Exploration of Medicine

Open Exploration

Open Access Original Article

Agreement between remote and in-person assessments of mobility and balance tests for children with cerebral palsy. A pilot study

Aspasia Mavronasou†, Georgia Kechagia†, Panagiotis Dalamarinis, Eleni A. Kortianou*

Clinical Exercise Physiology and Rehabilitation Laboratory, Physiotherapy Department, University of Thessaly, 35100 Lamia, Greece

†These authors contributed equally to this work.

***Correspondence:** Eleni A. Kortianou, Clinical Exercise Physiology and Rehabilitation Laboratory, Physiotherapy Department, University of Thessaly, 35100 Lamia, Greece. ekortianou@uth.gr

Academic Editor: Hua Su, University of California, USA

Received: April 30, 2026 **Accepted:** July 15, 2026 **Published:** July 28, 2026

Cite this article: Mavronasou A, Kechagia G, Dalamarinis P, Kortianou EA. Agreement between remote and in-person assessments of mobility and balance tests for children with cerebral palsy. A pilot study. Explor Med. 2026;7:1001419. https://doi.org/10.37349/emed.2026.1001419

Abstract

Aim: To determine the level of agreement between remote and in-person assessment for the Timed Up and Go test (TUG), the Lateral Step-Up test (LSU), the Five-Times-Sit-To-Stand test (FTSTS), and the Pediatric Balance Scale (PBS) in children with cerebral palsy (CP).

Methods: Fifteen children diagnosed with unilateral and bilateral spastic CP and Gross Motor Function Classification System levels I–II underwent these four tests in two distinct environments. These assessments were conducted 24 to 48 hours apart, first in person at the physiotherapy clinic and then remotely at home via the VSee platform. Before the remote assessment, all parents attended a preliminary training session that covered technical proficiency with VSee, the proper setup of home equipment, and essential safety protocols.

Results: Children had a median age of 7 (IQR = 6–10) years old. All children successfully finished all four tests in both the home and clinic settings without any reported safety issues. The agreement was good for the time of repetitions in the FTSTS: ICC = 0.884 (95% CI: 0.654 to 0.961; MD = –1.26 seconds, 95% CI: –2.81 to 0.28, p = 0.102), and excellent for the total time of the TUG: ICC = 0.998 (95% CI: 0.990 to 0.999; MD = –0.42 seconds, 95% CI: –0.76 to –0.08, p = 0.017), the total score in the PBS: ICC = 0.998 (95% CI: 0.992 to 0.999; MD = 0.26 units, 95% CI: 0.01 to 0.52, p = 0.041), and the total number of steps in the LSU: ICC = 0.988 (95% CI: 0.965 to 0.996; MD = 0.53 steps, 95% CI: –0.68 to 1.75, p = 0.364).

Conclusions: Remote evaluation of the FTSTS, TUG, LSU, and PBS demonstrates good-to-excellent agreement with in-person assessment. However, wide limits of agreement indicated individual variability.

Keywords

cerebral palsy, children, Five-Times-Sit-To-Stand, Lateral Step-Up, pediatric balance scale, remote assessment, Timed Up and Go

Introduction

Spasticity and musculoskeletal impairments limit balance and functional mobility in children with cerebral palsy (CP) [1, 2]. While over 60% of children with CP possess the ability to walk, the use of thorough functional assessments within the rehabilitation framework is essential for identifying specific physical impairments [1, 2]. In individuals with CP, conservative treatment focuses on mitigating the progression of musculoskeletal impairments, thereby promoting the preservation of independent mobility and walking. Among treatment options, goal-directed, functional, and gait training seem to be effective in improving the functionality of children and young adults with CP [3]. Thus, a detailed functional assessment using reliable tools is fundamental not only for establishing a physiotherapy program but also for determining the program's efficiency [4].

Telehealth interventions are at the forefront of solutions to make rehabilitation programs more accessible to children and their caregivers [5]. Research has established that utilizing remote assessment for evaluating functional capacity is both a practical and accurate alternative to in-person assessment in typically developing children [6]. Motor skills (e.g., motor coordination and postural stability) can be evaluated via remote supervision at home using simple functional tests such as the Timed Up and Go test (TUG) and one-legged standing balance test, providing valid information as in the in-person evaluation [6].

A recent systematic review revealed 23 mobile applications that are specially designed for children with CP. These applications are designed to provide information, evaluation, and treatment services in the field of the management of this population [7].

For assessing functional and lower limb capacity, as well as balance, practical tests such as the TUG, Lateral Step-Up test (LSU), the Five-Times-Sit-To-Stand test (FTSTS), and the Pediatric Balance Scale (PBS) are commonly used in pediatric patients with CP as these tests require short time, a small space, and can be easily performed at home [8–10]. Beyond their practical feasibility, these measures assess complementary domains of functional performance, including dynamic mobility (TUG), functional lower-limb strength (FTSTS), frontal-plane stability (LSU), and postural control (PBS). Together, they provide a multidimensional evaluation of activity-level performance and capture functional abilities relevant to everyday tasks such as walking, transfers, stair negotiation, and balance maintenance [11]. Their selection is further supported by evidence demonstrating their feasibility in pediatric rehabilitation settings [12, 13], while significant associations between TUG and PBS scores and Gross Motor Function Measure (GMFM) outcomes further support their clinical relevance for evaluating functional performance in children with CP [14].

In children and adolescents with CP, the aforementioned tests, when typically performed in-person, have excellent test-retest reliability [intraclass correlation coefficient, ICC = 0.92 (95% confidence interval, CI: 0.80 to 0.97) (LSU); ICC = 0.912 (95% CI: 0.821 to 0.956) (FTSTS); ICC = 0.958 (95% CI: 0.910 to 0.981) (PBS)] [4, 9, 15]. Furthermore, between raters, FTSTS and PBS have good to excellent reliability for the same population [ICC = 0.88 (95% CI: 0.758 to 0.941), and 0.901 (95% CI: 0.849 to 0.941), respectively] [9, 15]. The TUG has excellent inter- and intra-rater reliability [ICC = 0.992 (95% CI: 0.986 to 0.996)] in children and adolescents with autism spectrum disorder when administered in person [16]. Up to the present time, only a small selection of functional assessments has been adapted for remote use to evaluate functional capacity, core performance, and motor skills in both typically developing children and those with mobility impairments [6, 17]. The literature on remote assessment of functionality in pediatric populations with CP remains scarce.

Remote treatment procedures in pediatric populations are increasingly recognized as feasible and effective [18], but there is a major gap in the literature of evidence-based measurement and assessment procedures conducted remotely [18, 19]. Conducting remote assessments within a familiar home environment reduces children's anxiety and apprehension, fosters more natural functional performance, and facilitates active parental involvement in the evaluation process [20].

Therefore, the study aimed to examine the agreement between remotely supervised, parent-witnessed TUG, LSU, PBS, and FTSTS and the in-person assessment in a sample of children with CP. We hypothesized that the remote implementation of these tests, supervised by an expert pediatric physiotherapist, could produce similar results to the traditional assessment in this pediatric population.

Materials and methods

Participants

This prospective study was carried out between January and February 2022. A convenience sample of fifteen children with CP was recruited from a pool of patients of four private pediatric physiotherapy clinics in Greece. Children were eligible for the study when they met the following inclusion criteria: (1) age 6 to 12 years old, (2) diagnosed with unilateral or bilateral type of spastic CP, (3) functional mobility in categories I or II according to Gross Motor Function Classification System (GMFCS), (4) ability to understand verbal instructions. The exclusion criteria were: (1) orthopaedic surgery in the last year, (2) botulinumtoxin-A (Bont-A) injections within four months before the assessment, (3) categories III to IV on GMFCS, (4) difficulties in following verbal instructions, and (5) absence of electronic devices (laptop or smartphone). Parents of all participants were fully informed about the aims of the study, and they provided written informed consent. The study protocol followed the ethical standards of the World Medical Association (Declaration of Helsinki) and was approved by the Ethics Committee of the Physiotherapy Department, University of Thessaly (Protocol ID-719/23-9-2021).

Study design

To test the hypothesis, the study employed a within-subjects design in which children completed the assessments in two environments: remotely at home and in-person at the private physiotherapy clinic. These sessions were scheduled with 24 to 48 hours apart. For the remote evaluations, the VSee platform was utilized to facilitate real-time, two-way communication between children, parents, and the physiotherapist at home. VSee was selected for its high-resolution video quality, secure interface, and user-friendly accessibility, as it functions without requiring software installation on laptops or smartphones. To minimize potential learning effect, the functional tests order in both experimental conditions was randomized using the Random Sequence Generator software (https://www.random.org/sequences/). However, the order of experimental conditions remained the same for all participants, starting with the in-person assessment, followed 24 to 48 hours later by the remote assessment. To avoid fatigue, a resting period of 10 minutes between tests was provided in both experimental conditions. All children were instructed to avoid any vigorous activities (such as outdoor activities) or physiotherapy sessions between the two experimental conditions (remote and in-person). The study design is presented in Figure 1. To ensure thorough preparation and clarity, all parents were provided with comprehensive written guidelines. This documentation included instructions for equipment setup, navigation of the VSee platform, and explicit safety protocols for emergencies, complemented by instructional videos and illustrated guides for each specific test. Furthermore, supportive materials (ropes, meter tapes, goniometers, adhesive tapes, and a wooden step) were given to all children for the remote assessment. In both remote and in-person assessments, the same parent was present. All measurements were performed by the same physiotherapist (GK). The entire physiotherapy team possessed prior clinical expertise in conducting remote evaluations and delivering telerehabilitation services [21–23].

1-hour remote orientation session

One day before the remote evaluation, each participating parent attended a 1-hour virtual orientation with the same physiotherapist (GK). This session focused on organizing the home environment and mastering the VSee platform. The physiotherapist provided practical advice, such as optimizing internet stability by positioning devices near the router, and delivered a verbal walkthrough on the correct use of testing materials. Furthermore, the session dedicated a specific time to emergency preparedness, ensuring parents felt confident managing potential technical failures or safety concerns, such as a child losing their balance.

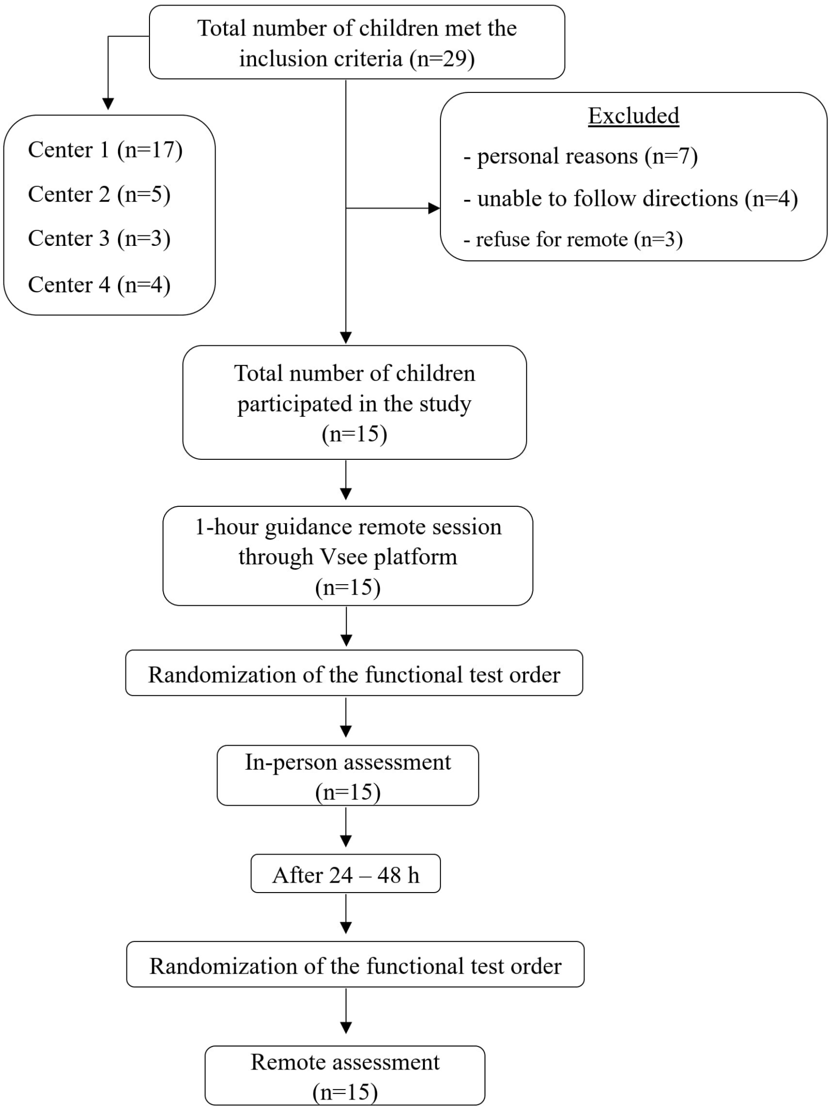

Figure 1. Flow chart of enrollment and completion of the study.

Assessment procedure

Prior to study enrollment, each child underwent a thorough clinical intake at a private physiotherapy center to document their medical history, anthropometric data (height and weight), and specific physical impairments related to CP. The functional evaluations were conducted across two consecutive days, with 24 to 48 hours apart, separating the in-person assessment and the remote one performed at the child's home. All children were asked to wear the same comfortable clothes and footwear for both experimental conditions. At the in-person assessment, children rested for 30–40 minutes before testing to avoid any fatigue due to their transportation to the physiotherapy clinic.

During the remote assessment, the children remained at home, maintaining no physical contact with the physiotherapist, who conducted the session from the clinic. Parents were only observed the assessment procedure and received strict instructions not to be involved unless an adverse event occurred. To ensure a stable and standardized viewing angle, all parents were supplied with a 50 × 50 cm wooden box to serve as a platform for their laptop or smartphone. Additionally, a wooden step that simulated a 17 cm stair step was provided to all participants for the remote assessment of the LSU. To conduct the TUG test, the laptop or smartphone was placed on the box at an angle, roughly 100 to 150 cm from the starting point. This specific positioning ensured the physiotherapist had a clear, unobstructed view of the entire 3-meter path, including the critical moments when the child initiated the movement from the chair and returned to a seated position. For the LSU, the laptop/smartphone was placed on the box, 100–150 cm away, at the back side of the child, and at an angle to fully capture the hip extension of the child while standing up and down from the wooden step. Similarly, for the FTSTS, the laptop/smartphone was placed beside each child to fully capture their performance during the five repetitions.

For PBS, the laptop/smartphone was placed at a distance that allowed the physiotherapist an unobstructed, full-body view of the child in both standing and seated postures.

Children were given a chance of two practice trials for each test before starting measurements. All the aforementioned functional tests were performed under the verbal guidance of the physiotherapist (GK) and in line with the recommended guidelines for each test [9, 10, 24, 25]

Adverse events

The physiotherapist was primed to record any adverse events for both the in-person and the remote assessment. Adverse events were defined as any physical or emotional incidents related to the assessment, including loss of balance resulting in a fall, dizziness, or significant distress. It was established that the occurrence of any such physical event would result in the immediate termination of the session. Additionally, the physiotherapist documented any factors that necessitated stopping a remote assessment, such as poor internet connectivity or a total loss of communication with the participant.

Measurements

Timed Up and Go test (TUG)

The TUG is a simple, valid, and reliable clinical test that is frequently used for assessing basic functional mobility in children with CP [24]. It involves consecutive activities such as getting up from a chair, walking 3 meters, turning 180 degrees, walking back to the chair, and sitting down. Timing begins when the child rises from the chair, walks 3 meters, turns around, walks back, and sits down again. The total time in seconds to complete the entire activity was measured by a stopwatch and was used for statistical analyses; longer duration is associated with decreased mobility [26].

Lateral Step-Up test (LSU)

LSU is a closed kinetic chain test that has been utilized to assess lower extremity muscular performance [10]. The test consists of 3, 30-second trials, where children are asked to perform as many steps as possible, stepping up and down on a 21cm (GMFCS I–II) or 12 cm (GMFCS III) step. For each trial, participants are instructed to stand with the more affected limb on the step and the less affected limb resting on the floor. Feet are parallel and shoulder-width apart. The hip and knee of the tested limb are moved into full extension, followed by flexion until the heel of the non-tested limb touches the floor. Steps are counted each time the heel of the resting limb touches the floor, without any support. In our study, the LSU was performed on the most affected limb to avoid fatigue and to include hemiplegic type of CP without causing further stress to participants. Furthermore, the research team elected to use a 17-cm step height, despite the recommendations outlined in the original testing protocol. This modification was implemented to better reflect the dimensions of steps commonly encountered in residential environments, thereby enhancing the ecological validity and applicability of the assessment within the home-based setting. During the remote evaluation, the supervising parent performed an independent step count to ensure accuracy. In the event of a discrepancy between the parent’s count and the physiotherapist’s, the physiotherapist’s recording was prioritized for the final data set. The best trial (number of steps) was included in the statistical analysis [10].

Five-Times-Sit-To-Stand test (FTSTS)

This test was performed using a height-adjustable chair without armrests, positioned against a wall. According to test guidelines, knee flexion while sitting should remain approximately at 90 degrees. To ensure this sitting position, parents adjusted the seat height by placing books. The structural stability, non-slip safety, and exact positioning of this modified chair setup were inspected and verified by the supervising physiotherapist during the 1-hour remote orientation session. For this test, children were tested barefoot on a firm mat. They were asked to keep their feet parallel, with the hip flexed at 90° and knee flexed at 105°, and fold their arms across their shoulders. They were instructed to fully stand up (complete knee extension) and sit down completely. On the command “go”, the time started counting, and the children performed five consecutive sit-to-stand repetitions as quickly as possible. The test was finished when the child returned to the seated position the fifth time. The time needed to complete the five repetitions in seconds was recorded and used in the statistical analysis [9].

Pediatric Balance Scale (PBS)

The PBS consists of 14 items that evaluate various activities essential for a child’s safe and independent participation in home, school, and community settings [25]. These items include sitting and standing balance, sit-to-stand, stand-to-sit, transfers, stepping, reaching forward, reaching to the floor, turning, and stepping on and off of an elevated surface. Each item is scored on a 5-point scale, and the total score ranges from 0 (severe balance impairment) to 56 (highest level of balance performance). It can be administered and scored in less than 20 minutes using equipment that can be easily found at a physiotherapy clinic or at a home setting (e.g., standard chair with armrests, measuring tape, or small lightweight object). The total score of the PBS was included in the statistical analysis. In the present study, the Greek adaptation of the PBS test for kinetic disorders of children aged 4 to 18 years old was used [27].

Statistical analysis

The normality of the data distribution was evaluated using the Shapiro-Wilk test. Descriptive statistics for categorical variables were presented as frequencies and percentages, while continuous variables were expressed as either means (with standard deviation, SD) or medians (with interquartile ranges, IQR), based on their distribution. The mean difference (MD) was expressed as in-person minus remote measurement. The required sample size was determined using the G*Power 3.1.9.4 software (University of Düsseldorf, Germany). Based on an “exact” bivariate normal model test to detect a moderate effect size (f = 0.7) with 80% power and a significance level of a = 0.05, a minimum of thirteen children was required. To account for a potential 20% attrition rate [28], common in studies of this nature, the final recruitment target was set at fifteen patients.

Differences between the two experimental conditions were calculated using the paired sample t-test. Statistical agreement for all continuous variables was evaluated using the ICC, specifically employing a two-way random-effects model to assess absolute agreement. Following the criteria established by Koo & Li, reliability was categorized as poor (< 0.5), moderate (0.5 to 0.75), good (0.75 to 0.9), or excellent (> 0.9) [29]. To further examine the level of agreement between the remote and in-person sessions and to detect possible outliers, Bland-Altman plots were generated [30]. Additionally, the standard error of measurement (SEM) was determined using the following formula to quantify the precision of the assessment [31]:

$$SEM = SD \times \sqrt{1 - ICC}$$

This value subsequently informed the calculation of the smallest detectable change (SDC95%), representing the threshold for true change at a 95% confidence interval (CI) [32]:

$$SDC95\% = SEM \times 1.96 \times \sqrt{2}$$

The SDC95% serves as the threshold for determining whether an observed difference reflects a “true” clinical improvement rather than a measurement error. As such, a lower SDC indicates a more sensitive assessment tool [31]. All statistical computations were conducted using SPSS for Mac, version 29.0 (SPSS Inc, Chicago, IL, USA), with the level of significance established at $p < 0.05$.

Results

While twenty-nine children from four physiotherapy clinics met the inclusion criteria, fourteen of those were subsequently excluded for reasons detailed in Figure 1. The clinical and demographic profiles of the remaining fifteen children who proceeded with the study are summarized in Table 1. A total of 9 children (60%) were boys, while 9 children (60%) were classified in category I at the GMFCS.

Table 1. Clinical and demographic characteristics of the fifteen children with cerebral palsy (*n* = 15).

Characteristic	Values
Weight (kg), mean (SD)	25.87 (1.78)
Height (cm), mean (SD)	125.87 (3.17)
Age (years), median (IQR: Q1–Q3)	7 (6–10)
GMFCS (category), median (IQR: Q1–Q3)	1(1–2)
Boys/girls, *n* (%)	9/6 (60/40)
GMFCS (category I), *n* (%)	9 (60)
GMFCS (category II), *n* (%)	6 (40)

Data are presented as mean (SD), median (IQR: Q1–Q3), numbers (*n*), and % percentage. GMFCS: Gross Motor Function Classification System; SD: standard deviation; IQR: interquartile range.

All children completed the trials of the TUG, LSU, FTSTS, and PBS, both in-person at the same physiotherapy clinic and remotely at home. During the measurements in both experimental conditions, no adverse bodily events or loss of internet signal were reported.

Excellent agreement between in-person and remote assessments was observed for the TUG (ICC = 0.998, 95% CI: 0.990 to 0.999), LSU (ICC = 0.988, 95% CI: 0.965 to 0.996), and PBS (ICC = 0.998, 95% CI: 0.992 to 0.999), while the FTSTS demonstrated good-to-excellent agreement (ICC = 0.884, 95% CI: 0.654 to 0.961).

The MD between the two experimental conditions, the SEM, and the SDC95% for the TUG, the LSU, the PBS, and the FTSTS are presented in Table 2. There were no differences between the mean in-person and the mean remote assessment scores for LSU ($p = 0.364$) and FTSTS ($p = 0.102$); however, TUG was significantly longer ($p = 0.017$) for the remote assessment than that of the in-person one. Additionally, the PBS score was significantly lower ($p = 0.041$) for the remote assessment than that of the in-person mode. For TUG and FTSTS tests, children scored higher for the remote measures compared with in-person values, indicating slower (in seconds) functional performance (Table 2).

Table 2. In-person and remote agreement of TUG, LSU, PBS, and FTSTS (*n* = 15).

Measurement	In-person assessment	Remote assessment	MD (95% CI)	*p*-value	SEM	SDC95%
TUG, time in sec	10.07 (7.70)	10.50 (8.16)	–0.42 (–0.76 to –0.08)	0.017*	0.35	0.98
LSU, number of steps	15.66 (9.88)	15.13 (10.16)	0.53 (–0.68 to 1.75)	0.364	1.10	3.05
PBS, total units	50.13 (5.57)	49.86 (5.82)	0.26 (0.01 to 0.52)	0.041*	0.25	0.69
FTSTS, time in sec	10.59 (4.32)	11.86 (4.80)	–1.26 (–2.81 to 0.28)	0.102	1.55	4.29

Data are presented as mean (SD). SD: standard deviation; MD: mean difference; SEM: standard error of measurement; SDC95%: smallest detectable change with 95% CI; TUG: Timed Up and Go test; LSU: Lateral Step-Up test; PBS: Pediatric Balance Scale; FTSTS: Five-Times-Sit-to-Stand test. *p*-value from paired sample *t*-test. *: $p < 0.05$.

The Bland-Altman plots revealed a good level of agreement between the in-person and the remote assessment for the total time in seconds of the FTSTS and the total number of steps of the LSU (Figure 2), identifying one child as an outlier on the FTSTS and one on the LSU. An excellent level of agreement was observed for the total time in seconds of the TUG and the total score of the PBS (Figure 3), identifying two children on the lower limit of agreement (TUG).

During the remote evaluation of the LSU, there were no significant differences ($p > 0.05$) between the step counts recorded by the physiotherapist and the parent.

Discussion

The study evaluated the level of agreement between the in-person and the home-based remote assessment of the TUG, LSU, FTSTS, and PBS specifically for children with CP. The findings demonstrate good to excellent agreement between in-person and remote evaluations, establishing that these four tests can produce comparable results for this population in a home setting and facilitating implementation of assessment procedures in a home setting when in-person assessment is not feasible.

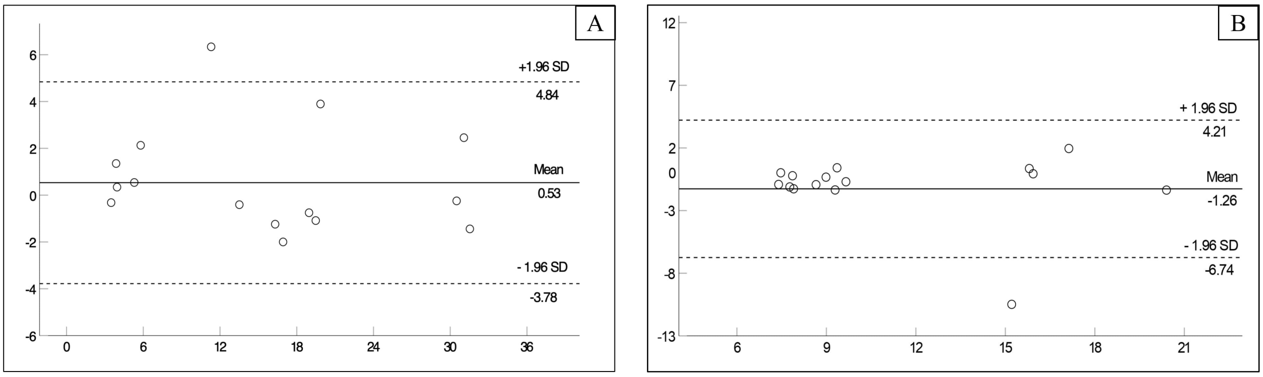

Figure 2. Bland-Altman plots for the total number of steps for LSU (A) and for the total time in seconds of the FTSTS (B). The x-axis represents the mean of the in-person and remote measurements for each participant, and the differences between scores are on the y-axis. The solid line represents the mean difference. The standard deviations of the paired differences were 2.19 and 2.79, respectively. The dotted lines represent the 95% limits of agreement, calculated as the mean difference ± 1.96 SD. Data were slightly jittered to prevent overplotting. SD: standard deviation. LSU: Lateral Step-Up test; FTSTS: Five-Times-Sit-To-Stand test.

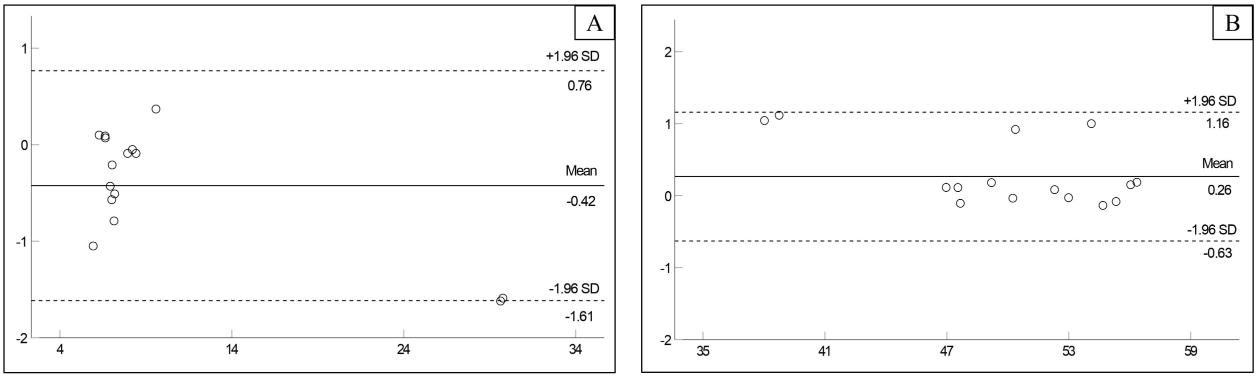

Figure 3. Bland-Altman plots for the total time in seconds of the TUG (A) and total PBS score (B). The x-axis represents the mean of the in-person and remote measurements for each participant, and the differences between scores are on the y-axis. The solid line represents the mean difference. The standard deviations of the paired differences were 0.60 and 0.45, respectively. The dotted lines represent the 95% limits of agreement, calculated as the mean difference ± 1.96 SD. Data were slightly jittered to prevent overplotting. SD: standard deviation. TUG: Timed Up and Go test; PBS: Pediatric Balance Scale.

Validity and reliability of the TUG and the FTSTS when these tests are performed in-person or remotely at the home setting have been previously assessed in the youth population with and without CP [17]. The study by Lai et al. [17] included 19 youths, 9 with CP aged 17.4 ± 1.9 years old, and 10 healthy ones, aged 19.3 ± 1.2 years old. The TUG and FTSTS tests were performed in-person in a laboratory, while the supervised remote performance was implemented using Zoom videoconferencing, in a specially designed space of the laboratory that resembled a living room at home. Both tests revealed excellent agreement [TUG: ICC = 0.92 (95% CI: 0.79 to 0.97); FTSTS: ICC = 0.95 (95% CI: 0.79 to 0.98)] between the two conditions [17]. These results are partially consistent with our results in terms of the TUG test. However, the FTSTS in our study revealed lower values of agreement, possibly explaining greater variability in young children's attention regulation and their slower information processing [33].

Functionality of children with CP may vary, often leading to impaired trunk control [34]. Trunk control plays a key role in posture, balance, and coordinated movements, enhancing gross and fine motor skills. Research examining the agreement of remotely administered PBS in children with CP remains limited. Reliability of the in-person and remote assessment of the trunk control has been previously assessed in children with CP aged 10 years old using the Trunk Control Measurement Scale (TCMS) [35]. The TCMS evaluates the static sitting balance, the dynamic sitting balance, and the dynamic reaching balance, giving a comprehensive overview of the child's ability to maintain posture and perform coordinated movements [36]. In the Üneş et al. study, the total score of the TCMS demonstrated excellent agreement [ICC = 0.91 (95% CI: 0.83 to 0.95)] between in-person and remote assessment [35]. Similarly, in our study, the PBS had excellent agreement between the two conditions [ICC = 0.998 (95% CI: 0.992 to 0.999)]. These results could be explained by similarities in the age group and comparable functional level, according to the GMFCS, in both studies.

Based on the available evidence, there is no study that investigates the agreement of the LSU between in-person and remote assessment in children with CP. A previous study in adult patients after chronic stroke investigated the reliability of performing the LSU in a laboratory setting [37]. The LSU demonstrated good to excellent test-retest reliability (ICC = 0.869 to 0.903) for the paretic and nonparetic legs, indicating consistent and reproducible measurements. Establishing good test-retest reliability is particularly crucial for remote assessment, as it ensures that the LSU, as a tool, provides consistent results in different settings. In our study, the LSU revealed excellent agreement [ICC = 0.988 (95% CI: 0.965 to 0.996)], giving similar results between in-person and remote performance in children with CP. Differences in the nature of the studies, the population's age, and the disease could describe the absence of further comparisons between the two studies.

When examining the absolute agreement indices, the observed MD values between in-person and remote assessments were small across all outcome measures, indicating limited differences between the two experimental conditions.

The Bland-Altman analyses further supported the agreement between modalities, as most observations were distributed around the MD and remained within the 95% limits of agreement. However, there were few outliers for the FTST test (one child) and the LSU (one child), indicating a notable between-individual variability.

For all tests, the SEM values showed a relatively low degree of measurement error across all outcome measures.

Among the assessed outcomes, the PBS demonstrated the lowest SEM and SDC95% values, whereas the FTSTS exhibited the highest values. Collectively, these findings indicate that the level of measurement error associated with remote assessment was modest and that the differences observed between assessment modalities were generally small. Therefore, remote administration of these outcome measures may provide results that are comparable to those obtained through in-person assessment.

Study limitations

There are several limitations to acknowledge in this study. Firstly, the small number of participants, the small age group (children from 6 to 12 years old), and their classification as GMFCS levels I and II, limit the generalizability of the findings in older children and in the broader pediatric CP populations, such as those with more severe limitations on motor skills (GMFCS levels III or higher). Although the initial power analysis showed almost fifteen children to be included in the study, functional, behavioral variations, and environmental factors may affect children's performance, and therefore, our findings should be considered as preliminary. It is important to note that because our primary objective was to determine the agreement between in-person and remote sessions, other forms of reliability, such as intra-rater or test-retest reliability within a specific setting, were not analyzed [38]. While this study focused on inter-modal consistency, test-retest reliability remains a vital metric for establishing the precision of individual measurements, which is necessary for monitoring clinical progress over time.

It is worth noting that in this study, parents acted strictly as passive witnesses to the assessment procedure and did not provide any physical assistance. However, even the purely observational presence of a parent in the home environment can psychologically influence the child's performance dynamics. This passive parental oversight may introduce assessment bias by subtly altering the child's motivation, compliance, or performance anxiety, which can consequently impact execution speed and balancing strategy compared to traditional clinic-based environments [39].

Despite the prior experience of the research team in conducting remote assessments utilizing functional tests among adult populations [22, 23], the administration of the tests by the same physiotherapist in both experimental conditions may have increased the risk of observer bias. However, this approach was intentionally adopted to reflect routine clinical practice, where the same clinician typically conducts repeated assessments and monitors patient progress over time.

An additional consideration when interpreting the present findings is the potential influence of a learning effect associated with repeated test administration. To minimize this possibility, the tests' order was randomized across participants in the two experimental conditions. Moreover, children did not demonstrate superior performance in the remote condition, suggesting that prior exposure to the tasks at the in-person assessment was unlikely to have substantially affected the results. Nevertheless, the potential contribution of familiarity with the assessment procedures should be considered in future studies examining agreement between in-person and remote testing modalities.

Furthermore, in our effort to minimize confounding factors of measurement variability, we standardized the home environment through parent training, equipment provisioning, and strict positioning guidelines. This rigorous experimental control was necessary to establish baseline agreement, as there were no previous data for this specific pediatric population. Therefore, our study does not truly reflect routine telehealth practice in a typical home setting, limiting the external validity. Overall, due to the absence of sensitivity analysis, the results of this study should be interpreted with caution as they may be influenced by individual variances.

Taking these limitations into account, future research should include higher sample sizes to assess these tests on different age and GMFCS level groups by incorporating test-retest and intra-rater reliability to obtain a broader perspective on their use in telehealth.

Conclusion

In conclusion, the preliminary findings demonstrate that the TUG, LSU, FTSTS, and PBS can produce similar results with the in-person assessment when performed remotely at home in children aged 6 to 12 years old with CP. The incorporation of these functional tests into remote assessment holds greater significance, emphasizing the opportunity of reaching more populations with limited access to in-person assessment due to multiple reasons. However, the results of this pilot study should be treated consciously, as at an individual level, there were some considerable variations. Clinicians wishing to implement these tests for remote assessment at home, when in-person assessment is not achievable, may need to consider greater standardization and provide training and facilitation to parents/caregivers. Future studies with larger sample sizes are needed to ensure the accuracy of these results.

Abbreviations

CI: confidence intervals

CP: cerebral palsy

FTSTS: Five-Times-Sit-To-Stand test

GMFCS: Gross Motor Function Classification System

ICC: intraclass correlation coefficient

IQR: interquartile range

LSU: Lateral Step-Up test

MD: mean difference

PBS: Pediatric Balance Scale

SD: standard deviation

SDC95%: smallest detectable change

SEM: standard error of measurement

TCMS: Trunk Control Measurement Scale

TUG: Timed Up and Go test

Declarations

Acknowledgments

The authors would like to thank all the children and their parents for their time and experiences.

Author contributions

AM: Data curation, Formal analysis, Methodology, Resources, Visualization, Writing—original draft. GK: Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Writing—original draft. PD: Visualization, Writing—original draft. EAK: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Resources, Supervision, Validation, Writing—review & editing. All authors read and approved the submitted version.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

The study protocol was approved by the Ethics Committee of the Physiotherapy Department, University of Thessaly (Protocol ID-719/23-9-2021).

Consent to participate

Parents of all participants were fully informed about the aims of the study, and they provided written informed consent.

Consent to publication

All parents of all participants were informed in advance, and consent was obtained for publication.

Availability of data and materials

Data from the study are available from the corresponding author upon reasonable request.

Funding

The authors reported that there is no funding associated with the work featured in this article.

Copyright

Publisher's note

References

1. Morgan P, McGinley JL. Cerebral palsy. Handb Clin Neurol. 2018;159:323–36. [DOI] [PubMed]
2. Cetin SY, Erel S. Investigation of the validity and reliability of the L test in children with cerebral palsy. Physiother Theory Pract. 2020;38:182–8. [DOI] [PubMed]
3. Das SP, Ganesh GS. Evidence-based Approach to Physical Therapy in Cerebral Palsy. Indian J Orthop. 2019;53:20–34. [DOI] [PubMed] [PMC]
4. VAN Tittelboom V, Alemdaroğlu-Gürbüz I, Hanssen B, Plasschaert F, Heyrman L, Feys H, et al. Reliability of functional tests of the lower limbs and core stability in children and adolescents with cerebral palsy. Eur J Phys Rehabil Med. 2021;57:738–46. [DOI] [PubMed]

5. Zheng J, Hou M, Liu L, Wang X. Knowledge Structure and Emerging Trends of Telerehabilitation in Recent 20 Years: A Bibliometric Analysis via CiteSpace. Front Public Health. 2022;10:904855. [DOI] [PubMed] [PMC]
6. Button AM, Webster EK, Kracht CL, Hendrick C, Okely A, Chong KH, et al. Validation of remote assessment of preschool children's anthropometrics and motor skills. Front Digit Health. 2023;5: 1168618. [DOI] [PubMed] [PMC]
7. Rodríguez Mariblanca M, Cano de la Cuerda R. Aplicaciones móviles en la parálisis cerebral infantil. Neurología. 2021;36:135–48. [DOI] [PubMed]
8. 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;55:988–99. [DOI] [PubMed]
9. Kumban W, Amatachaya S, Emasithi A, Siritaratiwat W. Five-times-sit-to-stand test in children with cerebral palsy: Reliability and concurrent validity. NeuroRehabilit: Int Interdiscip J. 2013;32:9–15. [DOI] [PubMed]
10. Chrysagis N, Skordilis EK, Tsiganos G, Koutsouki D. Validity evidence of the Lateral Step Up (LSU) test for adolescents with spastic cerebral palsy. Disabil Rehabil. 2012;35:875–80. [DOI] [PubMed]
11. Romin E, Lindgren A, Rodby-Bousquet E, Cloodt E. Sit-to-stand performance in children with cerebral palsy: a population-based cross-sectional study. BMC Musculoskelet Disord. 2024;25:460. [DOI] [PubMed] [PMC]
12. Nicolini-Panisson RD, Donadio MV. Normative values for the Timed 'Up and Go' test in children and adolescents and validation for individuals with Down syndrome. Dev Med Child Neurol. 2013;56: 490–7. [DOI] [PubMed]
13. Alsakhawi RS, Elshafey MA. Effect of Core Stability Exercises and Treadmill Training on Balance in Children with Down Syndrome: Randomized Controlled Trial. Adv Ther. 2019;36:2364–73. [DOI] [PubMed]
14. Kembhavi G, Darrah J, Magill-Evans J, Loomis J. Using the Berg Balance Scale to Distinguish Balance Abilities in Children with Cerebral Palsy. Pediatr Phys Ther. 2002;14:92–9. [DOI] [PubMed]
15. Her J, Woo J, Ko J. Reliability of the Pediatric Balance Scale in the Assessment of the Children with Cerebral Palsy. J Phys Ther Sci. 2012;24:301–5. [DOI]
16. Martín-Díaz P, Carratalá-Tejada M, Molina-Rueda F, Cuesta-Gómez A. Reliability and agreement of the timed up and go test in children and teenagers with autism spectrum disorder. Eur J Pediatr. 2023; 182:3577–85. [DOI] [PubMed] [PMC]
17. Lai B, Wadsworth D, Spring K, Jones CS, Mintz M, Malone LA, et al. Validity and Reliability of a Telehealth Physical Fitness and Functional Assessment Battery for Ambulatory Youth With and Without Mobility Disabilities: Observational Measurement Study. JMIR Rehabil Assist Technol. 2024; 11:e50582. [DOI] [PubMed] [PMC]
18. Ros-DeMarize R, Chung P, Stewart R. Pediatric behavioral telehealth in the age of COVID-19: Brief evidence review and practice considerations. Curr Probl Pediatr Adolesc Health Care. 2021;51: 100949. [DOI] [PubMed] [PMC]
19. Lima CRG, Verdério BN, de Abreu RWF, Brugnaro BH, Dos Santos AN, Dos Santos MM, et al. Telemonitoring of motor skills using the Alberta Infant Motor Scale for at-risk infants in the first year of life. J Telemed Telecare. 2022;30:885–94. [DOI] [PubMed]
20. Ogourtsova T, Boychuck Z, O'Donnell M, Ahmed S, Osman G, Majnemer A. Telerehabilitation for Children and Youth with Developmental Disabilities and Their Families: A Systematic Review. Phys Occup Ther Pediatr. 2022;43:129–75. [DOI] [PubMed]
21. Kortianou E, Tsimouris D, Mavronasou A, Lekkas S, Kazatzis N, Apostolara Z, et al. Application of a home-based exercise program combined with tele-rehabilitation in previously hospitalized patients with COVID-19: A feasibility, single-cohort interventional study. Pneumon. 2022;35:1–10. [DOI]

22. Mavronasou A, Asimakos A, Vasilopoulos A, Katsaounou P, Kortianou EA. Remote administration of the short physical performance battery, the 1-minute sit to stand, and the Chester step test in post-COVID-19 patients after hospitalization: establishing inter-reliability and agreement with the face-to-face assessment. Disabil Rehabil. 2023;46:5334–44. [DOI] [PubMed]
23. Panagiotopoulou E, Mavronasou A, Spanos S, Kortianou EA. Tele-assessment of Berg balance scale, timed-up, go and box and block test in patients with chronic stroke: establishing agreement with the face-to-face assessment. Int J Clin Trials. 2024;11:84–90. [DOI]
24. Dhote SN, Khatri PA, Ganvir SS. Reliability of "Modified timed up and go" test in children with cerebral palsy. J Pediatr Neurosci. 2012;7:96–100. [DOI] [PubMed] [PMC]
25. Yi SH, Hwang JH, Kim SJ, Kwon JY. Validity of Pediatric Balance Scales in Children with Spastic Cerebral Palsy. Neuropediatrics. 2012;43:307–13. [DOI] [PubMed]
26. Eagles D, Perry JJ, Sirois MJ, Lang E, Daoust R, Lee J, et al. Timed Up and Go predicts functional decline in older patients presenting to the emergency department following minor trauma. Age Ageing. 2016; 46:214–8. [DOI] [PubMed] [PMC]
27. Laspa V, Besios T, Xristara A, Tsigaras G, Milioudi M, Mauromoustakos S, et al. Reliability and Clinical Significance of the Pediatric Balance Scale (PBS) in the Greek Language in Children Aged 4 to 18 Years. Open J Prev Med. 2020;10:73–81. [DOI]
28. Güngör F, Ovacık U, Ertan Harputlu Ö, Yekdaneh AA, Kurt İ, Ertürk Uzunoğlu G, et al. Tele-assessment of core performance and functional capacity: Reliability, validity, and feasibility in healthy individuals. J Telemed Telecare. 2022;30:1017–25. [DOI] [PubMed]
29. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15:155–63. [DOI] [PubMed] [PMC]
30. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8:135–60. [DOI] [PubMed]
31. Weir JP. Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. J Strength Cond Res. 2005;19:231–40. [DOI] [PubMed]
32. Lu WS, Wang CH, Lin JH, Sheu CF, Hsieh CL. The minimal detectable change of the simplified stroke rehabilitation assessment of movement measure. J Rehabil Med. 2008;40:615–9. [DOI] [PubMed]
33. Bottcher L, Flachs EM, Uldall P. Attentional and executive impairments in children with spastic cerebral palsy. Dev Med Child Neurol. 2010;52:e42–7. [DOI] [PubMed]
34. Sadowska M, Sarecka-Hujar B, Kopyta I. Cerebral Palsy: Current Opinions on Definition, Epidemiology, Risk Factors, Classification and Treatment Options. Neuropsychiatr Dis Treat. 2020;16:1505–18. [DOI] [PubMed] [PMC]
35. Üneş S, Tunçdemir M, Özal C, Delioğlu K, Seyhan Bıyık K, Kerem Günel M. Tele-assessment of trunk control in children with cerebral palsy: Intra- and inter-rater reliability, and validity of the trunk control measurement scale. J Telemed Telecare. 2025;32:266–74. [DOI] [PubMed]
36. Heyrman L, Molenaers G, Desloovere K, Verheyden G, De Cat J, Monbaliu E, et al. A clinical tool to measure trunk control in children with cerebral palsy: The Trunk Control Measurement Scale. Res Dev Disabil. 2011;32:2624–35. [DOI] [PubMed]
37. Kwong PWH, Ng SSM. Reliability of the Lateral Step-Up Test and Its Correlation with Motor Function and Activity in Chronic Stroke Survivors. BioMed Res Int. 2020;2020:7859391. [DOI] [PubMed] [PMC]
38. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The Case for Using the Repeatability Coefficient When Calculating Test–Retest Reliability. PLoS ONE. 2013;8:e73990. [DOI] [PubMed] [PMC]
39. Rebold MJ, Kobak MS, Croall CA, Cumberledge EA, Hall M, Raumikaitis L. The Experimental Effect of Parental Attentiveness on Children's Physical Activity Behavior. Asian J Kinesiol. 2021;23:12–20. [DOI]