor markings used according to the Kids-BESTest administration and video protocols [27].

Analysis

Reproducibility is the degree to which repeated measures of the tests provide similar results. Reproducibility includes two components: (i) agreement and (ii) reliability [31]. Agreement assesses how close the results of repeated measurements are, and the margins that may be used to represent

real clinical change, as opposed to random measurement error [31]. Reliability evaluates how well children can be distinguished from one another despite measurement error [31]. Statistical analysis was performed using Stata statistical software v 13.0 (StataCorp, College Station, TX, USA).

Agreement analysis involved calculation of percentage agreement, Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) and 95% Confidence Intervals (CI) for Limits of Agreement (LoA) using Bland-Altman methods [31, 32]. Suitable percentage agreement was set *a priori* consistent with previous work with typically developing children. For the Kids-BESTest Total score it was defined as: excellent = >90% within 4 points, good = >80% within 4 points, fair = >60% within 4 points and poor = <60% within 4 points. For the Kids-Mini-BESTest Total score it was defined as: excellent = >90% within 2 points, good = >80% within 2 points, fair = >60% within 2 points and poor = <60% within 2 points. For the domains the *a priori* agreement values were set at 2 points for Kids-BESTest domain scores and 1 point for Kids-Mini-BESTest scores. The SEM was calculated to indicate the measurement error of both tools and this was used to calculate the SDC. The SDC is the smallest change in score that may be used to indicate real change, not just measurement error [31]. To allow comparison, the SDC was expressed as a percentage of the *Total score* and each *Domain score* for each test version. Consistent with previous work, the SDC was defined as excellent = 0-5%; good = >5-10%; fair = >10-15%; or poor = > 15% agreement [27]. The 95% CI LoA was calculated as the range within which different examiners or the same examiner produced similar scores on separate assessment occasions.

Reliability was calculated via Intra-class Correlation Coefficients (ICC) and 95% confidence intervals using analysis of variance models. Consistent with previous work [27, 33, 34], an ICC was defined as excellent = > 0.75; good = 0.74 – 0.60; fair = 0.59 – 0.40; and poor = < 0.4.

RESULTS

Parents of 21 children responded to the recruitment process. Three children were excluded (1 x recent surgery, 2 x intellectual disability). The remaining 18 children were: aged between 8 and 17 years; all were independently mobile; 13 had spastic hemiplegia, 4 had spastic diplegia and one had ataxia (Table 1). Of these 18 children, 13 children returned for a repeat assessment (also Table 1). All participants were able to complete and be scored on all items of the Kids-BESTest and Kids-Mini-BESTest so there were no missing items as a result.

Kids-BESTest Results

1. Intra-rater reproducibility (video assessment)

The Kids-BESTest *Total Score* showed good intra-rater agreement (89% within 4 points, Table 2) and excellent intra-rater reliability (ICC = 0.99, 95% CI 0.97 to 1.00, Table 3A) when assessed via video. The intra-rater reliability of all *Domains* was also excellent (ICC = 0.92-0.98, Table 3A). The SDC for the *Total Score* was excellent (5.5 points, or 5%, Table 2) and *Domains* ranged from good to fair (1.2 to 2.9 points, or 7% to 14%, Table 2). This means that children must improve by 6 points on the *Total score,* or 2-3 points depending on the *Domain* to demonstrate real clinical change when assessed by one examiner using video.

2. Inter-rater reproducibility (video assessment)

The Kids-BESTest *Total Score* showed good inter-rater agreement (83% within 4 points, Table 2) and excellent inter-rater reliability (ICC = 0.97, 95% CI 0.94 to 1.00, Table 3A) when assessed via video. The inter-rater reliability of *Domains* was good to excellent (ICC = 0.70 to 0.93, Table 3A). The SDC for the *Total Score* was good (9.3 points, 9%, Table 2) and *Domains* ranged from fair to poor (2 to 4.5 points, 12% to 21%, Table 2). This means that children must improve by 10 points on the *Total Score,* or 2-5 points depending on the *Domain* to demonstrate real clinical change when assessed by two different examiners via video.

3. Test-retest reproducibility (real-time and video assessment)

Test-retest reliability for the Kids-BESTest *Total score* was excellent for both video (ICC = 0.96, 95% CI 0.92 to 1.00, Table 3A) and real-time assessment (ICC = 0.97, 95% CI 0.95 to 1.00, Table 3A). Similarly, test-retest reliability for *Domain scores* was excellent using both video (ICC = 0.77 to 0.88, Table 3A) and real time assessment (ICC = 0.76 to 0.94, Table 3A).

Test-retest agreement for the Kids-BESTest *Total score* was excellent when assessed with video (92% within 4 points, Table 2) and good in real-time (84% within 4 points, Table 2). Test-retest agreement varied between *Domains* from 67% to 100% within 2 points. For example, the *Biomechanical constraints* domain demonstrated higher agreement when assessed in real time (100% within 2 points) compared to video (92% within 2 points). In contrast, the *Reactive* domain showed the opposite pattern.

The SDC for the Kids-BESTest *Total Score* was excellent for real-time (5.6 points, or 5%, Table 2) and good for video (6.1 points, or 6%, Table 2). The SDC for *Domains* ranged from fair to poor for video (1.9 to 3.5 points, or 11% to 21%, Table 2) and good to poor for real-time (1.5 to 3.9 points, or 10% to 21%, Table 2). This means that children must improve by 6 points on the *Total score,* or 2-4 points depending on the *Domain* to demonstrate real clinical change when scored on two different occasions using either video or real-time modes.

<u>*Kids Mini-BESTest Results*</u>

1. Intra-rater reproducibility (video assessment)

The Kids-Mini-BESTest *Total Score* showed excellent intra-rater agreement (94% within 2 points, Table 4) and excellent reliability (ICC = 0.98, 95% CI 0.96 to 1.00, Table 3B) when assessed using video. The SDC was good for the *Total Score* (2.4 points, 9%, Table 4) but fair to poor for the *Domains* (range 0.7 to 2.1 points, or 11% to 30%, Table 4). This means that children must improve by 3 points on the *Total score,* or 1-3 points depending on the *Domain* to demonstrate real clinical change when scored via video by one examiner.

2. Inter-rater reproducibility (video assessment)

The Kids-Mini-BESTest *Total Score* showed good inter-rater agreement (89% within 2 points, Table 4) and excellent reliability (ICC = 0.97, 95% CI 0.93 to 1.00, Table 3B) when scored via video . The SDC was fair for the *Total Score* (3.3 points, 12%, Table 4) and fair to poor for the *Domains* (0.9 to 2.2 points, 15% to 27%, Table 4). This means that children must improve by 4 points on the *Total score,* or 1-3 points depending on the *Domain* to demonstrate real clinical change when scored by via video two different examiners.

3. Test-retest reproducibility (real-time and video assessment)

The Kids-Mini-BESTest *Total Score* showed better test-retest agreement when scored via video (77% within 2 points) compared to in real-time (62% within 2 points, Table 4). Test-retest reliability for the *Total Score* was excellent when scored either via video (ICC = 0.90, 95% CI 0.79 to 1.00) or in real-time (ICC = 0.79, 95% CI 0.57 to 1.00, Table 3B). However, reliability for the *Domains* was better using video (ICC 0.70 to 0.88) than real-time (ICC 0.49 to 0.76, Table 3B). The SDC was fair for the *Total Score* using video (3.9 points, 14%) but poor for real-time (4.6 points, 16%, Table 4). This means that children must improve by 4 points on the *Total score* when scored by video, or 5 points in real-time to demonstrate real clinical change when scored by one examiner.

DISCUSSION

The Kids-BESTest is the first assessment to address all *Systems Approach* components for postural control in children with CP. Our results indicate that the Kids-BESTest and Kids-Mini-BESTest versions are feasible and reproducible for this population. The tests can detect real clinical change (high agreement and low SDC) and different abilities of postural control (high reliability) in school-aged children with CP, with the full Kids-BESTest showing the best overall results.

Both BESTest versions demonstrated excellent reliability for differentiating postural control function when administered on different days, or scored by different examiners, with the Kids-BESTest showing superior results. This finding is consistent with previous studies, which showed that the Full-BESTest was more reliable than the Mini-BESTest for assessing postural control in typically developing school-aged children [27] and adults with neurological conditions [35, 36]. Our results suggest that the Kids-BESTest may be better at differentiating postural control function of children with CP than typically developing children. Scores for some items reached a ceiling for typically developing children, which decreased the reliability score for that population [27]. In contrast, no child with CP reached a ceiling on any domain and the CP group showed greater variation in skill level enabling the test to effectively differentiate between children in this group.

In terms of whether individual Kids-BESTest *Domains* can be used to profile postural control dysfunction for children with CP, our data indicates that all Kids-BESTest *Domains* have a good to excellent ability to discriminate between different levels of performance and the Kids-BESTest would be preferred to the Mini-BESTest for a more comprehensive set of domains. As expected, agreement for the Kids-BESTest *Domains* was slightly better within-day compared to between days. The same trend was reported previously for typically developing children [27]. These results highlight performance variability that children with or without CP might show on different days and emphasises the need for consistency in test application. It also supports future research to confirm individual test item validity.

Our data for children with CP suggests that the Kids-BESTest may be better at detecting clinical change between days than the Kids-Mini-BESTest. This is suggested because of the smaller SDC seen for the *Total score* of the Kids-BESTest compared to the Kids-Mini-BESTest. The Kids-BESTest *Total score* SDC was consistent with the SDC reported for typically developing children [27]. The Kids-Mini-BESTest *Total score* SDC appeared to be higher for children with CP than previously reported for typical children [27] but similar to results for adults with neurological

disorders impacting gait [35]. In terms of clinical practice, if measuring change in performance pre-post intervention, we recommend that an increase of at least 6 points on the Kids-BESTest or at least 4 points on the Kids-Mini-BESTest needs to be seen to confirm a clinically significant improvement for children with cerebral palsy.

Finally, in terms of administration method, although our data showed little difference between test-retest SDCs for the Kids-BESTest for real-time or video evaluation, the SDCs for the Kids-Mini-BESTest were better when evaluated via video versus real time. This difference may occur because some items benefit from having an examiner feel subtle responses during real-time handling such as hip/trunk lateral strength and some items are best scored from video so that performance can be seen from a distance such as stability in gait. Therefore, for best results, we recommend a combination of real-time scoring plus retrospective video review to confirm scoring for children with CP. This is in keeping with the recommendation for the Kids-BESTest for typically developing children [27] and most other reliable motor assessments for children with CP (e.g. Assisting Hand Assessment [37] or Gross Motor Function Measure Challenge Module [38]).

Strengths, Limitations and Future Directions for Research

Although our data shows the comprehensive Kids-BESTest is a feasible and reliable battery for children with CP, it did demonstrate varying levels of agreement within the *Domains.* Future studies could investigate validity of test items and responsiveness to age or time between assessments, practice and intervention tailored for individual domains. Testing with children with other motor types is also recommended. Finally, although this test is appropriate for school-aged children, development of a similar comprehensive postural control assessment in younger children is needed.

CONCLUSION

The Kids-BESTest is the first assessment to evaluate all systems contributing to postural control in children with CP. The Kids-BESTest shows excellent ability to distinguish between different levels of postural control abilities for school-aged children with CP. It has a low SDC, indicating it has good potential for use as an outcome measure pre and post postural control interventions. Although the Kids Mini-BESTest is faster to administer, it has lower reproducibility and does not include two important domains. BESTest training as outlined in our methods is recommended for all examiners prior to using the Kids-BESTest for clinical practice or research. The Kids-BESTest warrants future research in children with cerebral palsy and other clinical populations to investigate its responsiveness in intervention trials.

REFERENCES

1. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D et al. A report: the definition and classification of cerebral palsy April 2006. Dev Med Child Neurol Suppl 2007;109(suppl 109):8-14.

2. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Öhrvall A-M et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol 2006;48(07):549-54.

3. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. Dev Med Child Neurol 1997;39(4):214-23.

4. Vos-Vromans DC, Ketelaar M, Gorter JW. Responsiveness of evaluative measures for children with cerebral palsy: the Gross Motor Function Measure and the Pediatric Evaluation of Disability Inventory. Disabil Rehabil 2005;27(20):1245-52.

5. Glazebrook CM, Wright FV. Measuring Advanced Motor Skills in Children With Cerebral Palsy: Further Development of the Challenge Module. Pediatr Phys Ther 2014;26(2):201-13.

6. Russell DJ, Avery LM, Walter SD, Hanna SE, Bartlett DJ, Rosenbaum PL et al. Development and validation of item sets to improve efficiency of administration of the 66 item Gross Motor Function Measure in children with cerebral palsy. Dev Med Child Neurol 2010;52(2):e48-e54.

7. Ko J, Kim M. Reliability and responsiveness of the gross motor function measure-88 in children with cerebral palsy. Phys Ther 2013;93(3):393-400.

8. Dewar R, Claus AP, Tucker K, Johnston LM. Perspectives on Postural Control Dysfunction to Inform Future Research: A Delphi Study for Children With Cerebral Palsy. Arch Phys Med Rehabil 2017;98(3):463-79.

9. Shumway-Cook A, Woollacott MH. Motor control: translating research into clinical practice. Fourth Edition ed.: Lippincott Williams & Wilkins; 2012.

10. Lowes LP, Westcott SL, Palisano RJ, Effgen SK, Orlin MN. Muscle force and range of motion as predictors of standing balance in children with cerebral palsy. Phys Occup Ther Pediatr 2004;24(1-2):57-77.

11. Lowes LP, Habib Z, Bleakney D, Westcott S. Relationship between clinical measures of balance and functional abilities in children with cerebral palsy. Pediatr Phys Ther 1996;8(4):176.

12. Westcott SL, Zaino CA, Miller F, Thorpe DE. Comparison of Postural Muscle Activity in Children of Different Ages During Stand and Reach From Firm. Compliant, and Narrow Surfaces: 9. Pediatr Phys Ther 1997;9(4):207.

13. Bigongiari A, De Andrade E Souza F, Franciulli PM, Neto SER, Araujo RC, Mochizuki L. Anticipatory and compensatory postural adjustments in sitting in children with cerebral palsy. Human movement science 2011;30:988-97.

14. Girolami GL, Shiratori T, Aruin AS. Anticipatory postural adjustments in children with hemiplegia and diplegia. J Electromyogr Kinesiol 2011;21:988-97.

15. Tomita H, Fukaya Y, Honma S, Ueda T, Yamamoto Y, Shionoya K. Anticipatory postural muscle activity associated with bilateral arm flexion while standing in individuals with spastic diplegic cerebral palsy: A pilot study. Neurosci Lett 2010;479(2):166-70.

16. Westcott SL, Zaino CA, Miller F, Thorpe D, Unanue R. Anticipatory postural coordination and functional movement skills by degree of cerebral palsy in children age 6-12 years. Proceedings of the Society for Neuroscience Abstracts; 1998.

17. Westcott SL, Zaino CA, Unanue R, Thorpe D, Miller F. Comparison of anticipatory postural control and dynamic balance ability in children with and without cerebral palsy. Dev Med Child Neurol 1998;40:14.

18. Westcott SL, Burtner P. Postural control in children. Phys Occup Ther Pediatr 2004;24(1-2):5-55.

19. Burtner P, Qualls C, Woollacott M. Muscle activation characteristics of stance balance control in children with spastic cerebral palsy. Gait Posture 1998;8(3):163-74.

20. Burtner P, Woollacott M, Craft G, Roncesvalles M. The capacity to adapt to changing balance threats: A comparison of children with cerebral palsy and typically developing children. Dev Neurorehabil 2007;10(3):249-60.

21. Roncesvalles M, Woollacott M, Burtner P. Neural factors underlying reduced postural adaptability in children with cerebral palsy. Neuroreport 2002;13(18):2407-10.

22. Katz-Leurer M, Rotem H, Keren O, Meyer S. Balance abilities and gait characteristics in post-traumatic brain injury, cerebral palsy and typically developed children. Dev Neurorehabil 2009;12(2):100-5.

23. Liao HF, Mao PJ, Hwang AW. Test-retest reliability of balance tests in children with cerebral palsy. Dev Med Child Neurol 2001;43(3):180-6.

24. Burns YR, Ensbey RM, Norrie MA. The Neuro-sensory motor developmental assessment part 1: development and administration of the test. Aust J Physiother 1989;35(3):141-9.

25. Saether R, Helbostad JL, Riphagen II, Vik T. Clinical tools to assess balance in children and adults with cerebral palsy: a systematic review. Dev Med Child Neurol 2013;55(11):988-99.

26. Franjoine MR, Gunther JS, Taylor MJ. Pediatric Balance Scale: a modified version of the Berg Balance Scale for the school-age child with mild to moderate motor impairment. Pediatr Phys Ther 2003;15(2):114-28.

27. Dewar R, Claus AP, Tucker K, Ware R, Johnston LM. Reproducibility of the Balance Evaluation Systems Test (BESTest) and the Mini-BESTest in school-aged children. Gait Posture 2017;55:68-74.

28. Horak FB, Wrisley DM, Frank J. The balance evaluation systems test (BESTest) to differentiate balance deficits. Phys Ther 2009;89(5):484-98.

29. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation System Test: the mini-BESTest. J Rehabil Med 2010;42(4):323.

30. Horak FB. [cited 2016 4th September]. Available from: URL: http://www.bestest.us.

31. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59(10):1033-9.

32. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17(4):571-82.

33. Auld ML, Ware RS, Boyd RN, Moseley GL, Johnston LM. Reproducibility of tactile assessments for children with unilateral cerebral palsy. Phys Occup Ther Pediatr 2012;32(2):151-66.

34. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 1981;86(2):127-37.

35. Di Carlo S, Bravini E, Vercelli S, Massazza G, Ferriero G. The Mini-BESTest: a review of psychometric properties. Int J Rehabil Res 2016;39(2):97-105.

36. Chinsongkram B, Chaikeeree N, Saengsirisuwan V, Viriyatharakij N, Horak FB, Boonsinsukh R. Reliability and validity of the Balance Evaluation Systems Test (BESTest) in people with subacute stroke. Phys Ther 2014;94(11):1632-43.

37. Krumlinde‐Sundholm L, Holmefur M, Kottorp A, Eliasson AC. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. Dev Med Child Neurol 2007;49(4):259-64.

38. Wilson A, Kavanaugh A, Moher R, McInroy M, Gupta N, Salbach NM et al. Development and pilot testing of the challenge module: a proposed adjunct to the Gross Motor Function Measure for high-functioning children with cerebral palsy. Phys Occup Ther Pediatr 2011;31(2):135-49.

Figure 1. Domains of the Kids-BESTest and the systems each assess from the Systems Approach Framework. Each domain may involve more than one systems however the measurement criteria contained within that item will predominately focus on one or two systems. Stability in Gait domain assesses functional integration of all systems.

Table 1. Summary of participant characteristics

Table 2. Agreement analyses for the Kids-BESTest

Table 3. Reliability analyses for A. the Kids-BESTest and B. Kids-Mini-BESTest

Table 4. Agreement of the Kids Mini-BESTest

Table 1. Summary of participant characteristics

		Day 1 Intra-rater video Inter-rater video (n=18)	Day 2 Test-retest real-time Test-retest video (n=13)
Male, n (%)		12 (67%)	10 (77%)
Age, mean (SD)		11.5 (2.8) years	10.9 (2.6) years
Body mass index, mean (SD)		17.8 (4.5)	16.4 (3.9)
Height, mean (SD)		145.1 (14.4) cm	141.3 (12.8) cm
Weight, mean (SD)		40.0 (16.0) kg	36.0 (13.4) kg
GMFCS	I	11	10
	II	7	3
MACS	I	14	10
	II	4	3
Hemiplegia		13	11
Diplegia		4	1
Ataxia		1	1

GMFCS, Gross Motor Function Classification System; MACS, Manual Ability Classification System.

Table 2. Agreement analyses for the Kids-BESTest

Kids-BESTest	Increment (Range)	Agreement (%)		SEM	SDC(%)*	95% CI for LoA
		Within 2 points	Within 4 points			
1. Intra-rater agreement (n= 18, video 1, one assessor)						
Biomechanical Constraints	1 (0-15)	100		0.48	1.3 (9%)	1.3 to -1.5
Stability Limits and Verticality	1 (0-21)	100		0.55	1.5 (7%)	1.2 to -2.0
Transitions/Anticipatory	1 (0-18)	100		0.48	1.3 (7%)	1.3 to -1.5
Reactive	1 (0-18)	94	100	0.84	2.4 (13%)	2.7 to -2.2
Sensory Orientation	1 (0-15)	100		0.44	1.2 (8%)	1.1 to -1.5
Stability in Gait	1 (0-21)	94	100	1.00	2.9 (14%)	3.1 to -3.0
Total score	1 (0-108)	67	89	1.98	5.5 (5%)	5.3 to -6.3
2. Inter-rater agreement (n=18, video 1, two assessors)						
Biomechanical Constraints	1 (0-15)	100		1.10	3.1 (21%)	3.0 to -3.5
Stability Limits and Verticality	1 (0-21)	100		1.16	3.2 (15%)	2.7 to -4.1
Transitions/Anticipatory	1 (0-18)	100		0.78	2.2 (12%)	1.5 to -3.1
Reactive	1 (0-18)	78	100	1.22	3.4 (19%)	4.4 to -2.7
Sensory Orientation	1 (0-15)	100		0.88	2.4 (16%)	2.0 to -3.1
Stability in Gait	1 (0-21)	67	94	1.62	4.5 (21%)	5.8 to -3.8
Total score	1 (0-108)	72	83	3.08	9.3 (9%)	9.3 to -10.3
3. Test-retest agreement (n=13, video 1 and 2, one assessor)						
Biomechanical Constraints	1 (0-15)	92	100	0.99	2.8 (19%)	2.8 to -3.1
Stability Limits and Verticality	1 (0-21)	92	100	1.01	2.8 (13%)	2.7 to -3.3
Transitions/Anticipatory	1 (0-18)	100		0.67	1.9 (11%)	1.7 to -2.3
Reactive	1 (0-18)	92	100	1.19	3.3 (18%)	3.3 to -3.6
Sensory Orientation	1 (0-15)	92	100	1.10	3.1 (21%)	3.4 to -3.1
Stability in Gait	1 (0-21)	84	100	1.27	3.5 (17%)	3.4 to -3.9
Total score	1 (0-108)	46	92	2.19	6.1 (6%)	5.4 to -7.4
4. Test-retest agreement (n= 13, real-time, one assessor)						
Biomechanical Constraints	1 (0-15)	100		0.53	1.5 (10%)	1.9 to -1.3
Stability Limits and Verticality	1 (0-21)	92	100	0.79	2.2 (10%)	1.9 to -2.7
Transitions/Anticipatory	1 (0-18)	100		0.74	2.0 (11%)	2.8 to -1.5
Reactive	1 (0-18)	84	100	1.39	3.8 (21%)	3.8 to -4.3
Sensory Orientation	1 (0-15)	92	100	0.98	2.7 (18%)	3.2 to -2.6
Stability in Gait	1 (0-21)	84	100	1.39	3.9 (19%)	3.9 to -4.3
Total score	1 (0-108)	69	84	2.03	5.6 (5%)	6.3 to -5.6

*SEM: standard error of the mean, SDC: smallest detectable change, CI: confidence interval, LoA: limits of agreement, * SDC is expressed as a percentage of the Total score or domain score to allow comparison of scores with different ranges*

Table 3. Reliability analyses for A. Kids-BESTest and B. Kids-Mini-BESTest

	Intra-rater video (n=18, Day 1, one assessor)		**Inter-rater video (n= 18, Day 1, two assessors)**		**Test-retest video (n=13, Day 1 and 2, one assessor)**		**Test-retest real-time (n=13, Day 1 and 2, one assessor)**	
A. Kids-BESTest	ICC	95% CI	ICC	95% CI	ICC	95% CI	ICC	95% CI
Biomechanical Constraints	0.97	0.94 to 1.00	0.85	0.71 to 0.98	0.83	0.66 to 1.00	0.94	0.89 to 1.00
Stability Limits & Verticality	0.92	0.86 to 0.99	0.70	0.47 to 0.94	0.78	0.56 to 1.00	0.89	0.77 to 1.00
Transitions/Anticipatory	0.98	0.95 to 1.00	0.89	0.79 to 0.99	0.84	0.69 to 1.00	0.82	0.65 to 1.00
Reactive	0.96	0.93 to 1.00	0.92	0.85 to 0.99	0.88	0.76 to 1.00	0.87	0.73 to 1.00
Sensory Orientation	0.98	0.97 to 1.00	0.93	0.87 to 0.99	0.78	0.57 to 1.00	0.79	0.59 to 1.00
Stability in Gait	0.93	0.87 to 1.00	0.83	0.69 to 0.98	0.77	0.54 to 1.00	0.76	0.53 to 0.99
Total score	0.99	0.97 to 1.00	0.97	0.94 to 1.00	0.96	0.92 to 1.00	0.97	0.95 to 1.00
B. Kids-Mini-BESTest								
Transitions/Anticipatory	0.95	0.91 to 1.00	0.87	0.76 to 0.98	0.77	0.55 to 1.00	0.74	0.48 to 0.99
Reactive	0.88	0.78 to 0.99	0.92	0.85 to 0.99	0.83	0.65 to 1.00	0.53	0.14 to 0.93
Sensory Orientation	0.97	0.95 to 1.00	0.92	0.84 to 0.99	0.88	0.75 to 1.00	0.76	0.53 to 1.00
Stability in Gait	0.90	0.80 to 0.99	0.88	0.78 to 0.99	0.70	0.41 to 0.98	0.49	0.06 to 0.91
Total score	0.98	0.96 to 1.00	0.97	0.93 to 1.00	0.90	0.79 to 1.00	0.79	0.57 to 1.00

ICC, Intra-class correlation coefficient; CI, Confidence Interval.

Table 4. Agreement of the Kids Mini-BESTest

Kids Mini-BESTest	Increment (Range)	Agreement (%)		SEM	SDC (%)*	95% CI of LoA (95% CI)
		Within 1 point	Within 2 points			
1. Intra-rater agreement (n=18, video 1, one assessor)						
Transitions/Anticipatory	1 (0-6)	100		0.24	0.7 (11%)	0.7 to -0.7
Reactive	1 (0-6)	100		0.66	1.8 (30%)	1.9 to -2.0
Sensory Orientation	1 (0-6)	100		0.30	0.8 (14%)	0.8 to -0.9
Stability in Gait	1 (0-10)	94	100	0.74	2.1 (21%)	2.6 to -1.8
Total score	1 (0-28)	78	94	0.88	2.4 (9%)	2.9 to -2.3
2. Inter-rater agreement (n=18, video 1, two assessors)						
Transitions/Anticipatory	1 (0-6)	100		0.33	0.9 (15%)	0.7 to -1.2
Reactive	1 (0-6)	89	100	0.59	1.6 (27%)	1.8 to -1.6
Sensory Orientation	1 (0-6)	94	100	0.46	1.3 (22%)	1.1 to -1.6
Stability in Gait	1 (0-10)	83	100	0.78	2.2 (22%)	2.5 to 2.1
Total score	1 (0-28)	50	89	1.20	3.3 (12%)	3.3 to 3.8
3. Test-retest agreement (n=13, video 1 and 2, one assessor)						
Transitions/Anticipatory	1 (0-6)	100	92	0.42	1.2 (20%)	1.0 to -1.5
Reactive	1 (0-6)	92	100	0.64	1.8 (30%)	2.0 to -1.7
Sensory Orientation	1 (0-6)	100		0.41	1.1 (18%)	1.2 to -1.2
Stability in Gait	1 (0-10)	85	100	0.91	2.5 (25%)	2.8 to -2.5
Total score	1 (0-28)	46	77	1.43	3.9 (14%)	4.3 to -4.1
4. Test-retest agreement (n=13, real-time, one assessor)						
Transitions/Anticipatory	1 (0-6)	100		0.39	1.1 (18%)	1.3 to -1.0
Reactive	1 (0-6)	77	92	0.91	2.5 (42%)	2.5 to -2.8
Sensory Orientation	1 (0-6)	100		0.35	1.0 (16%)	1.1 to -1.0
Stability in Gait	1 (0-10)	53	77	1.34	3.7 (37%)	3.4 to -4.5
Total score	1 (0-28)	54	62	1.67	4.6 (16%)	4.5 to -5.4

*SEM: standard error of the mean, SDC: smallest detectable change, CI: confidence interval, LoA: limits of agreement, * SDC is expressed as a percentage of the Total score or domain score to allow comparison of scores with different ranges*

Figure 1. Domains of the Kids-BESTest and the systems each assess from the Systems Approach Framework. Each domain may involve more than one systems however the measurement criteria contained within that item will predominately focus on one or two systems. Stability in Gait domain assesses functional integration of all systems.