

RESEARCH ARTICLE

Capture of Longitudinal Change in Real-Life Walking in Cerebellar Ataxia Increases Patient Relevance and Effect Size

Jens Seemann, MSc,^{1,2} Theresa Beyme, MD,^{1,3} Natalie John,^{1,3} Florian Harmuth, MSc,⁴ Martin Giese, PhD,^{1,2} Ludger Schöls, MD,^{5,6} Dagmar Timmann, MD,⁷ Matthis Synofzik, MD,^{3,5}  and Winfried Ilg, PhD^{1,2*} 

ABSTRACT: Background: With disease-modifying drugs for degenerative ataxias on the horizon, ecologically valid measures of gait performance that can detect patient-relevant changes in trial-like time frames are highly warranted.

Objectives: In this 2-year longitudinal study, we aimed to unravel ataxic gait measures sensitive to longitudinal changes in patients' real lives using wearable sensors.

Methods: We assessed longitudinal gait changes of 26 participants with degenerative cerebellar disease (Scale for the Assessment and Rating of Ataxia [SARA]: 9.4 ± 4.1) at baseline, 1-year, and 2-year follow-up using three body-worn inertial sensors in two conditions: (1) laboratory-based walking (LBW); and (2) real-life walking (RLW). In RLW, a context-sensitive analysis was performed by selecting comparable walking bouts according to macroscopic gait characteristics. Gait analysis focused on measures of spatio-temporal variability, particularly stride length variability, lateral step deviation, and a compound measure of spatial variability (SPCmp).

Results: Gait variability measures showed high test-retest reliability in both walking conditions (intraclass

correlation coefficient [ICC], ≥ 0.82). Cross-sectional analyses revealed high correlations of gait measures with ataxia severity (SARA, effect size $\rho \geq 0.75$); and with patients' subjective balance confidence (Activity-specific Balance Confidence scale [ABC]: $\rho \geq 0.71$).

Although SARA showed longitudinal changes only after 2 years, the gait measure SPCmp revealed changes after 1 year with high effect size ($r_{\text{prb}} = 0.80$). Sample size estimation for the gait measure SPCmp showed a required cohort size of $n = 42$ participants ($n = 38$; spinocerebellar ataxias [SCA]_{1/2/3} subgroup) to detect a 50% reduction in progression at 1 year with a hypothetical intervention, compared to $n = 147$ for SARA at 2 years.

Conclusions: Because of their ecological validity and larger effect sizes, real-life gait characteristics represent promising performance measures as outcomes for future treatment trials. © 2025 The Author(s). *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

Key Words: biomarker; cerebellar ataxia; digital health; real-life walking; wearable sensors

¹Section Computational Sensomotrics, Hertie Institute for Clinical Brain Research, Tübingen, Germany; ²Centre for Integrative Neuroscience (CIN), Tübingen, Germany; ³Division Translational Genomics of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research and Center for Neurology, University of Tübingen, Tübingen, Germany; ⁴Institute of Medical Genetics and Applied Genomics, University Hospital Tübingen, University of Tübingen, Tübingen, Germany; ⁵German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany; ⁶Department of Neurodegeneration, Hertie Institute for Clinical Brain Research and Centre of Neurology, Tübingen, Germany; ⁷Department of Neurology and Center for Translational and Behavioral Neurosciences (C-TNBS), Essen University Hospital, University of Duisburg-Essen, Essen, Germany

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

*Correspondence to: Dr. W. Ilg, Section Computational Sensomotrics, Hertie Institute for Clinical Brain Research, Otfried-Müller-Straße 25, 72076 Tübingen, Germany; E-mail: winfried.ilg@uni-tuebingen.de

Matthis Synofzik and Winfried Ilg contributed equally.

Relevant conflicts of interest/financial disclosures: Statistical analysis was conducted by J.S. and W.I. There is no conflict of interest for any of the authors.

Funding agencies: This work was supported by the International Max Planck Research School for Intelligent Systems (IMPRS-IS) (to J.S.) and the Else Kröner-Fresenius-Stiftung Medical Scientist program 'ClinbrAln' (to W.I. and M.G.), as well as the Else Kröner-Fresenius Stiftung Clinician Scientist program "PRECISE.net" (to M.S.). In addition, this work was supported by the European Union, project European Rare Disease Research Alliance, GA no.101,156,595, funded under call HORIZON-HLTH-2023-DISEASE-07 (to M.S. and J.S.).

Received: 26 November 2024; **Revised:** 20 March 2025; **Accepted:** 24 April 2025

Published online 21 May 2025 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.30230

Gait measures constitute promising candidates for performance outcome measures in upcoming therapeutic intervention studies in ataxias,^{1,2} because gait disturbances often present as the first signs of degenerative cerebellar disease (DCD)³⁻⁵ and represent one of the most patient-relevant disabling features throughout the disease course.^{6,7} It has been shown in laboratory-based assessments by different capture technologies that measures of spatio-temporal variability allow to characterize ataxic gait⁸⁻¹⁴ (for reviews, see Buckley et al,¹⁵ Milne et al,¹⁶ Ilg et al,¹⁷ and Cabaraux et al¹⁸) with high sensitivity to ataxia severity in cross-sectional and first longitudinal studies.^{19,20}

With progress in wearable sensor technology enabling gait recordings in patients' real life, it was hypothesized that those real-life gait measures could be potentially even more sensitive to disease-specific signatures of ataxic gait impairment compared to clinical and lab settings, because of the complexity and challenges of the environments,²¹⁻²⁵ but also because of the larger amount of available walking bouts.²⁶ In a first cross-sectional study on real-life gait in degenerative cerebellar ataxia, we have shown that ataxic-sensitive gait measures allow not only to capture the gait variability inherent in ataxic gait in real life, but also to demonstrate high sensitivity to small cross-sectional differences in disease severity, with higher effect sizes in real-life walking (RLW) compared to clinical gait assessment.²⁷ However, to serve as ecologically valid patient-focused progression and therapy response outcomes, these gait measures have to prove (1) their sensitivity to individual longitudinal change in a short period realistic for intervention trials,^{1,28} and (2) their meaningfulness by anchoring with patient-reported outcomes, as required by the United States Food and Drug Administration (FDA).²⁹

Noteworthy, to compare patients' real-life gait behavior at two measurement time points, the influence of context and environment on gait measures have to be considered. These contextual and environmental factors have been shown to have a significant impact on macroscopic (ie, behavioral, as opposed to microscopic, ie, spatio-temporal)^{30,31} gait characteristics, such as speed, length of walking bouts, and number of turns.^{25,32-34} These gait characteristics will differ for indoor (eg, in a small apartment) versus outdoor walking, and in turn will influence several gait measures,^{30,35} as it has been shown in healthy participants and for different patient populations (Parkinson's disease, multiple sclerosis, cerebral palsy).³⁵⁻³⁷ This holds for general performance measures like mean gait speed and even more for variability measures.³⁶⁻³⁸ Analysis of shorter walking bouts for indoor walking—compared to longer outdoor bouts—inherently provides increased variability for healthy participants and patients.^{37,39}

Therefore, if one compares measures of a patient's gait variability at two time points 1 year apart, increased variability could be caused not by increased

balance disturbances, but rather by mere differences in the context and environmental factors of the recorded gait behavior. Therefore, to identify disease progression or treatment-induced changes in longitudinal analyses of real-life gait behavior, it is necessary to consider these contextual and environmental factors, particularly their influence on macroscopic gait characteristics.

Here, we performed a longitudinal analysis of baseline, 1- and 2-year follow-up gait recordings in lab-based gait assessment and patients' real life. Matching of longitudinal walking bouts (baseline and follow-up) was performed according to macroscopic characteristics of walking behavior, namely the bout length and number of turns. We hypothesized that gait measures capturing longitudinal change in patients' real life could be (1) more sensitive to progression in short, trial-like time-frames (eg, 1 year) compared to lab-based gait assessments and clinical rating scales; and (2) more patient-relevant in terms of correlation with patient-reported outcomes of balance confidence in important activities of everyday living. This would be key for future treatment trials, as the targeted primary outcome is usually a slowing of disease progression in a limited study period, ideally capturable within 1 year, and by outcomes reflecting patient relevance, as emphasized by the FDA.²⁹

Patients and Methods

Participants and Clinical Outcome Assessments

Study Participants

Twenty-six participants at an early-to-moderate ataxic or pre-ataxic stage of DCD (age: 48 ± 9.5 years) were recruited from the Ataxia Clinics of the University Hospitals Tübingen and Essen. They consisted of 21 participants in the ataxic (ATX) stage of DCD as defined by a Scale for the Assessment and Rating of Ataxia (SARA) score of ≥ 3 (group ATX; SARA: 9.4 ± 3.2), and five participants with repeat-expansions in spinocerebellar ataxia SCA2, SCA3, or SCA6 at the pre-ataxic (PRE) stage of DCD (SARA < 3) (group PRE: SARA: 1.6 ± 0.65).⁴⁰ A total of 18 of 26 DCD participants carried a repeat expansion in SCA1, 2, or 3 (SCA_{1/2/3} subgroup). All main analyses were additionally performed in this subgroup, because these fast-progressing polyglutamine (polyQ) SCA types are a promising target in many upcoming intervention trials.^{1,2} Details of patient characteristics are shown in Table 1.

Patients were included based on the following inclusion criteria: (1) degenerative cerebellar ataxia in the absence of any signs of secondary central nervous system disease; (2) age between 18 and 75 years; and (3) ability to walk without walking aids. Exclusion criteria

TABLE 1 Patient characteristics at baseline assessment.

Patient	Diagnosis	SARA ^{BL}	SARA _{p&g} ^{BL}	Stage ^{BL}
PRE1	SCA6	2.5	0	0
PRE2	SCA2 ^{Conv}	2	0	0
PRE3	SCA3	1	0	0
PRE4	SCA3	1	0	0
PRE5	SCA3	1.5	0	0
ATX1	SCA6	3	0	0
ATX2	SCA3	8.5	4	2
ATX3	ADCD	4.5	2	1
ATX4	ADCK3	8.5	3	2
ATX5	SCA3	13	6	2
ATX6	SCA2	4.5	1	1
ATX7	SCA14	10	4	2
ATX8	ADCK3	10	5	2
ATX9	PNPLA6	9.5	4	2
ATX10	SCA1	12	5	2
ATX11	SCA2	12.5	5	2
ATX12	SCA3	13	5	2
ATX13	SCA3	9	4	2
ATX14	SCA6	8.5	2	1
ATX15	SCA1	17.5	6	2
ATX16	SCA1	10.5	4	2
ATX17	SCA3	6.5	3	2
ATX18	SCA1	9	3	2
ATX19	SCA3	7.5	3	2
ATX20	SCA3	9.5	5	2
ATX21	SCA3	10	4	2
Subgroup	PRE #5 age Ø45.2 ± 10.1 BMI Ø23.8 ± 2.5	SARA Ø1.6 ± 0.7	T2EDO Ø: −3.8 [−11,3]	
Subgroup	ATX #21 age Ø49.8 ± 9.1 BMI Ø24.4 ± 3.7	SARA Ø9.4 ± 3.2		
Group	DCD #26 age Ø48.9 ± 9.5 BMI Ø24.3 ± 3.5	SARA Ø7.9 ± 4.2		
Controls	HC #34 age Ø44.08 ± 14.8 BMI Ø25.3 ± 4.5			

Clinical ataxia severity was determined by the SARA.⁴⁰ Cerebellar patients include pre-ataxic (PRE: SARA<3) and ataxic participants (ATX: SARA≥3). Participant PRE2 converted after 1 year to the ataxic stage (Conv). SARA scores and FARS ataxia staging⁴¹ (stage) are shown at BL. T2EDO was defined for the PRE subgroup as the difference between current age and estimated age at onset,⁴² with estimated disease onset calculated based on the individual's CAG repeats, as described in Tezenas du Montcel et al.⁴³ Negative values denote estimated disease onset in the future, positive values denote estimated disease onset in the past. The following diagnosis denotes the gene underlying the respective ataxia type: ADCK3 (=ARCA 2, autosomal-recessive cerebellar ataxia type 2); *PNPLA6*.

Abbreviations: SARA, Scale for the Assessment and Rating of Ataxia; BL, baseline; SARA_{p&g}, Scale for the Assessment and Rating of Ataxia posture and gait; PRE, pre-ataxic; SCA, autosomal-dominant spinocerebellar ataxia of defined genetic type; ATX, ataxic participants; ADCD, autosomal dominant ataxia of still undefined genetic cause; BMI, body mass index; T2EDO, time to disease onset; HC, healthy controls.

were: severe visual or hearing impairment, cognitive impairment that limits the understanding of instructions or the performance of the gait tasks in the laboratory and real life, or orthopedic limitations (eg, severe arthrosis or lower limb fractures or hip/knee

replacements) that functionally affect gait. In addition, we recruited 34 healthy controls (HC) (age: 44.08 ± 14.78 years). Healthy participants had no history of any neurological or psychiatric disease, no family history of neurodegenerative disease, and did not

show any neurological signs on clinical examination. Participants were analyzed cross-sectionally at baseline and, where available, longitudinally at 1-year and 2-year follow-ups.

Clinical Outcome Assessments and Patient-Reported Outcomes

The severity of ataxia was rated using the SARA.⁴⁰ The three items rating gait and posture are grouped by the subscore SARA posture and gait (SARA_{p&g}).^{14,44} SARA assessments were performed by expert ataxia neurologists (M.S., L.S., and D.T.) To capture the impact of disease on daily living, DCD participants were asked to self-report their balance confidence in activities important in daily living using the activity-specific balance confidence scale (ABC).⁴⁵

Standard Protocol Approvals, Registration, and Patient Consent

The experimental procedure was approved by the local ethics committee (598/2011BO1; 303/2008BO2). All participants gave their informed consent before participation.

Gait Conditions

Walking movements were recorded in two different conditions. (1) Lab-based walking (LBW), where participants walked 60 m straight on a 30 m indoor floor (ie, including one turn) at their preferred self-selected speed on a pre-specified straight route in an institutional setting, supervised without any distractions. The turn and one stride before and after the turn were excluded from the analysis. (2) RLW, which is unconstrained walking during participants' usual individual everyday living, where participants were free to move how they wanted and were used to in their daily life, without supervision by any study personnel (total recording time: 4–6 hours within 2–3 days). Participants were instructed to wear the sensors inside and outside their homes and include at least a half-hour walk. Participants were instructed to wear the sensors for consecutive recording sessions, each lasting a maximum of 2 hours. Participants documented their recorded walking movements in an activity protocol (for details see Supplement S1).

Movement Recordings and Gait Measures

Three Opal inertial sensors (APDM, Portland, OR) were attached on both feet and posterior trunk at the level of L5 with elastic Velcro bands. Inertial sensor data was collected and wirelessly streamed to a laptop for the automatic generation of gait and balance metrics by Mobility Lab software (APDM). For the real-life condition, data was logged on board of each OPAL sensor and downloaded after the session. Step

events and spatio-temporal measures were extracted using Mobility Lab (Version 2),⁴⁶ which has been shown to deliver good-to-excellent accuracy and repeatability.^{47,48}

From the rich source of gait measures, we adopted a hypothesis-driven approach, focusing on those measures that have been considered promising candidate gait measures in previous work. Recent longitudinal studies from our group in laboratory-based gait analysis identified 1-year longitudinal gait change in a SCA3¹⁹ and an early SCA2 cohort,²⁰ revealing stride length variability (StrideL_{CV}), lateral body sway, lateral step deviation (LatStepDev) and a compound measure of spatial step variability (SPCmp, combining StrideL_{CV} and LatStepDev)²⁷ (see Supplementary S2) as most sensitive to longitudinal change in ataxia severity. In RLW, we showed that LatStepDev and SPCmp were sensitive to the cross-sectional ataxia severity,²⁷ and to short-term treatment-related improvements in SCA27B.⁴⁹

In addition, we included speed as an indicator of functional mobility and the variability of the lateral angle of the foot during the stance phase (ToeOutAngle_{SD}, inspired by results in Shah et al¹²). The lateral sway was determined by the coronal range of motion measured by the lumbar sensor (CorRoM_{SD}). StrideL_{CV} was determined using the coefficient of variation (CV) = σ/μ , normalizing the standard deviation (SD) with the mean value.

Selected Walking Bouts and Matching of RLW Behavior

The analysis focused on walking bouts >15 strides, as gait variability measures in short walking bouts are often estimated inaccurately.²⁶ The first and last two strides of each bout were removed to reduce the effects of gait initiation and termination on the variability measures.⁵⁰ A walking bout is here defined as a sequence of strides that is not interrupted by a turn or a complete halt of at least two median gait cycle durations of the individual. Finally, bouts that showed a jump in gait speed of at least 0.5 m/s between two strides were discarded.

Matching of RLW Behavior Based on Macroscopic Gait Characteristics

Because of the described influence of contextual and environmental factors on macroscopic gait characteristics,^{30,31} we used a linear regression model to confirm that gait variability measures like LatStepDev are influenced by bout length and the number of turns (number of turns, 60 seconds before and after the bout, see Supplement S3).

Second, we introduced a matching procedure for identifying for each subject comparable walking bouts in longitudinal assessments. This matching procedure compares walking bouts and identifies matching pairs

by optimizing a cost function with the two macroscopic gait descriptors: bout length and number of turns.

For matching, an Euclidean distance matrix D was created for each subject using bout length and number of turns for each baseline bout to each follow-up bout (see Fig. 1A). To weight both criteria equally, the data were previously standardized per subject (centered and normalized to one SD of the participants' baseline measurement). Using the distance matrix D for each subject as a cost matrix, a linear assignment problem was solved to obtain a 1:1 matching of similar bouts for baseline and follow-up. A cost for non-assignment of 0.5 was set, to obtain bouts that were as similar as possible, but at the same time to avoid losing any participants for the analysis. Therefore, for each subject, only bouts with similar macroscopic gait descriptors are compared. Mismatched bouts are not included in the analysis, reducing purely contextual longitudinal differences.

To determine test-retest reliability in everyday life, we used the matching procedure described above. Here,

the two baseline measurement days with the most bouts were selected and matched. For the LBW, we divided the 60 m task into two 30 m segments (before and after the turn) and calculated the split-half reliability (see Statistics). The test-retest reliability of gait measures was calculated using the intraclass correlation (ICC) (2, 1).⁵¹ ICC values <0.5, between 0.5 and 0.75, between 0.75 and 0.9, and >0.90 were considered as poor, moderate, good, and excellent reliability, respectively.⁵¹ Based on the ICC the minimum detectable change (MDC) was calculated, which is critical in determining whether a change can be reliably detected or is lost in the measurement noise.⁵²

$$\text{MDC}_{90} = 1.65 \times \text{SD}_{\text{baseline}} \times (\sqrt{2[1 - \text{ICC}]})$$

With 1.65 is the z-score of 90% level of confidence. Please note that we used a conservative method for calculating the MDC, which is suitable for individual measures. For groups, the original term can be divided by \sqrt{n} , with (n = group size).^{53,54}

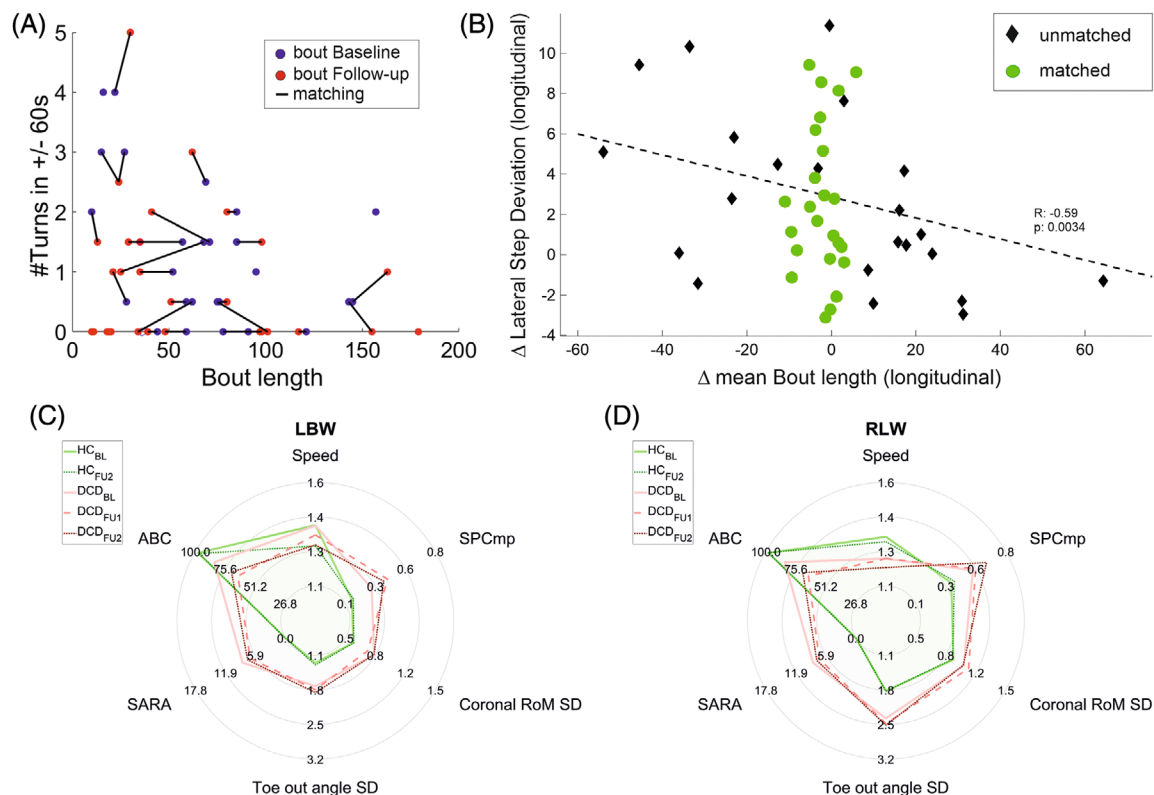


FIG. 1. (A) Illustration of the mapping procedure of corresponding real-life walking (RLW) trials at baseline (blue) and the 1-year follow-up (red). (B) Relationship between longitudinal differences (Δ = 1-year follow-up – baseline) differences in mean bout length and lateral step deviation for the degenerative cerebellar disease (DCD) participants (non-matched: red; matched: blue). In the non-matched condition, there is a significant correlation between the longitudinal difference in lateral step deviation and the difference in mean bout length ($r = 0.59$, $P = 0.0034^{**}$). (C,D) Radar plots illustrating cross-sectional and longitudinal differences on Scale for the Assessment and Rating of Ataxia, Activity-specific Balance Confidence scale, and four gait parameters for the gait conditions laboratory-based walking LBW (C) and RLW (D): speed, spatial variability compound measure (SPCmp), coronal range of motion variability (CorRoM_{SD}), and foot angle variability (ToeOutAngle_{SD}). Cross-sectional differences can be seen by comparison of healthy controls at baseline (HC_{BL}) and the DCD group at baseline (DCD_{BL}). Given are average values for each group. Longitudinal progression can be seen comparing DCD participants at baseline (DCD_{BL}) and follow-up (FU) assessments (DCD_{FU1}, DCD_{FU2}) as well as for healthy controls (HC_{BL} and HC_{FU2}). [Color figure can be viewed at wileyonlinelibrary.com]

Statistics

Between-group differences (DCD vs. HC) were determined using the non-parametric Mann–Whitney *U* test with effect sizes calculated by Cliff's Δ .⁵⁵ Repeated measurement analyses were performed for longitudinal analyses using the non-parametric Friedman-Test to determine within-group differences between assessments. When the Friedman-Test yielded an effect ($P < 0.1$), post hoc analysis was performed using a Wilcoxon-signed-rank-test for pairwise comparisons. Effect sizes r_{prb} for the repeated measurements analyses were determined by matched-pairs rank biserial correlation.⁵⁶

We report three significance levels: (1) uncorrected $*P < 0.05$; (2) Bonferroni-corrected for multiple comparisons $**P < 0.05/n = 6$: number of analyzed features; and (3) $***P < 0.001$. Spearman's ρ was used to examine the correlation between movement measures and SARA, between gait measures for different walking conditions, and between gait measures for matched and unmatched walking bouts. Effect sizes ρ were classified as ρ : 0.1 small effect, 0.3 medium effect, 0.5 large effect, 0.7 very large effect.⁵⁷ Statistical analysis was performed using MATLAB (The MathWorks, version R2024A, Natick, MA). Based on the effect sizes of longitudinal change of the gait measures and the SARA, a sample size estimation was performed using G*power 3.1⁵⁸ to determine the required cohort size for detecting a 50% reduction of progression by a hypothetical intervention.

Results

Real-life gait data was collected from 26 DCD participants at baseline (253 ± 206 minutes recorded within 2.1 ± 1.3 days), from 23 DCD at the 1-year follow-up after 380 ± 52 days (follow-up 1 [FU1], 227 ± 147 minutes recorded within 2.2 ± 1.1 days) and from 22 DCD at the 2-year follow-up after 816 ± 106 days (follow-up 2 [FU2], 326 ± 234 minutes recorded within 2.5 ± 1.1 days).

From the collected RLW gait data, we extracted 92,117 valid strides in 2854 bouts (length: 47.2 ± 32.6 strides). After the described matching procedure, 53,178 strides in 894 bouts (length: 58.6 ± 63.9 strides) were analyzed. For comparison, the amount of gait data extracted from LBW was in total 1021 strides in 52 bouts (length: 24.7 ± 4.2 strides).

Sensitivity of Gait Measures to Ataxia Severity: Cross-Sectional Results

Cross-sectional analyses revealed group differences between DCD versus HC for several examined gait variability measures in both walking conditions, constrained LBW (StrideL_{CV}: $P = 0.002^{**}$; LatStepDev: $P = <0.001^{***}$; SPCmp: $P = <0.001^{***}$) and RLW

(LatStepDev: $P = <0.001^{***}$; SPCmp: $P = <0.001^{***}$) (see Table 2, Fig. 1C,D). High sensitivity of gait measures StrideL_{CV}, LatStepDev, and SPCmp to cross-sectional ataxia severity was indicated by correlations with the SARA and the SARA_{p&g} subscore, with large effect size in both conditions $\rho > 0.75$. In addition, gait measures revealed high correlations with the patient-reported balance confidence in activities of daily living, assessed by the ABC (StrideL_{CV}, LatStepDev, SPCmp: $P < 0.001^{***}$) (Table 2). Effect sizes were higher for RLW than LBW (SPCmp: $\rho = -0.81$ RLW vs. $\rho = -0.71$ LBW).

Gait Measures for Matched Walking Bouts Show Good-to-Excellent Test–Retest Reliability

To identify suitable macroscopic gait characteristics for matching walking bouts, linear regression analysis showed significant contributions of both, bout length and number of turns in explaining the variability for walking bouts of HC (see Supplement S3). Performing the presented matching procedure on the real-life baseline assessment (see Patient and Methods), revealed good to excellent test–retest reliability (ICC) for several gait measures like LatStepDev, StrideL_{CV}, and SPCmp (ICC (1, 2) ≥ 0.82) (see Table 2).

Sensitivity of Gait Measures to Longitudinal Change

Longitudinal analyses revealed for the SARA significant changes only in the second follow-up (see Table 3). The patient-reported outcome measure of balance confidence (ABC score) as well as a few laboratory-based gait measures revealed significant changes after 1 year with modest effect sizes (ABC: FU1: $P = 0.03^{*}$, $r_{prb} = -0.52$; speed: FU1: $P = 0.04^{*}$, $r_{prb} = -0.47$; CorRoM_{SD}: FU1: $P = 0.03^{*}$, $r_{prb} = 0.51$), whereby the latter grow substantially after 2 years (eg, speed: FU2: $P = 0.0005^{***}$, $r_{prb} = -0.84$; StrideL_{CV}: FU2: $P = 0.005^{**}$, $r_{prb} = 0.68$) (Table 3).

In contrast, in condition RLW, several gait measures—in particular StrideL_{CV} and SPCmp—revealed significant changes with high effect sizes already after 1 year (StrideL_{CV}: FU1: $P = 0.0047^{**}$, $r_{prb} = 0.67$; SPCmp: FU1: $P = 0.0007^{***}$, $r_{prb} = 0.80$). The robustness of these results is supported by the second follow-up (SPCmp: FU2: $P = 0.016^{*}$, $r_{prb} = 0.77$) (Figure 2A).

Moreover, these longitudinal changes after 1 year for the gait measures StrideL_{CV} and SPCmp can be observed and validated in the subgroup SCA_{1/2/3} with higher effect size (FU1: $n = 18$; FU2: $n = 18$) (SPCmp: FU1: $P = 0.0023^{**}$, $r_{prb} = 0.86$; FU2: $P = 0.0067^{**}$, $r_{prb} = 0.76$) (Supplement S4). For both the DCD group and the SCA_{1/2/3} subgroup, these annual changes were larger than the MDC (see Tables 3 and Supplement-S4)

TABLE 2 Cross-sectional analyses: Between-group differences of HC and DCD participants for clinician reported outcomes, patient-reported outcomes, and gait measures in the LBW and RLW conditions at baseline assessment.

		Group difference DCD vs. HC		Correlations DCD				ICC
				SARA		ABC		
		Measure	P	δ	ρ	P	ρ	
Clinician-reported outcomes	SARA	<0.001***	0.99	–	–	0.76	0.001**	–
	SARA _{p&g}	<0.001***	0.75	–	–	0.73	0.002**	–
Patient-reported outcome	ABC	<0.001***	0.70	–	–	–	–	–
Gait–LBW	Speed	0.67	0.06	0.03	0.88	–0.22	0.45	0.94
	StrideL _{CV}	0.002**	0.47	0.79	<0.001***	–0.55	0.04*	0.73
	LatStepDev	<0.001***	0.64	0.81	<0.001***	–0.71	0.005**	0.91
	SPCmp	<0.001***	0.60	0.81	<0.001***	–0.71	0.005**	0.83
	CorRoM _{SD}	<0.001**	0.54	0.74	<0.001***	–0.72	0.004**	0.75
	ToeOutAng _{SD}	<0.001**	0.62	0.64	<0.001***	–0.70	0.006*	0.49
Gait–RLW	Speed	0.19	0.2	–0.47	0.02*	0.05	0.86	0.83
	StrideL _{CV}	0.06	0.29	0.75	<0.001***	–0.35	0.23	0.84
	LatStepDev	<0.001***	0.61	0.77	<0.001***	–0.81	<0.001***	0.82
	SPCmp	<0.001***	0.54	0.78	<0.001***	–0.81	<0.001***	0.85
	CorRoM _{SD}	0.13	0.23	0.59	0.002**	–0.62	0.02*	0.84
	ToeOutAng _{SD}	0.008**	0.41	0.59	0.002**	–0.27	0.36	0.79

δ Denotes the effect sizes determined by Cliff's δ . Correlations between gait measures and clinician-reported ataxia severity (SARA, SARA_{p&g}) and patient-reported outcomes (ABC) are given for the DCD group. Effect sizes of correlations are given using Spearman's ρ . Test–retest reliability is analyzed by determining the ICC (see Patients and Methods).

Indicates significant differences between groups ($\equiv P < 0.05$, ** $\equiv P < 0.0083$ Bonferroni-corrected, *** $\equiv P < 0.001$).

Abbreviations: HC, healthy controls; DCD, degenerative cerebellar disease; LBW, lab-based waling; RLW, real-life walking; ICC, intraclass correlation coefficient; SARA, Scale for the Assessment and Rating of Ataxia; ABC, Activity-specific Balance Confidence scale; SARA_{p&g}, Scale for the Assessment and Rating of Ataxia posture and gait; m, mean; SD, standard deviation; StrideL_{CV}, stride length variability; LatStepDev, lateral step deviation; SPCmp, compound measure of spatial variability; CorRoM_{SD}, coronal range of motion measured by the lumbar sensor; ToeOutAng_{SD}, foot angle variability.

(eg, SPCmp: Δ_{BL} vs. FU1 = 0.12 > MDC₉₀ = 0.02). In the HC group, longitudinal changes in the RLW condition were observed in gait speed, but not in the ataxia-specific measures of spatio-temporal variability (see Fig. 1C,D).

Based on observed effect sizes in the DCD population, sample size estimation for SPCmp showed for the RLW condition a required cohort size of $n = 42$ for detecting a 50% reduction of progression after 1 year by a hypothetical intervention (90% power and one-sided 5% type I error) in comparison to $n = 147$ for the SARA (Fig. 2B, Table 3). For the SCA_{1/2/3} subgroup, because of the larger effect size, the required cohort size was even smaller ($n = 38$). In the LBW condition, the required cohort size for SPCmp was $n = 39$ after 2 years (Fig. 2B, Table 3).

Compared to the matched walking trials in RLW, analysis of the non-matched walking trials (bout length > 15) showed significant longitudinal differences, yet with smaller effect sizes (SPCmp: FU1:

$P = 0.0089^*$, $r_{prb} = 0.62$; FU2: $P = 0.01^*$, $r_{prb} = 0.58$) (Supplement-S5). Moreover, the analysis of the non-matched walking bouts revealed a correlation between longitudinal differences in variability measures like StrideL_{CV} and longitudinal differences in mean bout length ($r = 0.59$, $P = 0.0034^{**}$), potentially influencing the observed changes. Such correlations were not observed in the matched analysis (Figure 1B).

Discussion

This study aimed to test the hypothesis that the progression of ataxia-related gait impairments can be reliably and sensitively captured in real-life walking by analysis of walking bouts, which are matched according to macroscopic gait characteristics reflecting environmental factors. We showed that real-life gait measures can capture longitudinal change within short trial-like time frames like 1 year with high effect size, therefore,

TABLE 3 Longitudinal within-subject comparison for the group of DCD.

Group	Friedman test		Baseline		1-year follow-up				2-year follow-up					
	DCD	Measure	χ^2	P	M \pm SD	M \pm SD	P	r _{pb}	ESS _{50%}	M \pm SD	P	r _{pb}	ESS _{50%}	MDC ₉₀
Clinician-reported outcomes		SARA	1.42	0.49	8.0 \pm 4.5	8.1 \pm 4.5	0.66	0.12	–	9.1 \pm 5.5	0.02*	0.50	147	–
		SARA _{p&g}	3.04	0.21	3.1 \pm 2.1	3.0 \pm 2.1	0.79	0.17	–	3.5 \pm 2.5	0.03*	0.53	179	–
Patient-reported outcome		ABC	1.87	0.39	75.6 \pm 20.5	67.7 \pm 23.4	0.03*	–0.52	73	68.7 \pm 31.2	0.12	–0.51	–	–
Gait-LBW		Speed	12.0	0.001 ⁺	1.39 \pm 0.14	1.35 \pm 0.12	0.04*	–0.47	84	1.26 \pm 0.19	0.0005***	–0.84	66	0.019 ^{<FU1+2}
		Stridel _{CV}	1.9	0.38	0.02 \pm 0.007	0.023 \pm 0.016	0.30	0.24	–	0.029 \pm 0.02	0.005**	0.68	67	0.0065 ^{<FU2}
		LatStepDev	3.4	0.17	0.032 \pm 0.011	0.034 \pm 0.013	0.44	0.18	–	0.036 \pm 0.015	0.04*	0.47	51	0.003 ^{<FU2}
		SPCmp	6.0	0.04 ⁺	0.38 \pm 0.21	0.41 \pm 0.21	0.27	0.26	–	0.47 \pm 0.23	0.04*	0.48	39	0.04 ^{<FU2}
		CorRoM _{SD}	6.5	0.03 ⁺	0.87 \pm 0.34	0.99 \pm 0.34	0.03*	0.51	83	0.91 \pm 0.33	0.7	0.11	–	0.2
		ToeOutAng _{SD}	0.6	0.7	1.81 \pm 0.5	1.79 \pm 0.61	0.54	–0.14	–	2.00 \pm 0.72	0.13	0.36	–	0.54
Gait-RLW		Speed	4.1	0.12	1.23 \pm 0.14	1.17 \pm 0.13	0.09	–0.39	–	1.15 \pm 0.2	0.009**	–0.63	50	0.048 ^{<FU1+2}
		Stridel _{CV}	5.5	0.06 ⁺	0.034 \pm 0.012	0.042 \pm 0.016	0.0047**	0.67	41	0.048 \pm 0.03	0.04*	0.49	69	0.0025 ^{<FU1+2}
		LatStepDev	5.1	0.07 ⁺	0.044 \pm 0.009	0.047 \pm 0.01	0.05	0.44	45	0.048 \pm 0.008	0.006**	0.66	43	0.0021 ^{<FU1+2}
		SPCmp	10.9	0.004 ⁺	0.58 \pm 0.17	0.67 \pm 0.2	0.0007***	0.80	42	0.72 \pm 0.31	0.001**	0.78	54	0.02 ^{<FU1+2}
		CorRoM _{SD}	7.6	0.02 ⁺	1.0 \pm 0.21	1.09 \pm 0.2	0.006**	0.65	43	1.07 \pm 0.24	0.18	0.32	–	0.13
		ToeOutAng _{SD}	3.37	0.18	2.35 \pm 0.57	2.56 \pm 0.59	0.04*	0.47	62	2.68 \pm 0.9	0.12	0.37	–	0.22 ^{<FU2}

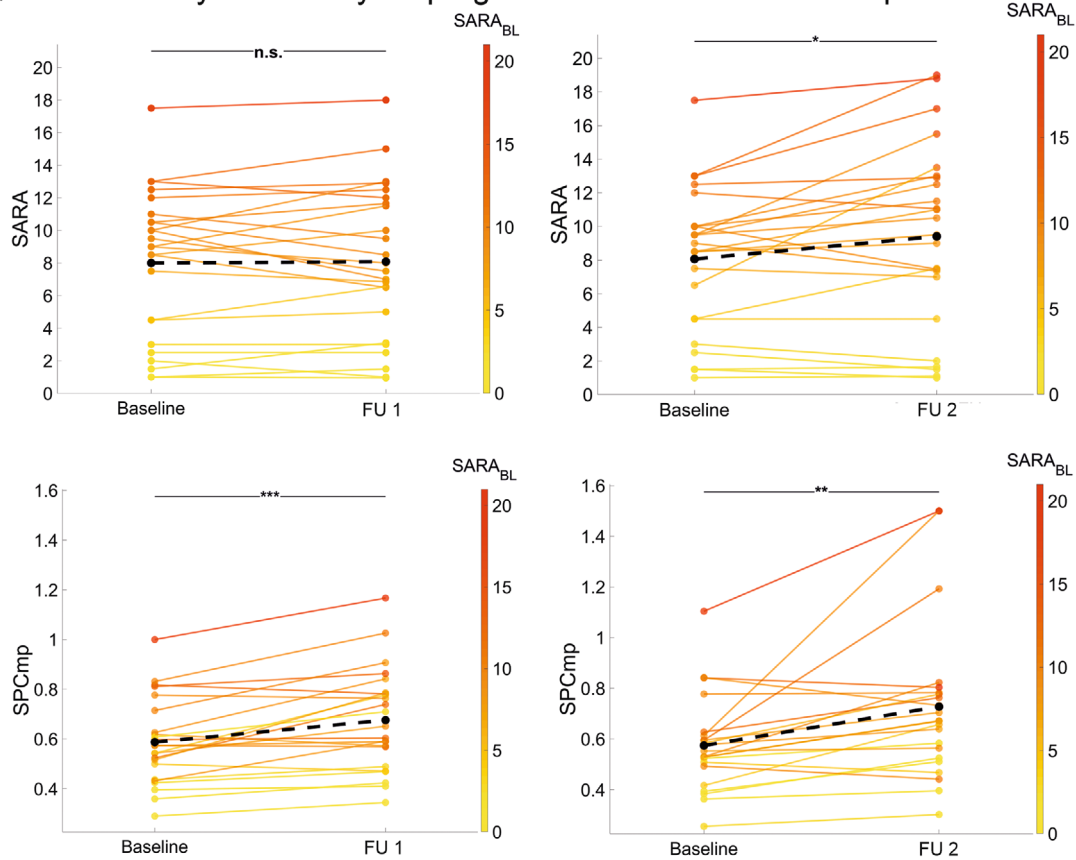
Shown are results of clinician-reported ataxia ratings (SARA score and SARA_{p&g} subscore⁴¹) as well as of gait measures in clinical assessment LBW and RLW for baseline, 1-year, and 2-year follow-up assessments. Friedman test determined within-group longitudinal differences (**P* < 0.1). Post hoc test *P*-values determined by Wilcoxon signed-rank test for both follow-up assessments relative to baseline. Effect sizes *r*_{pb} determined by matched-pairs rank-biserial correlation.⁵⁶ Shown are analyses for the whole group of DCD, both consisting of pre-ataxic and ataxic participants. For significant longitudinal changes, we determined the sample size estimate for a cohort required to detect a 50% reduction in progression ESS_{50%}. MCD₉₀ values denote the smallest reliable detectable change (90% confidence interval). Superscripts indicate whether the MDC is smaller than the change between baseline and 1-year follow-up (<²ESS_{50%}) or/and between baseline and 2-year follow-up.

*Indicate significant differences between groups (**P* < 0.05, ***P* < 0.0083 Bonferroni-corrected, ****P* < 0.001). Abbreviations: DCD, degenerative cerebellar disease; MDC, minimum detectable change; M, mean; SD, standard deviation; SARA, Scale for the Assessment and Rating of Ataxia; SARA_{p&g}, Scale for the Assessment and Rating of Ataxia posture and gait; ABC, Activity-specific Balance Confidence scale; LBW, lab-based walking; StrideL_{CV}, stride length variability; LatStepDev, lateral step deviation; SPCmp, compound measure of spatial variability; CorRoM_{SD}, coronal range of motion measured by the lumbar sensor; ToeOutAng_{SD}, foot angle variability; RLW, real-life walking.

outperforming in their sensitivity both clinician-reported outcomes like the SARA and lab-based gait measures. These results indicate an increased sensitivity

and ecological validity for real-life gait measures as promising performance outcomes in upcoming treatment trials.

(A) Individual 1-year and 2-year progression of SARA and SPCmp in RLW



(B) Sample size estimation for future treatment trials

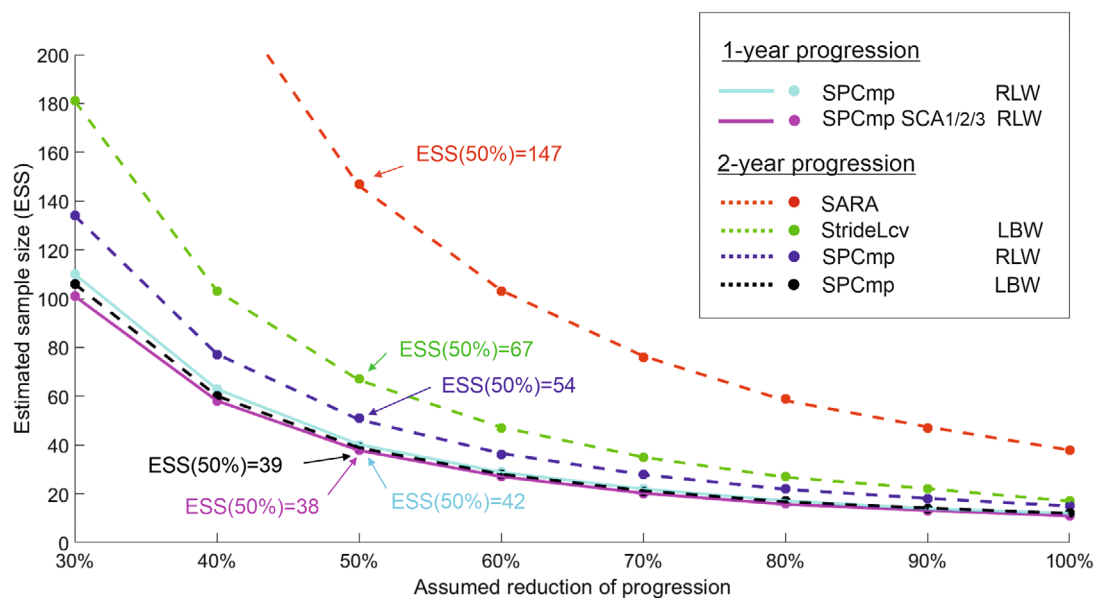


FIG. 2. Legend on next page.

Increased Spatio-Temporal Gait Variability as a Consistent Feature of Ataxic Gait in Laboratory-Based Assessments

Our findings in the constrained walking condition LBW confirm the results of previous studies by our group and others using different motion capture technologies.^{8,9,12,13,19} These studies showed that spatio-temporal variability measures serve as reliable and valid measures for cerebellar ataxia in constrained walking conditions, and correlate with gait and posture ataxia severity and with patients' subjective balance confidence in important activities of daily living (ABC score).

Measures of Ataxic Gait in Real Life: Cross-Sectional Sensitivity to Ataxia Severity

Although real-life gait is inherently more variable in both HC and cerebellar patients,³⁹ several of our gait variability measures (eg, StrideL_{CV}, LatStepDev, and SPCmp), allow to capture the ataxia-specific gait variability in straight walking episodes of real life.²⁷ The compound measure SPCmp—integrating variability in the anterior–posterior and in the mediolateral dimension—hereby seems to benefit from capturing individual differences in gait variability.²⁷

Consistent with previous studies in lab-based gait assessments, cross-sectional sensitivity of real-life gait measures (eg, StrideL_{CV}, LatStepDev, and SPCmp) for ataxia severity was shown by the high correlation with clinical severity of ataxia ($P < 0.002^{**}$) (see Table 2).

Even more meaningful for patient relevance and ecological validity, however, is the high correlation between gait measures and patients' subjective balance confidence in daily living, as assessed by the ABC score. The high effect sizes—even higher for real-life than lab-based gait assessments (SPCmp: RLW: $\rho = 0.81$, LBW: $\rho = 0.71$)—emphasize the relevance of ataxia-related gait measures to patients' everyday life—which is key to FDA-conform patient-focused outcome and drug development.²⁹

Matching Walking Bouts According to Bout Length and Number of Turns

The influence of contextual and environmental factors on gait measures during real-life walking is

currently under intense investigation in various movement disorders.^{25,34,36,38} To date, matching or selection procedures based on contextual factors or macroscopic gait characteristics have been used exclusively in cross-sectional studies, for example, to compare a group of patients with HC in activity monitoring and real-life walking, to differentiate patient subgroups,²⁵ or to compare patients' real-life walking behavior with clinical gait assessments.^{38,59} Moreover, in most approaches, the selection of walking sessions was based on bout length alone.

In contrast, the focus of our study was to investigate a matching approach as a novel strategy for longitudinal change analysis, allowing to identify comparable walking bouts at baseline and follow-up visits. As turns are an important part of real-life walking behavior, and typically differ between indoor and outdoor walking,³² we included the number of turns and the bout length in our macroscopic characterization of walking behavior to identify comparable walking bouts. The number of turns contributes to the macroscopic walking characteristics in explaining step variability, as observed by linear regression analysis (Supplement S3). For the longitudinal analyses, the matching procedure ensures that observed effects are not predominantly because of differences in bout length, which is important to avoid false positive results regarding the progression of the gait variability measures. Using the proposed matching procedure, we showed good-to-excellent test–retest reliability for real-life recordings on different days.

In summary, this approach allows comparison of real-life gait assessments without requiring participants to perform similar gait behavior in all home assessments or capturing home data over a long period (where contextual differences can be assumed to average out). At the same time, it highlights the need for other longitudinal gait studies to control for changes in macroscopic walking characteristics like bout length and the number of turns, as potential confounders.

Gait Measures in RLW Capture Longitudinal Change with Higher Effect Size

Using these matching procedures, and consistent with the cross-sectional results, measures of spatio-temporal variability in real-life walking (and in particular the

FIG. 2. (A) Within-subject changes between baseline and follow-up assessments for the group of degenerative cerebellar disease (DCD) participants. Shown are longitudinal differences in Scale for the Assessment and Rating of Ataxia (SARA) (upper panel) and the spatial variability compound measure (SPCmp) (lower panels). For both measures, the left panel shows the change between baseline and 1-year follow-up (FU1), and the right panel the change between baseline and 2-year follow-up (FU2). In all panels, SARA scores of individual cerebellar participants are color-coded. Black dotted line = mean change across all participants. Stars indicate significant differences between time points (* $\equiv P < 0.05$, ** $\equiv P < 0.0083$ Bonferroni-corrected, *** $\equiv P < 0.001$). (B) Sample size estimates were made for future treatment trials that showed different levels of progression reduction for the various outcome measures: SARA as well as the gait measures SPCmp and stride length variability (StrideL_{CV}) in the laboratory (LBW) and real life (RLW). The estimated number of participants per study arm is plotted against the hypothesized therapeutic effect on reducing 1-year progression or 2-year progression in DCD patients or the SCA_{1/2/3} subgroup (SCA_{1/2/3}), respectively. Solid lines indicate 1-year estimates, and dashed lines indicate 2-year estimates. [Color figure can be viewed at wileyonlinelibrary.com]

compound measure SPCmp) show high responsiveness to change at 1 year. Their effect size in longitudinal sensitivity to change outperforms the clinician-reported outcome (SARA) and laboratory gait measures.

Importantly, the high test–retest reliability under real-life conditions resulted in a longitudinal change that was higher than the MDC (see Table 3), which is a critical requirement for reliable detection in clinical trials.⁶⁰

The subgroup SCA_{1/2/3} showed higher effect sizes and, consequently smaller sample size estimates. This finding is consistent with previous studies showing faster progression in SCA 1, 2, 3⁶¹ compared to a cross-genotype DCD population, which includes slower progressing DCD types (eg, SCA6⁶² and non-PolyQ SCAs).⁶³

The reduction in sample size inferred by these digital performance outcomes could be decisive for the feasibility of a treatment trial: whereas trials with, for example, 147 SCA participants per trial arm (as required for SARA as outcome for 2 years) are almost impossible, 38 SCA participants (as required for the gait measure SPCmp in SCA_{1/2/3} for 1 year) (Figure 2B) are well feasible.

The potential of real-life gait assessments is also supported by the comparison of the results presented here with a previous study on laboratory-based gait assessment in an SCA2 population, which found significant changes in laboratory-based gait measures.²⁰ Although the smaller number of SCA patients and the addition of slower progressing DCD types in the current study resulted in a non-significant change in laboratory-based scores after 1-year, real-life scores showed a significant change.

Limitations

Overall, the present study aimed to explore and longitudinally validate digital performance measures in real-life ataxic gait with a cross-genotype cohort, as this approach allows validation across DCDs. This approach was based on the assumption that our measures would capture functional impairment generically across DCDs, given that they qualitatively affect the same ataxia-related functions. This assumption is corroborated by our previous work demonstrating validity of ataxia-specific gait measures across various DCDs^{20,49,64} and by the comparable longitudinal results for the DCD cohort and SCA_{1/2/3} subgroup observed in the current study. However, given that different DCD genotypes gradually differ in their specific progression rates, the sensitivity of digital gait measures to detect longitudinal changes formally remains to be demonstrated within specific genotypes and pre-ataxic populations only. Furthermore, additional work should investigate longer—completely uninstructed—real-life

assessments to (1) include other important components of walking behavior (eg, turning,²³ gait initiation,⁶⁵ and termination⁶⁶) in addition to straight walking episodes; and (2) quantify daily fluctuations in patients' walking behavior.

Multicenter studies need to further investigate the potential of real-life gait assessments regarding their increased sensitivity and relevance,^{24,67} but also consider multicenter challenges such as the necessity for patient compliance, the impact of walking aids,⁶⁰ and the more complex data analysis compared to lab-based gait assessments.

Conclusion

This study unraveled methods and measures that allow quantifying longitudinal changes in real-life ataxic gait in early to moderate disease stages with high effect sizes and high correlations with patient-reported outcomes of daily living. This provides promising ecologically valid, patient-centered outcome measures for natural history and treatment trials in degenerative cerebellar ataxias. In addition to the higher effect sizes gained from real-life assessments, these measures allow for the objective quantification of patients' real-life gait performance—instead of clinical assessment in partly artificial settings, for example, by gait tasks as part of clinical scores or under lab conditions, which serve as surrogate parameters at best.

In conclusion, measures of real-life gait performance add ecological validity, and therefore, help to inform upcoming treatment trials in cerebellar ataxias and FDA-compatible development of patient-focused outcomes and approval of novel treatments.^{21,28} ■

Author Roles: (1) Research project: A. Conception, B. Organization, C. Execution; (2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; (3) Manuscript Preparation: A. Writing of the First Draft, B. Review and Critique.

J.S.: 1A, 1B, 1C, 2B, 3A

T.B.: 1C, 3B

N.J.: 1C, 3B

F.H.: 1C, 3B

M.G.: 1B, 1C, 3B

L.S.: 1B, 1C, 3B

D.T.: 1B, 1C, 3B

M.S.: 1A, 2C, 3A

W.I.: 1A, 1B, 1C, 2A, 2B, 3A

Acknowledgment: Open Access funding enabled and organized by Projekt DEAL.

Financial Disclosure: W.I. received consultancy honoraria by Ionis Pharmaceuticals, unrelated to the present work. J.S., T.B., N.J., F.H., M.G., and D.T. report no disclosures. L.S. served as advisor for Alexion, Novartis, and Vico. He participates as a principal investigator in clinical studies sponsored by Vigil Neuroscience (VGL101-01.001; VGL101-01.002), Vico Therapeutics (VO659-CT01), PTC Therapeutics (PTC743-NEU-003-FA), and Stealth BioTherapeutics (SPIMD-301), all unrelated to the present work. M.S. has received consultancy honoraria from Ionis, UCB, Prevail, Orphazyme, Biogen, Servier, Reata, GenOrph, AviadoBio, Biohaven, Zevra, Lilly, and Solaxa, all unrelated to the present manuscript.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

- Klockgether T, Ashizawa T, Brais B, et al. Paving the way toward meaningful trials in ataxias: an ataxia global initiative perspective. *Mov Disord* 2022;37:1125–1130.
- Ashizawa T, Oz G, Paulson HL. Spinocerebellar ataxias: prospects and challenges for therapy development. *Nat Rev Neurol* 2018;14:590–605.
- Globas C, du Montcel ST, Baliko L, et al. Early symptoms in spinocerebellar ataxia type 1, 2, 3, and 6. *Mov Disord* 2008;23:2232–2238.
- Ilg W, Branscheidt M, Butala A, et al. Consensus paper: neurophysiological assessments of ataxias in daily practice. *Cerebellum* 2018;17:628–653.
- Luo L, Wang J, Lo RY, et al. The initial symptom and motor progression in spinocerebellar ataxias. *Cerebellum* 2017;16:615–622.
- NAF. The voice of the patient: Friedreich's Ataxia. Summary report resulting from an externally led patient-focused drug development meeting; <https://www.ataxia.org/ataxiapfdd/2017>.
- Rosen A, Hagen S, Trace K, Compton A. The Voice of the Patient: Living with Polglutamine Spinocerebellar Ataxias (SCA) and Dentatorubal-Pallidolysion Atrophy (DRPLA). <https://www.ataxia.org/ataxiapfdd/>. Golden Valley, MN, USA: National Ataxia Foundation; 2021.
- Ilg W, Golla H, Thier P, Giese MA. Specific influences of cerebellar dysfunctions on gait. *Brain* 2007;130:786–798.
- Serrao M, Pierelli F, Ranavolo A, et al. Gait pattern in inherited cerebellar ataxias. *Cerebellum* 2012;11:194–211.
- Schniepp R, Wuehr M, Schlick C, et al. Increased gait variability is associated with the history of falls in patients with cerebellar ataxia. *J Neurol* 2014;261:213–223.
- Zhou H, Nguyen H, Enriquez A, et al. Assessment of gait and balance impairment in people with spinocerebellar ataxia using wearable sensors. *Neurol Sci* 2022;43:2589–2599.
- Shah VV, Rodriguez-Labrada R, Horak FB, et al. Gait variability in spinocerebellar ataxia assessed using wearable inertial sensors. *Mov Disord* 2021;36:2922–2931.
- Rochester L, Galna B, Lord S, Mhiripiri D, Eglon G, Chinnery PF. Gait impairment precedes clinical symptoms in spinocerebellar ataxia type 6. *Mov Disord* 2014;29:252–255.
- Ilg W, Fleszar Z, Schatton C, et al. Individual changes in preclinical spinocerebellar ataxia identified via increased motor complexity. *Mov Disord* 2016;31:1891–1900.
- Buckley E, Buckley C, Storm F, Mazza C, McNeill A. Spatiotemporal and Upper Body Variables to Quantify Gait Characteristics in Cerebellar Ataxia 8th World Congress of Biomechanics 2018.
- Milne SC, Murphy A, Georgiou-Karistianis N, Yiu EM, Delatycki MB, Corben LA. Psychometric properties of outcome measures evaluating decline in gait in cerebellar ataxia: a systematic review. *Gait Posture* 2018;61:149–162.
- Ilg W, Timmann D. Gait ataxia-specific cerebellar influences and their rehabilitation. *Mov Disord* 2013;28:1566–1575.
- Cabaraux P, Agrawal SK, Cai H, et al. Consensus paper: ataxic gait. *Cerebellum* 2023;22:394–430.
- Ilg W, Muller B, Faber J, et al. Digital gait biomarkers allow to capture 1-year longitudinal change in spinocerebellar ataxia type 3. *Mov Disord* 2022;37:2295–2301.
- Seemann J, Daghsen L, Cazier M, et al. Digital gait measures capture 1-year progression in early-stage spinocerebellar ataxia type 2. *Mov Disord* 2024;39(5):788–797.
- Manta C, Patrick-Lake B, Goldsack JC. Digital measures that matter to patients: a framework to guide the selection and development of digital measures of health. *Digit Biomark* 2020;4:69–77.
- Orendurff MS, Schoen JA, Bernatz GC, Segal AD, Klute GK. How humans walk: bout duration, steps per bout, and rest duration. *J Rehabil Res Dev* 2008;45:1077–1089.
- Thierfelder A, Seemann J, John N, et al. Real-life turning movements capture subtle longitudinal and Preataxic changes in cerebellar ataxia. *Mov Disord* 2022;37:1047–1058.
- Buckley C, Alcock L, McArdle R, et al. The role of movement analysis in diagnosing and monitoring neurodegenerative conditions: insights from gait and postural control. *Brain Sci* 2019;9:34.
- McArdle R, Del Din S, Donaghy P, Galna B, Thomas AJ, Rochester L. The impact of environment on gait assessment: considerations from real-world gait analysis in dementia subtypes. *Sensors (Basel)* 2021;21:813.
- Mico-Amigo ME, Bonci T, Paraschiv-Ionescu A, et al. Correction: assessing real-world gait with digital technology? Validation, insights and recommendations from the mobilise-D consortium. *J Neuroeng Rehabil* 2024;21:71.
- Ilg W, Seemann J, Giese M, et al. Real-life gait assessment in degenerative cerebellar ataxia: toward ecologically valid biomarkers. *Neurology* 2020;95:e1199–e1210.
- Byrom B, Watson C, Doll H, et al. Selection of and evidentiary considerations for wearable devices and their measurements for use in regulatory decision making: recommendations from the ePRO consortium. *Value Health* 2018;21:631–639.
- FDA. Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making.
- McArdle R, Morris R, Hickey A, et al. Gait in mild Alzheimer's disease: feasibility of multi-center measurement in the clinic and home with body-worn sensors: a pilot study. *J Alzheimers Dis* 2018;63:331–341.
- Del Din S, Yarnall AJ, Barber TR, et al. Continuous real-world gait monitoring in idiopathic REM sleep behavior disorder. *J Parkinsons Dis* 2020;10:283–299.
- Glaister BC, Bernatz GC, Klute GK, Orendurff MS. Video task analysis of turning during activities of daily living. *Gait Posture* 2007;25:289–294.
- Rehman RZU, Guan Y, Shi JQ, et al. Investigating the impact of environment and data aggregation by walking bout duration on Parkinson's disease classification using machine learning. *Front Aging Neurosci* 2022;14:808518.
- Roth N, Wieland GP, Kuderle A, et al. Do we walk differently at home? A context-aware gait analysis system in continuous real-world environments. *Annu Int Conf IEEE Eng Med Biol Soc* 2021;2021:1932–1935.
- Del Din S, Godfrey A, Galna B, Lord S, Rochester L. Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. *J Neuroeng Rehabil* 2016;13:46.
- Shah VV, McNames J, Harker G, et al. Effect of bout length on gait measures in people with and without Parkinson's disease during daily life. *Sensors* 2020;20:5769.
- Storm FA, Nair KPS, Clarke AJ, Van der Meulen JM, Mazza C. Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS One* 2018;13:e0196463.
- Carcreff L, Gerber CN, Paraschiv-Ionescu A, et al. Comparison of gait characteristics between clinical and daily life settings in children with cerebral palsy. *Sci Rep* 2020;10:2091.
- Tamburini P, Storm F, Buckley C, Bisi MC, Stagni R, Mazza C. Moving from laboratory to real life conditions: influence on the assessment of variability and stability of gait. *Gait Posture* 2018;59:248–252.
- Schmitz-Hubsch T, du Montcel ST, Baliko L, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. *Neurology* 2006;66:1717–1720.
- Subramony SH, May W, Lynch D, et al. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. *Neurology* 2005;64:1261–1262.

42. Jacobi H, Reetz K, du Montcel ST, et al. Biological and clinical characteristics of individuals at risk for spinocerebellar ataxia types 1, 2, 3, and 6 in the longitudinal RISCA study: analysis of baseline data. *Lancet Neurology* 2013;12:650–658.
43. Tezenas du Montcel S, Durr A, Rakowicz M, et al. Prediction of the age at onset in spinocerebellar ataxia type 1, 2, 3 and 6. *J Med Genet* 2014;51:479–486.
44. Lawerman TF, Brandsma R, Verbeek RJ, et al. Construct validity and reliability of the SARA gait and posture sub-scale in early onset ataxia. *Front Hum Neurosci* 2017;11:605.
45. Powell LE, Myers AM. The activities-specific balance confidence (ABC) scale. *J Gerontol A Biol Sci Med Sci* 1995;50A:M28–M34.
46. Mancini M, King L, Salarian A, Holmstrom L, McNames J, Horak FB. Mobility lab to assess balance and gait with synchronized body-worn sensors. *J Bioeng Biomed Sci* 2011;suppl 1:7.
47. Washabaugh EP, Kalyanaraman T, Adamczyk PG, Claffin ES, Krishnan C. Validity and repeatability of inertial measurement units for measuring gait parameters. *Gait Posture* 2017;55:87–93.
48. Morris R, Stuart S, McBarron G, Fino PC, Mancini M, Curtze C. Validity of mobility lab (version 2) for gait assessment in young adults, older adults and Parkinson's disease. *Physiol Meas* 2019;40: 095003.
49. Seemann J, Traschutz A, Ilg W, Synofzik M. 4-aminopyridine improves real-life gait performance in SCA27B on a single-subject level: a prospective *n*-of-1 treatment experience. *J Neurol* 2023;270: 5629–5634.
50. Kroneberg D, Elshehabi M, Meyer AC, et al. Less is more—estimation of the number of strides required to assess gait variability in spatially confined settings. *Front Aging Neurosci* 2019;10:435.
51. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15:155–163.
52. Beckerman H, Roebroek ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res* 2001;10:571–578.
53. Schmitz-Hubsch T, Fimmers R, Rakowicz M, et al. Responsiveness of different rating instruments in spinocerebellar ataxia patients. *Neurology* 2010;74:678–684.
54. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
55. Cliff N. Answering ordinal questions with ordinal data using ordinal statistics. *Multivar Behav Res* 1996;31:331–350.
56. Kerby DS. The simple difference formula: an approach to teaching nonparametric correlation. *Comprehensive Psychology* 2014;3:11.
57. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
58. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;41:1149–1160.
59. Shah VV, McNames J, Mancini M, et al. Quantity and quality of gait and turning in people with multiple sclerosis, Parkinson's disease and matched controls during daily living. *J Neurol* 2020;267: 1188–1196.
60. Ilg W, Milne S, Schmitz-Hubsch T, et al. Quantitative gait and balance outcomes for ataxia trials: consensus recommendations by the ataxia global initiative working group on digital-motor biomarkers. *Cerebellum* 2024;23:1566–1592.
61. Diallo A, Jacobi H, Tezenas du Montcel S, Klockgether T. Natural history of most common spinocerebellar ataxia: a systematic review and meta-analysis. *J Neurol* 2021;268:2749–2756.
62. Jacobi H, du Montcel ST, Bauer P, et al. Long-term disease progression in spinocerebellar ataxia types 1, 2, 3, and 6: a longitudinal cohort study. *Lancet Neurology* 2015;14:1101–1108.
63. Durr A. Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. *Lancet Neurology* 2010;9:885–894.
64. Beichert L, Seemann J, Kessler C, et al. Towards patient-relevant, trial-ready digital motor outcomes for SPG7: a cross-sectional prospective multi-center study (PROSPAX). *Neurology* 2024;103(12): e209887.
65. Timmann D, Horak FB. Perturbed step initiation in cerebellar subjects: 2. Modification of anticipatory postural adjustments. *Exp Brain Res* 2001;141:110–120.
66. Conte C, Serrao M, Casali C, et al. Planned gait termination in cerebellar ataxias. *Cerebellum* 2012;11:896–904.
67. Geh CL, Beauchamp MR, Crocker PR, Carpenter MG. Assessed and distressed: white-coat effects on clinical balance performance. *J Psychosom Res* 2011;70:45–51.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.