

Original Article

Reliability and Validity of Gait Analysis Using Mechanography in Pediatric Patients with CP

Andriani Bougadaki^{1,2}, Ibrahim Duran^{1,3}, Eckhard Schoenau^{1,3}, Kyriakos Martakis^{3,4,5}

¹Center of Prevention and Rehabilitation, Medical Faculty and University Hospital, University of Cologne, Cologne, Germany;

²Center for Social Pediatrics and Child Development, Lebenszentrum-Königsborn, Unna, Germany;

³Department of Pediatrics, Medical Faculty and University Hospital, University of Cologne, Cologne, Germany;

⁴Department of Pediatric Neurology, University Children's Hospital (UKGM) and Medical Faculty, Justus Liebig University of Giessen, Giessen, Germany;

⁵Third Department of Paediatrics, 'Attikon' University General Hospital, Chaidari, Attica, Greece; Medical School, National and Kapodistrian University of Athens, Athens, Attica, Greece

Abstract

Objectives: To determine the reliability and construct validity of the mechanographic gait analysis (Leonardo Mechanograph®-Gangway) in pediatric cerebral palsy (CP) patients compared to standardized clinical tests for the evaluation of the gross motor skills (GMFM-66, 1MWT, 6MWT). **Methods:** Retrospective study of 846 CP patients (GMFCS I-III, mean age 7.5 years, female 42.6%), examined between May 2007 and January 2020. Construct validity (Spearman's rho) and within-session repeatability (ICC, SEM, MDC across the best three consecutive trials selected for analysis) were assessed for the Gangway parameters: "average step length," "average horizontal velocity" and "pathlength/distance". **Results:** All parameters correlated most strongly with GMFM-66. "Average horizontal velocity" showed the strongest validity overall, reaching its highest correlation with GMFM-66 (rho 0.737, CI:0.715-0.757), but poor reliability (ICC 0.477, CI: 0.445-0.509). "Average step length" showed moderate-to-strong correlation with GMFM-66 (rho 0.636, CI:0.609-0.662) and strong consistency (ICC 0.827, CI:0.813-0.841). **Conclusions:** Three consecutive trials, conducted under identical conditions on Gangway, considering "average step length" provided reproducible and valid data in pediatric CP patients, whereas "average horizontal velocity" demonstrated construct validity, but should be evaluated in conjunction with more mechanically stable parameters, because of its high natural variance. Further research is needed to explore broader clinical applications of mechanographical parameters and across different age groups.

Keywords: Cerebral Palsy, Gait Analysis, Mechanography, Reliability, Validity

Introduction

Cerebral palsy (CP) outlines a group of irreversible, but non-progressive neuromuscular disorders characterized by abnormal muscle tone and posture, so as restrictions of the locomotor system, that are in many cases accompanied by sensory, perceptual, cognitive and behavioral

disturbances as well as dysfunctional neuropsychological development¹⁻³. CP is divided according to the leading clinical symptoms in spastic hemiplegia, spastic diplegia, ataxic, dyskinetic and mixed type^{2,3}. The birth prevalence of CP in high-income-countries is 1.6 per 1000 live births and in low- and middle-income countries remarkably higher^{3,4}. Although CP is not a very common condition, it is the primary cause of early-onset physical impairment^{5,6}.

The medical treatment of patients with CP focuses primarily on optimizing the quality of life, improving the motor skills and preventing the complications due to immobilization, especially in view of lack of a cure. A very important parameter with the aim of improving the patient's everyday life is the evaluation of their gait pattern. In this regard, the motor function is evaluated not only with the method of subjective inspection of the patient's gait pattern used since antiquity, but also by means of

The authors have no conflict of interest.

Corresponding author: Andriani Bougadaki, Address: UniReha – Zentrum für Kinder- und Jugendrehabilitation, Lindener Allee 44, 50931 Köln, Germany

E-mail: andriani.bougadaki@gmail.com

Edited by: G. Lyritis

Accepted 30 July 2026



objective, instrumental procedures and methods, which enable quantitative results for monitoring the progress of the therapies initiated. Over the last decades, several special gait analysis systems have been developed in view of this clinical need. Mechanography is a methodology that can be used for gait analysis. It refers to a medical diagnostic measurement method for motion analysis and assessment of muscle function by means of physical parameters. By utilizing a ground reaction force plate (GRFP), for example Gangway, to record force dynamics, this method quantifies functional movements associated with motion patterns of typical everyday movements. From these ground reaction forces center of gravity related physical parameters, like relative maximum forces and velocity are calculated. Mechanography is established as a safe, reliable, and accessible instrument for assessing lower extremity muscle function⁷. Furthermore, another measurement methodology available, than the use of force plates (ground reaction forces) is by using ground pressure sensors. Ground pressure assessment is performed using a high-resolution sensor system, located under the foot, to quantify the plantar pressure distribution and contact area during the stance phase⁸.

Gait analysis has become a subject of intensive research over the past few years. According to Veillieux et al. in 2011, gait analysis using Gangway showed high reliability (average velocity ICC 0.97, average step length ICC 0.99) in healthy young adults (aged 21 to 45 years) in within-session trials, comparing walking with shoes to barefoot walking and exhibiting a five percent greater step length⁹. Furthermore, another study by Veillieux et al. in 2014 demonstrated good reliability of Gangway (average horizontal velocity ICC 0.98, average step length ICC 0.99) in gait analysis in healthy children and adolescents aged 7 to 17 years in within-day trials¹⁰. Another research team, Wiegman et al., 2019, has provided evidence of Gangway as a valid tool for balance measurements¹¹.

Additionally, the analysis of spatial and temporal parameters of gait in healthy children and adults using other walkways has already been researched by Stolze et al. in 1998 and a high reliability of this type of gait analysis has been shown in healthy adults. The reliability of spatial gait analysis was shown in the same study to be better in healthy adults than in healthy children¹². Moreover, spatiotemporal characteristics of gait have been evaluated by Dussault-Picard et al. in 2025 using different floor panels. This study demonstrated the gait changes (such as increased mean stride width, gait speed variability, and walk ratio variability) of children with CP and their typically developing peers when walking on an uneven surface¹³.

However, as of the current time and to the best of our knowledge, there is no published data regarding the reliability of different gait analysis' parameters with ground reaction force measurements, such as those obtained by Gangway, in children and adolescents with CP. According to our literature review, there are no published

scientific papers regarding the validity of gait analysis' parameters using Gangway for either healthy children and adolescents or those suffering from CP. Hence, the aim of this study was to investigate the reliability and validity of the gait analysis using Gangway in children and adolescents with CP.

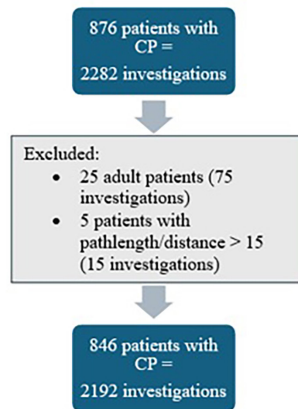
Materials and Methods

Study population

The study population entailed children and adolescents with CP, who participated in the rehabilitation program "Auf die Beine" of the Center of Prevention and Rehabilitation of University of Cologne in Germany, between May 2007 and January 2020 and had the cognitive ability to understand instructions and the walking capacity to complete the gait analysis. The age of the patients ranged from 2 to 18 years old. Participants who reached adulthood (>18 years of age) at any point during their time in the registry were entirely excluded from all analyses, including their earlier pediatric measurements, to ensure a strictly pediatric cohort. A formal a priori sample size calculation was not performed. To maximize statistical precision, we included all consecutive, eligible patient records, that met our inclusion criteria, yielding a final sample size of n=846. The rehabilitation concept "Auf die Beine" (meaning "On your feet") has already been thoroughly presented in a previous publication¹⁴. We conducted a retrospective analysis of database material from patients with Gross Motor Function Classification System (GMFCS) Level I-III. The GMFCS is one of the most accessible and commonly used tools to assess the severity of motor dysfunction in terms of CP. It was developed by Palisano et al¹⁵. This is a validated five level ordinal scale based on the functional independence of a patient with CP regarding basic motor activities and the need for accessory equipment (crutches, walking frames, wheelchairs). The GMFCS is adapted in accordance with the patient's age (younger than 2 years, 2–4 years, 4–6 years, 6–12 years and 12–18 years old)². The GMFCS lacks though in assessing the quality of the movement. Patients with GMFCS level I and II can walk independently or with certain little limitations respectively, but slower and less coordinated than healthy individuals of the same age. At level III the patients can walk only using accessory equipment and at level IV only with the use of powered devices. Patients with GMFCS level V can only be transported by a care provider and lack head control^{2,16}.

To ensure clarity, a "trial" is defined as one single walk along the Gangway. An "investigation" refers to all measurements (Gangway variables and clinical scores) collected during a single trial. During data collection, patients performed a variable number of trials (a minimum of three, but often more, depending on clinical compliance). While the total number of walk attempts was not systematically recorded for every patient, a "session"

Validity-Calculations



Reliability-Calculations

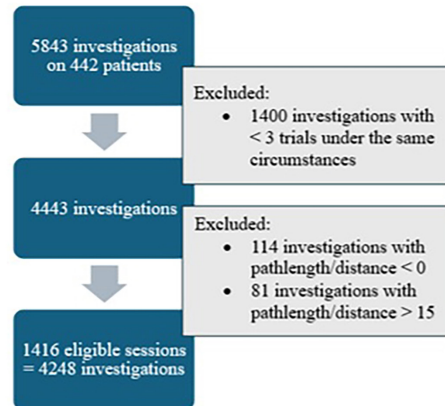


Figure 1. Flow charts on sample selection and criteria. 1416 eligible reliability sessions from 442 participants generated 4248 investigations.

for data analysis was strictly defined as the best three consecutive trials (identified as the continuous sequence of three trials yielding the lowest pathlength/distance values). These three consecutive trials were always conducted under identical circumstances. In total, for the validity analysis, 2282 investigations were performed on 876 patients with CP (GMFCS I-III) in the studied period (Figure 1). Only investigations in children and adolescents with plausible values were further evaluated (846 patients, corresponding to 2192 investigations). By plausible values we refer to the range of values a population parameter could reasonably take, given the sample data. For example, the parameter pathlength/distance could not take negative values or a value of more than 15 could not be reasonably corresponding to reality, so that those values were regarded as unrealistic and probably wrong and were excluded from the analysis. Those values occurred primarily in severely affected patients (GMFCS III), where minimal forward progression combined with lateral sway leads to an exponentially inflated, biomechanically impossible ratio. To remove outliers from our sample we used Tukey's interquartile range method, which automatically excluded pathlength/distance values more than 4.12 for the validity analysis and more than 2.2 for the reliability analysis. To be more specific, in the validity analysis 38 investigations for step length, 6 for velocity and 161 for pathlength/distance were excluded. In the reliability analysis: 50, 98 and 1118 respectively. To avoid sample bias, we conducted sensitivity analyses with and without Tukey outlier exclusion, so as to evaluate the robustness of our findings

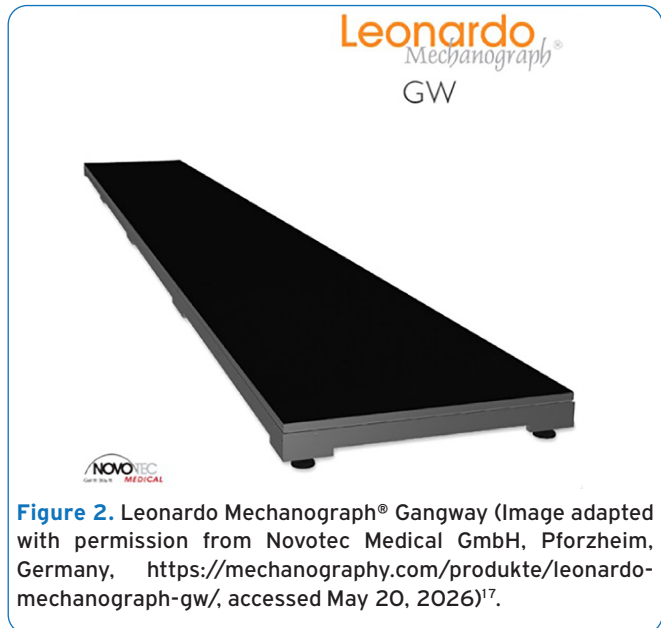


Figure 2. Leonardo Mechanograph® Gangway (Image adapted with permission from Novotec Medical GmbH, Pforzheim, Germany. <https://mechanography.com/produkte/leonardo-mechanograph-gw/>, accessed May 20, 2026)¹⁷.

and to assess the impact of these extreme values -which may represent either technical measurement artifacts or true, severe pathological gait deviations typical in this patient cohort- on the overall validity and reliability of the system. Regarding the reproducibility analysis, in total 5843 investigations on 442 patients were

Table 1. Number of sessions per participant involved in the reliability analysis.

Number of sessions per participant involved in the reliability analysis	Number of participants	Percentage (%)
1	97	21.9
2	109	24.7
3	82	18.6
4	57	12.9
5	38	8.6
6	25	5.7
7	9	2
8	10	2.3
9	9	2
10	4	0.9
11	1	0.2
12	1	0.2
Total	442	100

Table 2. Participants' demographics (Validity group).

	Validity group
Participants, N	846
Female, %	42.6%
Age, mean in years (standard deviation)	7.5 (3.67)
Height, mean in cm (standard deviation)	118.6 (20.89)
BMI, mean in kg/m² (standard deviation)	16.09 (3.01)
GMFCS, %	
GMFCS I	12.6
GMFCS II	28.7
GMFCS III	58.6
CP Subtype, %	
tetraplegic spastic (G80.0)	51.2
bilateral spastic (G80.1)	24.7
unilateral (G80.2)	10.4
dyskinetic (G80.3)	2.6
ataxic (G80.4)	3.1
mixed type (G80.8)	3.4
not classified (G80.9)	4.6

Table 3. Participants' demographics (Reliability group)

	Reliability group
Participants, N	442
Female, %	42.1%
Age, mean in years (standard deviation)	6.95 (3.32)
Height, mean in cm (standard deviation)	115.1 (19.45)
BMI, mean in kg/m² (standard deviation)	15.74 (2.73)
GMFCS, %	
GMFCS I	10.2
GMFCS II	29.4
GMFCS III	60.4
CP Subtype, %	
tetraplegic spastic (G80.0)	52
bilateral spastic (G80.1)	22.6
unilateral (G80.2)	10.2
dyskinetic (G80.3)	2.9
ataxic (G80.4)	2.7
mixed type (G80.8)	5.2
not classified (G80.9)	4.3

overall in the period studied, performed. After applying the additional exclusion criteria (1400 investigations without three under the same circumstances, consecutive conducted trials, 195 investigations with implausible results) 4248 (3x1416) investigations were eligible for

the reproducibility evaluation (Figure 1). In the analysis 97 (21.9%) participants were involved once, 109 (24.7%) twice, 82 (18.6%) three times and 57 (12.9%) four times (Table 1). Participants' demographics are given in Tables 2 and 3.

Table 4. Abbreviations and definitions.

Abbreviations and definitions	
CI	Confidence interval
CP	Cerebral palsy
GMFCS	Gross motor function classification system
GMFM-66	Gross motor function measure-66
GRFP	Ground reaction force platform
GW	Gangway
1MWT	One-Minute walk test
6MWT	Six-Minute walk test
ICC	Intraclass correlations coefficient
MDC	Minimum detectable change
SEM	Standard error of measurement
Average step length (in meters)	mean step length
Average velocity horizontal (in m/s)	mean velocity in the direction of the length of the GW
Pathlength/Distance (dimensionless ratio)	pathlength in relation to the walked distance in the direction of the length of the GW

Variables, investigations and GRFP

All trials were conducted using “Leonardo Mechanograph® GW” (Gangway) (Novotec Medical GmbH, Pforzheim, Germany), a 6-meter-long and 70-centimeter-wide platform (Figure 2), which consists of 4 x 1.5 m measure-modules, each measuring the ground reaction forces via four force sensors (in total 16 sensors) with a sample rate of 800 Hz. Data processing was performed using the software Leonardo GRFP Mechanography® v.4.4b03.58 (Novotec Medical GmbH, Pforzheim, Germany)¹⁷.

The patients were positioned on the Gangway and were instructed to walk in their own way and at their own pace, at least three times along the Gangway. Casual footwear and -when needed- orthoses or minimal physical support, such as intermittent hold from one or both arms, from the accompanying person were allowed. At the beginning as well as at the end of every walking trial, the patient’s weight was measured by the Gangway. Detailed description of the conduction of the investigations in our center has been presented in a previous publication⁶. The system enables simple and rapid gait analysis as the patient walks along the Gangway, evaluating ground reaction forces, captured with spatial resolution. It automatically analyzes typical gait analysis parameters such as stride length, asymmetries, gait regularity, peak forces, and path length. In our study only the parameters “average step length,” “average horizontal velocity” and “pathlength/distance” were evaluated, that are in clinical use easier comprehensible. The study variables are summarized in Table 4. Investigations were excluded from evaluation, if not all three walking trials were conducted under the same circumstances (e.g. one trial with

orthoses, two without) or if the results were not plausible, as explained above (Figure 1). All paired assessments (Gangway and the corresponding clinical tests: GMFM-66, 1MWT, and 6MWT) were performed within five days during the exact same rehabilitation visit. Thus, the risk of confounding due to long-term developmental changes or external interventions is effectively eliminated. The clinical assessments (GMFM-66, 1MWT, and 6MWT) were conducted by physiotherapists who received standardized training in accordance with our rehabilitation center’s standard operating protocols. To prevent bias, the physiotherapists were blinded to the Gangway results. Furthermore, the clinical assessments were always performed prior to the Gangway measurements. The Gangway measurements were conducted by medical technical assistants or sports scientists, who were also trained according to our standardized protocol. Routine safety monitoring was performed by the respective rater during all walking trials on the Gangway.

Anthropometry

The patients’ height was measured in 0.1 cm increments using the stadiometer Seca 213 (SECA GmbH, Hamburg, Germany). The weight was measured by the GRFP in 0.1 kg increments.

Reliability and validity

The reproducibility of the Gangway has been calculated by evaluating three walking trials, conducted under standardized conditions in rapid succession (i.e., within a few minutes).

To evaluate the construct validity of the Gangway, the representative trials for each investigation were

identified as the ones exhibiting the minimum “pathlength/distance”-value, as it reflects the most efficient and stable gait pattern achieved by the participant. This trial was subsequently used to determine the correlation coefficients. Our construct validity analysis used trials as the unit of analysis. This method follows the established clinical practice for the gait analysis using mechanography, where the “best-of-three” approach and not -for example- the mean value of three approaches, is routinely utilized. For the correlations, the results of the device were compared with the score results of Gross Motor Function Measure 66 (GMFM-66), One-Minute Walk Test (1MWT) and Six-Minute Walk Test (6MWT) of the pediatric patients with CP. The GMFM-66 consists of 66 motor tasks and is a standardized clinical measurement method, updated in 2000, for assessing gross motor function in children and adolescents with CP^{15,18,19}. Avery et al. suggested that the examination of 13 items is adequate for the GMFM-66. Nevertheless, they highlight that the GMFM-66 assessment is more accurate when a greater quantity of items is evaluated²⁰. Duran et al. generated reference centiles for the GMFM to objectively monitor a child's gross motor development over time and allow to separate this natural development from any extra benefit gained specifically from a new or added intervention²¹. The One-Minute Walk Test (1MWT) is a well-established, accurate and easy-to-perform assessment tool, frequently used for evaluating walking capacity in children with CP²². Furthermore there are 1MWT reference centiles, a specialized clinical tool to track and evaluate changes in the walking ability of children with CP and identify therapeutic effects, generated in the study of Martakis et al. in 2019. The Six-Minute Walk Test (6MWT) is a reliable test for pediatric ambulant patients with CP^{23,24}. For this clinical test as well, reference centiles were generated from Martakis et al., that enable to trace changes in the walking capacity of children with CP and differentiate between expected baseline changes and treatment effects²⁴. GMFM-66, 1MWT, and 6MWT assessments were conducted from physiotherapists, working in our rehabilitation center with experience in children with CP, blinded to Gangway results.

Reliability was assessed as within-session, trial-to-trial repeatability. For each patient, the best three consecutive trials (selected based on the lowest pathlength/distance value) constituted a session. Intraclass Correlation Coefficients (ICCs) as well as measurement error metrics (SEM, MDC) were calculated across these three consecutive trials to evaluate intra-session consistency. The number of sessions per participant that were involved in the reliability analysis is stated in Table 1.

Statistical Analysis

The statistical analysis was conducted using IBM SPSS Statistics Version 29.0.1.1, 2023 (IBM Corp, Armonk, NY, USA).

After refining the sample using the exclusion criteria described above (Figure 1), data preprocessing began with the identification and exclusion of outliers then based on Tukey's interquartile range method. According to this method, data points falling outside of 1.5xIQR below the first quartile (Q1) or above the third quartile (Q3) were considered outliers and subsequently excluded from further analysis²⁵. Then, the Kolmogorov-Smirnov-Test was used to define if our dataset was normally distributed. As a result, it was shown that the values did not have a normal distribution ($p < 0.001$). Therefore, Spearman's correlation coefficient was performed to assess the findings of our investigations on the selected Gangway parameters: “average step length,” “average horizontal velocity” and “pathlength/distance” and the clinical scores: GMFM-66, 1MWT and 6MWT. Average step length was normalized to body height to account for the high variance in anthropometrics across our pediatric cohort. This standardization allowed for a valid comparison of gait mechanics between children of different ages and sizes. A stratification by height was also performed to verify that the system's construct validity remains robust across the different gait dynamics and step sizes of the pediatric spectrum. To evaluate whether the correlations remain meaningful across varying degrees of motor impairment and severe gait deviations we also performed a stratification analysis by GMFCS level. Taking into consideration that multiple trials per child were included in the analyses, we conducted a sensitivity analysis using the first available trial (for the validity assessment) and the first available session (for the reliability assessment) per participant. For the interpretation of the correlation coefficient there are no strict guidelines. A rho of 0.5-0.8 in medical research is often used to indicate a significant correlation between two variables¹⁶. We used the conventional approach to interpreting a correlation coefficient suggested by Schober et al., according to which a rho of 0.70–0.89 indicate a strong correlation²⁶.

Within-session repeatability refers to the consistency of investigations taken by the same rater (observer or instrument) on multiple occasions. Unfortunately, there are no standard values for acceptable reliability using ICC. We conducted investigations at a sample of 442 individuals, involving 3 gait trials at the same time under the same circumstances (in total 4248 investigations). Under such conditions according to the publication of Koo et al., it is suggested that ICC values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability²⁷. To quantify absolute reliability and measurement error, the Standard Error of Measurement (SEM) and the Minimum Detectable Change (MDC) were calculated. The SEM provides an estimate of the measurement precision in the same units as the original parameter and was calculated using the sample's standard deviation (SD) and the Intraclass Correlation

Table 5. Validity of the Gangway parameters evaluated in this study. $p < 0.001$ in all variables.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
GMFM-66	0.636 (0.609-0.662) N=1991	0.643 (0.616-0.669) N=1991	0.737 (0.715-0.757) N=1991	-0.525 (-0.557- -0.492) N=1991
1MWT	0.394 (0.348-0.439) N=1410	0.400 (0.354-0.444) N=1410	0.476 (0.433-0.517) N=1410	-0.300 (-0.348- -0.251) N=1410
6MWT	0.429 (0.384-0.473) N=1371	0.407 (0.360- 0.451) N=1371	0.505 (0.463-0.544) N=1371	-0.415 (-0.459- -0.368) N=1371

Table 6. Validity of the Gangway parameters evaluated in this study, including Tukey's outliers. $p < 0.001$ in all variables.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
GMFM-66	0.665 (0.640-0.688) N=2192	0.664 (0.639-0.687) N=2191	0.752 (0.733-0.770) N=2192	-0.575 (-0.603- -0.445) N=2192
1MWT	0.414 (0.370-0.456) N=1520	0.414 (0.370-0.456) N=1520	0.487 (0.446-0.525) N=1520	-0.325 (-0.370- -0.278) N=1520
6MWT	0.446 (0.403-0.487) N=1473	0.417 (0.373- 0.460) N=1473	0.516 (0.466-0.554) N=1473	-0.443 (-0.484- -0.400) N=1473

Coefficient (ICC) according to the formula:

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

The MDC represents the smallest amount of change in a parameter that can be considered a true clinical alteration, distinguishing it from random measurement error. The MDC was calculated at the 95% confidence level (MDC_{95}) using the following formula:

$$MDC_{95} = SEM \times 1.96 \times \sqrt{2}$$

Because both SEM and MDC are mathematically derived from the SD, they share the identical unit of measurement as the evaluated spatial-temporal parameter. The only exception is the variable "pathlength/distance," which is calculated as a dimensionless ratio (Table 4).

ICC estimates and their 95% confidence intervals as well as SEM and MDC were calculated using SPSS statistical package version 29 based on single measures, absolute-agreement, 2-way mixed-effects model.

The specific Gangway parameters, clinical comparators, and primary statistical tests, as well as handling of statistical outliers were prespecified in an internal protocol prior to data extraction. The stratified subgroup analyses by GMFCS level and height, as well as the evaluation of the normalized step length to height were conducted post

hoc to ensure data integrity and address confounding variables.

Results

Study population

The mean age of the patients in the validity group was 7.5 (± 3.67) years. 42.6% (= 360/846) of the children were girls and 12.6% (= 102/846) had a GMFCS level I. 51.2% (= 433/846) of the examined patients suffered from tetraplegic spastic CP (Table 2). Similar demographics were observed in the reliability group, with mean age 6.95 (± 3.32) years. 42.1% (= 186/442) girls, 10.2% (= 45/442) GMFCS level I and 52% (= 230/442) with tetraplegic spastic CP (Table 3).

Validity

In the pooled analysis, all correlations were $p < 0.001$. In stratified analyses, p-values were < 0.001 unless otherwise indicated in Tables 11–12. All three clinical scores (GMFM-66, 1MWT, 6MWT) showed positive correlations with "average step length" and "average horizontal velocity" and negative correlations with "pathlength/distance." The variables "average step length" (rho 0.636, CI: 0.609-0.662) and

Table 7. Validity of the Gangway parameters evaluated in this study, including only one trial per participant. $p < 0.001$ in all variables.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
GMFM-66	0.662 (0.621-0.700) N=837	0.642 (0.642-0.682) N= 837	0.728 (0.693-0.759) N=837	-0.552 (-0.599- -0.501) N=837
1MWT	0.616 (0.510-0.703) N=171	0.475 (0.346-0.587) N=171	0.589 (0.478-0.682) N=171	-0.466 (-0.579- -0.336) N=171
6MWT	0.682 (0.572-0.767) N=126	0.447 (0.290- 0.580) N=126	0.540 (0.398-0.656) N=126	-0.495 (-0.619- -0.345) N=126

Table 8. Validity of the Gangway parameters evaluated in this study, including only one trial per participant, as well as Tukey's outliers. $p < 0.001$ in all variables.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
GMFM-66	0.668 (0.628-0.705) N=846	0.648 (0.606-0.686) N= 846	0.733 (0.699-0.763) N=846	-0.558 (-0.604- -0.508) N=846
1MWT	0.610 (0.505-0.698) N=176	0.476 (0.349-0.586) N=176	0.590 (0.481-0.681) N=176	-0.473 (-0.583- -0.345) N=176
6MWT	0.699 (0.595-0.780) N=129	0.477 (0.327- 0.604) N=129	0.557 (0.421-0.669) N=129	-0.499 (-0.622- -0.353) N=129

Table 9. Validity of the Gangway parameters evaluated in this study, stratified by body height. $p < 0.001$ in all variables.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
Height 83.5-105 cm				
GMFM-66	0.625 (0.574-0.672) N=638	0.609 (0.556-0.657) N=638	0.671 (0.624-0.712) N=638	-0.456 (-0.517- -0.390) N=638
1MWT	0.356 (0.270-0.437) N=449	0.354 (0.268-0.435) N=449	0.405 (0.322-0.482) N=449	-0.270 (-0.356- -0.179) N=449
6MWT	0.274 (0.182-0.361) N=438	0.270 (0.178- 0.357) N=438	0.377 (0.291-0.457) N=438	-0.359 (-0.440- -0.272) N=438
Height 105.3-126 cm				
GMFM-66	0.615 (0.565-0.661) N=673	0.615 (0.564-0.661) N=673	0.722 (0.682-0.757) N=673	-0.465 (-0.524- -0.402) N=673
1MWT	0.331 (0.246-0.411) N=475	0.332 (0.247-0.412) N=475	0.433 (0.355-0.506) N=475	-0.197 (-0.284- -0.106) N=475
6MWT	0.350 (0.264-0.431) N=452	0.351 (0.265- 0.431) N=452	0.447 (0.367-0.520) N=452	-0.300 (-0.384- -0.211) N=452
Height 126.4-193.7 cm				
GMFM-66	0.691 (0.648-0.729) N=673	0.698 (0.656-0.736) N=673	0.775 (0.741-0.804) N=673	-0.554 (-0.606- -0.498) N=673
1MWT	0.464 (0.389-0.534) N=482	0.466 (0.391-0.535) N=482	0.519 (0.449-0.584) N=482	-0.362 (-0.439- -0.280) N=482
6MWT	0.538 (0.469-0.600) N=478	0.515 (0.444- 0.580) N=478	0.586 (0.522-0.644) N=478	-0.487 (-0.554- -0.413) N=478

Table 10. Validity of the Gangway parameters evaluated in this study, stratified by body height, including Tukey's outliers. $p < 0.001$ in all variables.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
Height 83.5-105 cm				
GMFM-66	0.632 (0.586-0.674) N=753	0.620 (0.573-0.664) N=753	0.684 (0.643-0.722) N=753	-0.494 (-0.547- -0.436) N=753
1MWT	0.343 (0.261-0.420) N=503	0.342 (0.260-0.419) N=503	0.399 (0.321-0.472) N=503	-0.267 (-0.349- -0.181) N=503
6MWT	0.276 (0.189-0.358) N=489	0.271 (0.184- 0.354) N=489	0.379 (0.298-0.454) N=489	-0.364 (-0.441- -0.282) N=489
Height 105.3-126 cm				
GMFM-66	0.629 (0.581-0.673) N=710	0.627 (0.579-0.671) N=710	0.728 (0.690-0.762) N=710	-0.486 (-0.541- -0.425) N=710
1MWT	0.341 (0.258-0.418) N=493	0.340 (0.257-0.418) N=493	0.440 (0.364-0.511) N=493	-0.191 (-0.277- -0.102) N=493
6MWT	0.340 (0.255-0.420) N=467	0.341 (0.255- 0.421) N=467	0.441 (0.363-0.513) N=467	-0.310 (-0.392- -0.223) N=467
Height 126.4-193.7 cm				
GMFM-66	0.717 (0.679-0.751) N=728	0.721 (0.683-0.755) N=728	0.790 (0.760-0.816) N=728	-0.606 (-0.652- -0.557) N=728
1MWT	0.485 (0.414-0.549) N=524	0.482 (0.411-0.547) N=524	0.523 (0.456-0.584) N=524	-0.400 (-0.471- -0.323) N=524
6MWT	0.547 (0.481-0.606) N=517	0.522 (0.454- 0.584) N=517	0.594 (0.533-0.648) N=517	-0.516 (-0.579- -0.448) N=517

Table 11. Validity of the Gangway parameters evaluated in this study, stratified by GMFCS-level. $p < 0.001$ in all variables, if not others mentioned. “*” refers to non-significant p-values.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
GMFCS I				
GMFM-66	0.606 (0.521-0.679) N=265	0.397 (0.287-0.496) N=265	0.363 (0.251-0.466) N=265	-0.446 (-0.540- -0.341) N=265
1MWT	0.277 (0.148-0.397) N=224	0.113 (-0.022-0.245), $p=0.091$ N=224	0.157 (0.023-0.286), $p=0.019$ N=224	-0.304 (-0.422- -0.177) N=224
6MWT	0.441 (0.325-0.543) N=224	0.151 (0.016- 0.280), $p=0.024$ N=224	0.286 (0.158-0.406) N=224	-0.400 (-0.508- -0.280) N=224
GMFCS II				
GMFM-66	0.582 (0.525-0.634) N=609	0.508 (0.445-0.566) N=609	0.536 (0.475-0.592) N=609	-0.437 (-0.501- -0.368) N=609
1MWT	0.261 (0.171-0.346) N=467	0.181 (0.090-0.270) N=467	0.214 (0.123-0.301) N=467	-0.118 (-0.209- -0.025), $p=0.011$ N=467
6MWT	0.280 (0.190-0.366) N=449	0.168 (0.074- 0.259) N=449	0.287 (0.197-0.372) N=449	-0.300 (-0.385- -0.211) N=449
GMFCS III				
GMFM-66	0.386 (0.333-0.436) N=1110	0.378 (0.325-0.429) N=1110	0.457 (0.408-0.504) N=1110	-0.377 (-0.428- -0.324) N=1110
1MWT	0.160 (0.086-0.233) N=715	0.187 (0.113-0.259) N=715	0.288 (0.217-0.356) N=715	-0.149 (-0.222- -0.074) N=715
6MWT	0.087 (0.010-0.162), $p=0.022$ N=695	0.134 (0.058- 0.208) N=695	0.139 (0.064-0.214) N=695	-0.128 (-0.203- -0.052) N=695

Table 12. Validity of the Gangway parameters evaluated in this study, stratified by GMFCS-level, including Tukey's outliers. $p < 0.001$ in all variables, if not others stated. "***" refers to non-significant p-values.

	Spearman rho (95% CI)			
	Average step length	Normalized step length	Average velocity horizontal	Pathlength/Distance
GMFCS I				
GMFM-66	0.609 (0.529-0.678) N=294	0.419 (0.317-0.511) N=294	0.405 (0.302-0.499) N=294	-0.446 (-0.540- -0.341) N=294
1MWT	0.277 (0.155-0.391) N=250	0.123 (-0.005-0.247), $p=0.052^*$ N=250	0.157 (0.030-0.279), $p=0.013$ N=250	-0.304 (-0.422- -0.177) N=250
6MWT	0.440 (0.330-0.538) N=248	0.181 (0.054- 0.302), $p=0.004$ N=248	0.314 (0.193-0.425) N=248	-0.400 (-0.508- -0.280) N=248
GMFCS II				
GMFM-66	0.597 (0.543-0.645) N=650	0.523 (0.463-0.578) N=650	0.566 (0.510-0.618) N=650	-0.459 (-0.519- -0.394) N=650
1MWT	0.286 (0.201-0.368) N=492	0.214 (0.125-0.299) N=492	0.246 (0.158-0.329) N=492	-0.141 (-0.230- -0.051), $p=0.002$ N=492
6MWT	0.284 (0.196-0.367) N=470	0.175 (0.083- 0.264) N=470	0.295 (0.208-0.378) N=470	-0.319 (-0.400- -0.233) N=470
GMFCS III				
GMFM-66	0.429 (0.381-0.474) N=1248	0.413 (0.364-0.459) N=1248	0.487 (0.443-0.530) N=1248	-0.417 (-0.464 - -0.369) N=1248
1MWT	0.148 (0.076-0.218) N=778	0.170 (0.099-0.240) N=778	0.261 (0.192-0.327) N=778	-0.110 (-0.181- -0.038), $p=0.002$ N=778
6MWT	0.087 (0.014-0.160), $p=0.017$ N=755	0.131 (0.058- 0.202) N=755	0.141 (0.068-0.212) N=755	-0.120 (-0.191- -0.047) N=755

"average horizontal velocity" (rho 0.737, CI: 0.715-0.757) were the variables with the strongest correlation to GMFM-66. Both variables showed moderate correlations to 1MWT (rho 0.394, CI: 0.348-0.439 and 0.476, CI: 0.433-0.517 respectively) and 6MWT (rho 0.429, CI: 0.384-0.473 and 0.505, CI: 0.463-0.544 respectively). "Normalized step length to height" showed very similar correlations to those of "average step length" (rho 0.643, CI: 0.616-0.669 to GMFM-66, rho 0.400, CI: 0.354-0.444 to 1MWT and rho 0.407, CI: 0.360-0.451 to 6MWT). "Pathlength/distance" demonstrated moderate correlation to GMFM-66 (rho -0.525, CI: -0.557 - -0.492), as well as to 6MWT (rho -0.415, CI: -0.459 - -0.368) (Table 5). Including the Tukey outliers in the analysis values maintained the significance of our primary findings regarding construct validity (Table 6). To account for potential bias introduced by repeated measurements, a sensitivity analysis restricting the dataset to a single, initial investigation per participant was conducted (Table 7). The Spearman rank correlations remained highly robust and consistent with the primary analysis across all parameters. For instance, the correlation between "average horizontal velocity" and the GMFM-66 score remained exceptionally strong (rho 0.728, CI: 0.693-0.759), confirming that within-subject dependencies did

not artificially inflate the observed validity. The inclusion of Tukey outliers had no substantial impact on the robustness of this outcome (Table 8).

Stratified analysis by height showed tendentially stronger correlations between all parameters and clinical assessments in taller children (Table 9) and the robustness of this result was unaffected by the inclusion of Tukey outliers (Table 10). Stratification analysis by GMFCS level demonstrated consistent, significant correlations between Gangway parameters and GMFM-66 across all three evaluated functional levels (GMFCS I-III). However, the correlations between gait metrics and the endurance-based functional walk tests (1MWT, 6MWT) exhibited a distinct, level-dependent pattern (Table 11). Retaining statistical outliers identified via the Tukey method did not compromise the stability of these results (Table 12).

Reliability

The ICC value for the parameter "average step length" 0.827 (CI: 0.813-0.841) indicated very good reliability. For the parameter "pathlength/distance" the ICC value 0.598 (CI: 0.569-0.626) was indicative for moderate reliability, while the value for "average horizontal velocity" demonstrated poorer reliability (ICC 0.477, CI: 0.445-

Table 13. Reliability of the Gangway parameters evaluated in this study, $p < 0.001$ in all variables.

Parameter	ICC in single measures (95% CI), N trials for each ICC = 3
Average step length	0.827 (0.813-0.841), SEM: 0.04 m, MDC: 0.11 m, N sessions=1366 N participants=437
Average velocity horizontal	0.477 (0.445-0.509), SEM: 0.07 m/s, MDC: 0.20 m/s, N sessions=1318 N participants=435
Pathlength/ Distance	0.598 (0.569-0.626), SEM: 0.14, MDC: 0.40, N sessions=1204 N participants=423

0.509) (Table 13).

To evaluate the impact of extreme clinical presentations on measurement stability, a sensitivity analysis was conducted retaining the Tukey outliers (Table 14). Including these extreme values did not significantly affect the ICC for “average step length” (ICC 0.795, CI: 0.779-0.811) but notably increased the ICC for “average velocity horizontal” (ICC 0.637, CI: 0.612-0.662). Conversely, the reliability of the “pathlength/distance” decreased substantially when outliers were retained (ICC 0.297, CI: 0.263-0.331), accompanied by a four-fold increase in the SEM (0.57 vs. 0.14) and MDC (0.40 vs. 1.59). This severe degradation in reliability confirms that extreme “pathlength/distance” values largely represent unstable measurement artifacts in patients with minimal forward progression.

To ensure that the reliability estimates were not artificially influenced by within-subject dependencies, a sensitivity analysis restricting the data to a single session per participant was performed (Table 15). The absolute measurement errors (SEM and MDC) for “average step length” and “pathlength/distance” remained identical to the pooled dataset (Table 13). Interestingly, for “average horizontal velocity”, restricting the analysis to a single session per participant improved the ICC (from 0.477 to 0.598) and slightly reduced the standard error of measurement (SEM from 0.07 to 0.06, MDC from 0.20 to 0.18). Retaining the Tukey’s outliers revealed a clear disparity between relative (ICC) and absolute (SEM, MDC) reliability metrics. For instance, retaining outliers artificially inflated the ICC for “average horizontal velocity” from 0.598 to 0.745, while the absolute trial-to-trial measurement error (SEM) concurrently worsened from 0.06 to 0.08. For the “pathlength/distance” lead this to an almost five-fold increase in the SEM (0.14 to 0.65), whereas the ICC decreased considerably from 0.615 to 0.352, confirming the severe intra-session instability of these specific outlier values (Table 16).

Table 14. Reliability of the Gangway parameters evaluated in this study, including Tukey’s outliers, $p < 0.001$ in all variables.

Parameter	ICC in single measures (95% CI), N trials for each ICC = 3
Average step length	0.795 (0.779-0.811), SEM: 0.05 m, MDC: 0.14 m, N sessions=1416 N participants=442
Average velocity horizontal	0.637 (0.612-0.662), SEM: 0.08 m/s, MDC: 0.22 m/s, N sessions=1416 N participants=442
Pathlength/ Distance	0.297 (0.263-0.331), SEM: 0.57, MDC: 1.59, N sessions=1416 N participants=442

Table 15. Reliability of the Gangway parameters evaluated in this study, including only one session per participant. $p < 0.001$ in all variables.

Parameter	ICC in single measures (95% CI), N trials for each ICC = 3
Average step length	0.819 (0.792-0.844), SEM: 0.04 m, MDC: 0.11 m, N sessions=N participants=428
Average velocity horizontal	0.598 (0.547-0.647), SEM: 0.06 m/s, MDC: 0.18 m/s, N sessions=N participants=397
Pathlength/ Distance	0.615 (0.564-0.663), SEM: 0.14, MDC: 0.39, N sessions=N participants=383

Table 16. Reliability of the Gangway parameters evaluated in this study, including only one session per participant and Tukey’s outliers, $p < 0.001$ in all variables.

Parameter	ICC in single measures (95% CI), N trials for each ICC = 3
Average step length	0.795 (0.765-0.823), SEM: 0.05 m, MDC: 0.13 m, N sessions=N participants=442
Average velocity horizontal	0.745 (0.710-0.778), SEM: 0.08 m/s, MDC: 0.21 m/s, N sessions=N participants=442
Pathlength/ Distance	0.352 (0.293-0.412), SEM: 0.65, MDC: 1.80, N sessions=N participants=442

Discussion

The primary finding of this study demonstrated varying reliability among the examined parameters, but construct validity of the Leonardo Mechanograph® GW in children and adolescents with CP. While instrumental gait analysis is widely recognized as a fundamental tool for identifying pathological gait patterns and guiding rehabilitation, its clinical utility depends heavily on reproducibility and accuracy. In our study, “average horizontal velocity” showed the strongest correlation with GMFM-66 and can therefore be interpreted as the parameter with the strongest convergent association with GMFM-66 in this construct-validity analysis. However, it also demonstrated the lowest within-session reliability among the examined parameters. “Average step length” on the other side demonstrated both moderate convergent association with GMFM-66 and good reliability. This finding of paradox reliability between the parameters reflects the biomechanical realities of pediatric CP. Step length is dictated by structural and physiological constraints, such as spasticity and contractures making it highly consistent across repeated trials. Walking speed, however, is modulated by altering cadence, which is heavily influenced by behavioral factors, fatigue, and motivation, leading to higher trial-to-trial variability. Nevertheless, a child's overall capacity to achieve a higher walking velocity remains the ultimate indicator of their global motor function.

Reliability

To estimate the reliability of Gangway, we conducted 4248 investigations in total at a sample of 442 individuals, involving 3 gait trials at the same time under uniform circumstances and analyzed the parameters “average step length,” “average horizontal velocity” and “pathlength/distance” of the device. In every gait trial, the selected investigation involved in the analysis, is the one with the lowest “pathlength/distance”-value, as established in clinical practice of mechanographic gait analysis. As a result, the “true” reliability of the entire measurement process cannot be precisely assessed using the available methods. Our analysis is therefore restricted to the reliability of the best three of all trials a child performed during data collection, which form a single session for that patient. It is important to note that our analysis reports the ICC for single measures. This specific metric indicates the reliability of any single, standalone trial out of the three recorded passes. From a clinical perspective, this is highly relevant, as it demonstrates whether one individual walking trial is sufficiently reliable for decision-making, or if averaging multiple trials is mandatory to compensate for intra-individual variance. The research team of Veilleux et al. also investigated the reproducibility of Gangway parameters in 2011, but in 15 healthy adults. Each one walked three times barefoot and three times with shoes

over Gangway to investigate the intrasession trial-to-trial reproducibility under different conditions (barefoot vs. with shoes). Among other things, Veilleux et al. found a low variability (“average horizontal velocity” ICC 0.97, “pathlength/distance” ICC 0.94, “average step length” ICC 0.99) in the same three Gangway parameters examined in this work. The results from the investigations by Veilleux et al. consequently suggest that three investigations of the Gangway within one day provide reproducible data in healthy young adults⁹. In accordance, very good repeatability of “average step length” ICC 0.827 was shown in our study population as well. However, “average horizontal velocity” demonstrated poor reliability (ICC 0.477) and “pathlength/distance” moderate consistency (ICC 0.598) in our study population. Furthermore, another study by Veilleux et al. in 2014 examined the reproducibility of Gangway in gait analysis in within-day trials in 19 healthy children and adolescents. Each participant walked in three blocks three times barefoot and in three blocks three times with shoes over the Gangway. No differences were detected between trials within each condition. All in this study measured force and spatiotemporal parameters had ICC more than 0.87 and coefficients of variation in the order of 1-6%. These values indicated repeatable within-day results of the Gangway for healthy children and adolescents¹⁰. Moreover Stolze et al. in 1998 examined the repeatability of the gait analysis using the method of footprint analysis in 12 healthy children and 12 healthy adults. The retest-reliability was evaluated in that study by calculating Pearson's product-moment correlation coefficients and intraclass correlation coefficients. The results showed in healthy children good reliability (ICC 0.72) for step length and poor (ICC 0.35) for gait velocity, congruent to our study (ICC 0.827 and 0.477 in our study population, respectively)¹². Regarding that children with CP might have a physical limit on how long their stride can be (due to bone structure and permanent muscle contractures for example), their consistent step length (“average step length”) could be explained. When it comes to healthy children and young adults, like those investigated in the mentioned studies, the gait is an automated process. On the other hand, gait speed (“average horizontal velocity”) of children with CP as well as their ability to walk in a straight line (measured as “pathlength/distance”) is a complex procedure, constantly changing based on balance, fatigue, momentary muscle spasticity, because of the conscious effort and concentration needed for them to walk. Therefore, the results of the last two parameters are expected to be unstable in every trial in this study population.

To summarize, the ICC value in our large study population of pediatric patients with CP for the examined Gangway parameters revealed: good (0.827, CI: 0.813-0.841), poor to moderate (0.477, CI: 0.445-0.509) and moderate (0.598, CI: 0.569-0.626) within-session repeatability for “average step length,” “average horizontal velocity” and “pathlength/distance” respectively. The narrow

confidence intervals further strengthened the reliability of these findings. The good reliability of “average step length” is particularly remarkable, as this parameter is often critical in assessing gait symmetry and efficiency. The poor reliability of “average horizontal velocity” suggested that it would be an insufficient metric for tracking small improvements over time or evaluating the success of a specific intervention, as the error or natural variance is too high. This high variance in pediatric cerebral palsy is likely multifactorial, driven by behavioral, physical, and neurological fluctuations. Unlike healthy adults who use a highly automated and consistent habitual walking speed, children exhibit significant behavioral variability across repeated trials, often altering their pace due to distraction, fluctuating motivation, or difficulty adhering to instructions to walk at a normal speed. Furthermore, the inherently high energy expenditure of pathological gait leads to rapid fatigue, potentially causing velocity to drop across sequential trials. Neurologically, fluctuating spasticity and the constant need to decelerate to recover balance introduce further inconsistencies. Moreover, for more severely affected children (GMFCS II and III), the use of external supports introduces substantial trial-to-trial inconsistency. The management of walking aids, orthoses, and variable standby assistance from caregivers creates unpredictable time delays. A child may briefly pause to reposition a walker, adjust their grip, or react to a caregiver’s guidance, drastically altering the average velocity of a specific trial. Additionally, while orthoses provide essential stability, the continuous biomechanical adaptation they demand can lead to fluctuating gait efficiency and discomfort over repeated passes. The moderate within-session repeatability of the variable “pathlength/distance” reflects the motor inconsistency and postural instability, i.e. high stride-to-stride variability, which characterizes CP and contrasts with the stable gait patterns observed in the healthy reference groups.

Comparing the reliability analyses with and without statistical outliers revealed critical insights into the trial-to-trial consistency of the evaluated parameters. Notably, for “average horizontal velocity”, retaining outliers mathematically increased the ICC (from 0.477 to 0.637) but simultaneously increased the absolute measurement error between the three repeated trials (SEM increased from 0.07 to 0.08 and MDC from 0.20 to 0.22). This indicates that the higher ICC in the outlier-inclusive dataset is an artifact of inflated between-subject variance (due to the presence of extremely slow walkers), rather than an actual improvement in trial-to-trial consistency. Because patients must exceed this large MDC threshold (up to 0.22 units) to demonstrate true improvement, this parameter’s measurement error is likely too high to reliably track subtle therapeutic changes. The repeated measurements themselves were more stable when statistical outliers were excluded. This effect is most drastically observed in the parameter “pathlength/distance”. When extreme values were included, the trial-to-trial measurement error

quadrupled (SEM increased from 0.14 to 0.57 and MDC from 0.40 to 1.59), causing the ICC to drop to 0.297. This confirms that these extreme physiological values are highly unstable across repeated trials within the same session, largely representing unstable measurement artifacts in patients with minimal forward progression. This disparity between relative (ICC) and absolute (SEM, MDC) reliability metrics proves that utilizing absolute error metrics and appropriately managing statistical outliers is essential to ascertain the true clinical precision of gait analysis systems.

To evaluate whether the within-session reliability of the spatial-temporal parameters was robust and not artificially inflated by the inclusion of multiple sessions per participant, we conducted a sensitivity analysis using one session per child. In fact, for “average horizontal velocity”, the reliability markedly improved when restricted to a single evaluation per patient. This suggests that mixing longitudinal data across different developmental stages (e.g., sessions spanning several years of growth or surgical interventions for the same child) introduces temporal variance that mathematically suppresses the intra-session ICC. Thus, the restricted single-session dataset likely provides a more accurate reflection of the true trial-to-trial reliability of “average horizontal velocity” during a single clinical assessment. Conversely, “average step length” remained highly robust; its absolute measurement errors were entirely unaffected by session pooling, confirming that its variance represents true physiological differences rather than measurement noise.

In conclusion, three consecutive, under identical conditions trials on Gangway considering the variables “average step length” and “pathlength/distance” provided reproducible data for children and adolescents with CP, whereas for longitudinal monitoring, tracking small clinical changes and treatment response “average horizontal velocity” should be evaluated in conjunction with more mechanically stable parameters, because of its high natural variance.

Validity

Regarding the validity of Gangway in gait analysis, to our knowledge no other published studies exist. In the literature, evidence of Gangway as a valid tool has only been provided for balance measurements from the research team Wiegman et al., 2019¹¹. To evaluate the construct validity of Gangway in gait analysis in children and adolescents with CP, we examined the correlations between the selected Gangway parameters: “average step length,” “average horizontal velocity” and “pathlength/distance” and the well-established and frequently used clinical scores for assessing walking and standing ability in children with movement impairments: GMFM-66, 1MWT and 6MWT, in a sample of 846 children and adolescents with CP. The significant correlations ($p < 0.001$) observed between Gangway parameters and clinical scores provided

compelling evidence for the system's construct validity. In particular, "average step length" and "average horizontal velocity" exhibited positive correlations with all three clinical scores, indicating that improvements in gait function, as measured by these established tests, were reflected in increased step length and walking speed. "Pathlength/distance" correlated, as expected, negatively to all clinical scores, meaning that as a patient's gross motor function or walking endurance increases (i.e. higher clinical scores), their pathlength/distance ratio decreases (i.e. the patient walks with better directional control). In other words, a longer "pathlength/distance" (indicating more deviation from a straight-line during walking) was associated with poorer performance in gross motor function (GMFM-66) and walking endurance (6MWT, 1MWT). All parameters correlated most strongly with GMFM-66. This suggested that Gangway was particularly sensitive to changes in gross motor function, as assessed by the GMFM-66. An important consideration in our validity analysis is the inherent nature of the GMFM-66. It is a comprehensive assessment of global gross motor function rather than an exclusive measure of walking ability. The GMFM-66 evaluates multiple dimensions of motor performance, including lying, sitting, standing, jumping and running. Because gait represents only one component of this global score, a very high correlation with specific spatial-temporal gait parameters was not anticipated. Instead, the moderate correlation we hypothesized and subsequently observed aligns exactly with the construct of the GMFM-66. Consequently, rather than reflecting a measurement weakness, this moderate correlation actively supports the construct validity of the evaluated parameters, demonstrating that they accurately capture the gait-related variance within the broader context of a patient's global motor function. The strongest correlations were found between "average horizontal velocity" and GMFM-66 (ρ 0.737, CI: 0.715-0.757) as well as "average step length" and GMFM-66 (ρ 0.636, CI: 0.609-0.662). Both variables showed moderate correlations to 1MWT (ρ 0.476, CI: 0.433-0.517 and 0.394, CI: 0.348-0.439 respectively) and 6MWT (ρ 0.505, CI: 0.463-0.544 and 0.429, CI: 0.384-0.473 respectively). The consistent correlations across different clinical measures underscored Gangway's ability to provide a comprehensive assessment of gait function, validating its use in clinical and research settings. The parameter "pathlength/distance" correlated modest to GMFM-66 and 6MWT (ρ -0.525, CI: -0.557 - -0.492 and -0.415, CI: -0.459 - -0.368 respectively) and weak to 1MWT (ρ -0.300, CI: -0.348 - -0.251). These correlations indicated that "pathlength/distance" was a relevant factor in assessing motor abilities, but less useful when evaluating short walking distances. In a short-distance test, such as 1MWT, where the patients in our department are asked to walk in circle with a length of 40 m, the required permanent turning becomes a confounding factor that weakens this parameter's utility. Interestingly, normalizing

step length to height did not yield stronger correlations than absolute "average step length". However, height stratification revealed tendentially stronger correlations across all parameters in taller children. This is consistent with the biomechanical realities of pediatric CP. Younger, smaller children inherently exhibit higher intra-individual gait variability, shorter step lengths, and lower compliance during testing, which introduces more signal noise and mathematically dampens correlations. In contrast, older children typically demonstrate a more habitual, stable gait pattern and longer steps, leading to cleaner data capture and consequently stronger statistical associations. A major methodological strength of these findings is their resilience against data artifacts. Our sensitivity analyses confirmed that the strong correlations were not artificially inflated by within-subject dependencies from repeated measurements. Restricting the dataset to a single initial investigation per participant maintained exceptionally strong correlations (e.g., "average horizontal velocity" to GMFM-66 remaining at ρ 0.728). Furthermore, retaining Tukey outliers did not compromise the stability of these primary findings. This suggests that the relationship between Gangway parameters and the GMFM-66 is fundamentally physiological rather than a mathematical artifact. Finally, the stratification by GMFCS level highlighted an important distinction between measuring motor capacity versus endurance. While Gangway metrics consistently correlated with the GMFM-66 across all severities (GMFCS I-III), their correlation with endurance-based tests (1MWT, 6MWT) showed a distinct, level-dependent pattern, including a non-significant correlation between normalized step length and the 1MWT in GMFCS I. This divergence underscores that while short-distance spatial-temporal analysis, like the gait analysis using Gangway, provides a robust representation of global motor capability as measured by the GMFM-66, it cannot fully capture the complex cardiovascular and muscular fatigue variables that dictate performance in endurance testing, particularly at the extreme ends of the functional spectrum. It is interesting to mention that stratified analysis by GMFCS level revealed that correlations between both absolute and normalized average step length and the GMFM-66 are substantially attenuated in patients with GMFCS level III compared to those in levels I and II. Because of this reduced associative strength, the construct validity of "average step length" is insufficient to support its clinical application for gait evaluation in this more severely impaired subgroup. This attenuated correlation likely reflects the severe biomechanical constraints in this cohort -such as pronounced structural contractures and the obligatory use of walking aids- which mechanically restrict step expansion and decouple it from broader motor capacity. According to all these findings, it can be assumed that regarding the parameters "average step length" and "average horizontal velocity", Gangway provides construct validity in children and adolescents with CP.

The combined findings of strong convergent associations with clinical function and good reliability on Gangway's "average step length" highlight the potential of this Gangway-parameter as a valuable tool in clinical practice. "Average horizontal velocity" remains a highly valid indicator of global motor function, but necessitates caution in its use for tracking clinical changes.

Limitations

While this study demonstrated the reliability and validity of Gangway in gait analysis, several limitations should be acknowledged. To begin with, our study was conducted as a monocentric study with data from a population consisting only of participants in the rehabilitation program "Auf die Beine", which treats children with several motor limitations. Consequently, there is a low number of children with CP GMFCS-level I in this cohort. Secondly, while clinical mechanography routinely utilizes the best of three walk trials, establishing the true within-session repeatability of this standard approach would require comparing the maximal performance from two independent series of three walk trials. However, due to the retrospective design of this study, our analysis was limited to determining the reproducibility of a single-walk protocol. To this extent, in a previous study from our Rehabilitation Center, the reliability and validity of the Leonardo Mechanograph® GRFP STD (jumping platform) were also statistically proven using retrospective analysis, and the method proved to be statistically reliable¹⁶. Moreover, while the three consecutive gait trials for every patient and the clinical tests were conducted under the same circumstances, a variability in the orthoses, when used or in the assistance provided from a helper, when needed, cannot be excluded. The specific details were not systematically recorded for group-level analysis. Thus, a notable limitation of this study is the lack of stratified analysis regarding the use of orthoses, varying footwear, and physical assistance. Because these external factors can significantly influence gait kinematics -particularly in patients with higher GMFCS levels- future studies should systematically document and adjust for orthotic use and manual support to further refine the validity of mechanographic gait parameters. Additionally, a certain examiner-dependent variability in the measurements, because of the extended data collection period (2007-2020), cannot be completely ruled out. In this regard, during this long data collection period (13 years) it has come to software updates of the Gangway, but the system remained compatible. Although no selection bias was identified by the authors, it still cannot be excluded. The exclusion of irregular, not plausible gait data might have biased the sample. A further limitation is the retrospective nature of data collection from the Gangway, so that discrepant or doubtful results could not be fundamentally checked again. For this reason, as well, detailed documentation regarding safety monitoring during the trials, like the occurrence of fatigue, minor

stumbles, or the need to temporarily stop testing was not available. Moreover, as a retrospective analysis, this study lacks comparison against a 3D motion capture criterion standard or a validated walkway system and is subject to unadjusted confounding from clinical variables such as specific orthotic configurations, prior interventions, and daily fluctuations in patient fatigue. Lastly, to improve the quality of our study level, participants whose gait analysis' parameters were not reliably measurable, due to various technical or personal reasons, i.e. more assistance from the helper than personal gait effort or not enough foot contact to the walkway, were excluded from the evaluation to ensure and maintain the quality standards of the work.

Conclusion

This study revealed that the gait analysis on a Leonardo Mechanograph GRFP® GW may be useful as a complementary clinical tool to evaluate physical performance of pediatric patients with CP, particularly for "average step length" and mainly in GMFCS I–II. Regarding the reliability of the GW the parameter "average step length" should be preferred in the gait analysis of this population, whereas "average horizontal velocity" should be evaluated in conjunction with more mechanically stable parameters, because of its high natural variance. Future, multicenter studies are needed to determine the effectiveness of gait analysis using Gangway in a broader clinical context (e.g. other neurological disorders or orthopedic conditions) and within diverse age ranges.

Ethics approval

This study is a retrospective secondary analysis of data collected prospectively within the single-center patient registry of the children rehabilitation center at the University of Cologne. The registry and the collection and storage of clinical data were approved by the Ethics Committee of the University of Cologne (approval no. 16-269) and registered in the German Clinical Trials Register (DRKS00011331). The Ethics Committee confirmed that approval no. 16-269 covered the retrospective analysis of the registry data presented in the current study, including data collected between May 2007 and January 2020. No additional ethics approval was required for the present retrospective analysis. All procedures were performed in accordance with the ethical standards of the institutional research committee and with the Declaration of Helsinki 1964 and its later amendments.

Consent to participate

Written informed consent for participation in the rehabilitation program and for the scientific use of the collected clinical data was obtained from the parents or legal guardians of all participants, while assent was additionally obtained from children and adolescents whenever age and cognitive capacity allowed.

Authors' contributions

Andriani Bougadaki: conception and design of the work, data collection, methodology development, data processing (analysis, interpretation), critical review of the manuscript, responsible for the integrity of the data analysis, wrote the initial draft of the manuscript,

read and approved the final version of the paper. Ibrahim Duran: conception and design of the work, data collection, methodology development, data processing (analysis, interpretation), critical review of the manuscript, read and approved the final version of the paper. Eckhard Schoenau: conception and design of the work, data collection, methodology development, critical review of the manuscript, read and approved the final version of the paper. Kyriakos Martakis: conception and design of the work, data collection, methodology development, data processing (analysis, interpretation), critical review of the manuscript, read and approved the final version of the paper.

Acknowledgements

We are grateful to the young patients and their families for their valuable participation, as well as to the dedicated physiotherapists and the Gangway-raters at UniReha GmbH, Centre of Prevention and Rehabilitation, University of Cologne, for their crucial role in the program "Auf die Beine".

References

- Richards CL, Malouin F. Cerebral palsy: definition, assessment and rehabilitation. *Handb Clin Neurol*. 2013;111:183-95.
- 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.
- Patel DR, Neelakantan M, Pandher K, Merrick J. Cerebral palsy in children: a clinical overview. *Transl Pediatr*. 2020;9(Suppl 1):S125-S135.
- McIntyre S, Goldsmith S, Webb A, Ehlinger V, Hollung SJ, McConnell K, et al. Global prevalence of cerebral palsy: A systematic analysis. *Dev Med Child Neurol*. 2022;64(12):1494-506.
- Arnaud C, Ehlinger V, Perraud A, Kinsner-Ovaskainen A, Klapouszczak D, Himmelmann K, et al. Public health indicators for cerebral palsy: A European collaborative study of the Surveillance of Cerebral Palsy in Europe network. *Paediatr Perinat Epidemiol*. 2023;37(5):404-12.
- Steven S, Spiess K, Lentzen B, Reinhart N, Schoenau E, Duran I. Effect of a Rehabilitation Program Including Home-Based Vibration-Assisted Therapy on Gait Parameters in Children with Cerebral Palsy. *J Musculoskelet Neuronal Interact*. 2024;24(3):259-66.
- Pilz F, Vill K, Rawer R, Bonfert M, Tacke M, Heussinger N, et al. Mechanography in children: pediatric references in postural control. *J Musculoskelet Neuronal Interact*. 2022;22(4):431-54.
- Claverie L, Ille A, Moretto P. Discrete sensors distribution for accurate plantar pressure analyses. *Med Eng Phys*. 2016;38(12):1489-94.
- Veilleux LN, Robert M, Ballaz L, Lemay M, Rauch F. Gait analysis using a force-measuring gangway: intrasession repeatability in healthy adults. *J Musculoskelet Neuronal Interact*. 2011;11(1):27-33.
- Veilleux LN, Ballaz L, Robert M, Lemay M, Rauch F. Analysing gait using a force-measuring walkway: intrasession repeatability in healthy children and adolescents. *Comput Methods Biomech Biomed Engin*. 2014;17(13):1447-51.
- Wiegmann S, Felsenberg D, Gast U, Böst H, Armbrrecht G, Dietzel R. Balance Performance across the Lifespan Assessed by the Leonardo Mechanograph®: A Cross-Sectional Study. *J Funct Morphol Kinesiol*. 2019;5(1).
- Stolze H, Kuhtz-Buschbeck JP, Mondwurf C, Jöhnk K, Friege L. Retest reliability of spatiotemporal gait parameters in children and adults. *Gait Posture*. 1998;7(2):125-30.
- Dussault-Picard C, Cherni Y, Dixon PC. Spatiotemporal characteristics of gait when walking on an uneven surface in children with cerebral palsy. *Sci Rep*. 2025;15(1):4912.
- Stark C, Hoyer-Kuhn HK, Semler O, Hoebing L, Duran I, Cremer R, et al. Neuromuscular training based on whole body vibration in children with spina bifida: a retrospective analysis of a new physiotherapy treatment program. *Childs Nerv Syst*. 2015;31(2):301-9.
- 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.
- Duran I, Martakis K, Stark C, Alberg E, Bossier C, Semler O, et al. Experience with jumping mechanography in children with cerebral palsy. *J Musculoskelet Neuronal Interact*. 2017;17(3):237-45.
- Novotec Medical GmbH. Leonardo Mechanograph® GW . 2026 Available from: <https://mechanography.com/produkte/leonardo-mechanograph-gw/>.
- Russell DJ, Rosenbaum PL, Cadman DT, Gowland C, Hardy S, Jarvis S. The gross motor function measure: a means to evaluate the effects of physical therapy. *Dev Med Child Neurol*. 1989;31(3):341-52.
- Russell DJ, Avery LM, Rosenbaum PL, Raina PS, Walter SD, Palisano RJ. Improved scaling of the gross motor function measure for children with cerebral palsy: evidence of reliability and validity. *Phys Ther*. 2000;80(9):873-85.
- Avery LM, Russell DJ, Raina PS, Walter SD, Rosenbaum PL. Rasch analysis of the Gross Motor Function Measure: validating the assumptions of the Rasch model to create an interval-level measure. *Arch Phys Med Rehabil*. 2003;84(5):697-705.
- Duran I, Stark C, Martakis K, Hamacher S, Semler O, Schoenau E. Reference centiles for the gross motor function measure and identification of therapeutic effects in children with cerebral palsy. *J Eval Clin Pract*. 2019;25(1):78-87.
- Martakis K, Stark C, Rehberg M, Semler O, Duran I, Schoenau E. One-Minute Walk Test in Children with Cerebral Palsy GMFCS Level 1 and 2: Reference Values to Identify Therapeutic Effects after Rehabilitation.

- Dev Neurorehabil. 2020;23(4):201-9.
23. Maher CA, Williams MT, Olds TS. The six-minute walk test for children with cerebral palsy. *Int J Rehabil Res.* 2008;31(2):185-8.
 24. Martakis K, Stark C, Rehberg M, Semler O, Duran I, Schoenau E. Reference Centiles to Monitor the 6-minute-walk Test in Ambulant Children with Cerebral Palsy and Identification of Effects after Rehabilitation Utilizing Whole-body Vibration. *Dev Neurorehabil.* 2021;24(1):45-55.
 25. Tukey JW. *Exploratory data analysis*: Springer; 1977.
 26. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg.* 2018;126(5):1763-8.
 27. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;15(2):155-63.