Journal of Neurology (2019) 266:1611–1622
https://doi.org/10.1007/s00415-019-09299-9

ORIGINAL COMMUNICATION

15-White Dots APP-Coo-Test: a reliable touch-screen application for assessing upper limb movement impairment in patients with cerebellar ataxias

Giuseppe Arcuria[1] · Christian Marcotulli[1] · Claudio Galasso[1] · Francesco Pierelli[1] · Carlo Casali[1]

Received: 26 January 2019 / Revised: 23 March 2019 / Accepted: 27 March 2019 / Published online: 6 April 2019

Abstract

Background The use of objective measurements is essential to assess disease progression and to evaluate the effectiveness of rehabilitation protocols and clinical treatments.

Aim The purpose of this study was to develop a touch-screen application, that we named *15-White Dots APP-Coo-Test* (15-WDACT), able to carry out quantitative and objective measurements of the rapid and coordinated upper limb movements, typically impaired in patients with cerebellar ataxias (CA).

Methods A total of 87 CA patients and 170 healthy subjects participated in this study. The subject was asked to touch with their index finger a white dot, appearing consecutively and randomly on the screen at different positions, for a total of 15 dots per session. The score is the execution time of a single session.

Results 15-WDACT measurements have highly correlated with the scores obtained with the Scale for the Assessment and Rating of Ataxia (SARA), with the Composite Cerebellar Functional Severity (CCFS) and with the measurements obtained using two validated evaluating systems, i.e., the Nine Hole Pegboard test (9HPT) and the Click Test. We also observed high internal consistency and an excellent intra-rater and test–retest reliability. We found a small Standard Error of Measurement (SEM) and an excellent Minimal Detectable Change (MDC), indicating that even small variations in the 15-WDACT measurements are to be associated with real changes in performance.

Conclusions We have concluded that 15-WDACT is an easy, fast and reliable tool to assess the severity of the upper limb ataxia in patients with CA.

Keywords Cerebellar ataxia · Neurodegenerative diseases · Upper limb movement impairment · Quantitative assessment · Touch-screen application · SARA

Introduction

The key motor symptoms of cerebellar ataxias are limb and gait ataxia, along with other symptoms [1]. The neurodegenerative CA are progressive disorders in which these symptoms get worse over time, with different progression rate according to the genotype [2–4]. Several validated clinical scales are available to assess the severity and progression of ataxia, such as the Scale for the Assessment and Rating of Ataxia (SARA) [5–9], the International Cooperative Ataxia Rating Scale (ICARS) [8, 10, 11], the Spinocerebellar Ataxia Functional Index (SCAFI) [12, 13] and the Composite Cerebellar Functional Severity (CCFS) [14]. The quantitative assessment of upper limb ataxia is different in each scale. In SARA it is assessed in ítems 5, 6 and 7 by the Finger-to-nose test, Finger-to-finger test and Fast alternating hands movements, respectively. The SCAFI and CCFS scales include the 9-Hole Pegboard Test (9HPT), but with different execution rules in each scale [14–16]: in SCAFI, the timing of 9HPT begins when the patient touches the first peg and ends when the last peg is placed in the container; in CCFS, the timing of 9HPT begins when the first peg is

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00415-019-09299-9) contains supplementary material, which is available to authorized users.

✉ Giuseppe Arcuria
giuseppe.arcuria@uniroma1.it

[1] Department of Medical and Surgical Sciences and Biotechnologies (DSBMC), Polo Pontino-University of Rome "Sapienza", Via Faggiana 34, 40100 Latina, Italy

placed in a hole and ends when the last peg is placed in the last hole. In both scales, the test is not easy to perform for patients with an advanced state of the disease (SARA > 30) [14, 17, 18]; the patient drops the pegs many times, forcing the examiner to stop the timer and ask the former to repeat the test again, thus generating fatigue and distress in the patient. We aimed to develop an easy-to-perform test that could be used also in an advanced state of the disease. The 15-WDACT is an application developed for touch-screen devices (tablet and mobile phone), freely downloadable from the web, compatible with Android and Apple operating systems. The quantitative measurement of each upper limb movement is the motion execution time. In this application the time is measured by the system and not by the examiner, thus no equipment other than a tablet is needed. At the end of the test, the application can generate a Report in a PDF format containing the execution time of each upper limb (dominant and non-dominant hand), an average time, the Standard Deviation (SD) and the Coefficient of Variation (CV) of the executed trials. Finally, the PDF report can be downloaded, printed and signed by the examiner to be stored. Some examples of 15-WDACT Report and the instructions how to generate it are given in Online Resources 1 and 2.

Materials and methods

Patients and controls

This study was conducted in the Department of Medical and Surgical Sciences and Biotechnologies (DSBMC), "Sapienza" University of Rome, Polo Pontino, at "Marco Pasquali" Institute—ICOT (Latina), from October 2016 to September 2018. We have recruited eighty-seven ataxic patients and one hundred and seventy healthy subjects. Out of the eighty-seven patients included, thirty-six had Friedreich's Ataxia (FRDA), nine had Spinocerebellar Ataxia type 1 (SCA1), six had SCA2, three had Myoclonic Epilepsy with Ragged-Red Fibers (MERRF), two had SCA3, two had SCA8, one had Autosomal Recessive Spastic Ataxia (ARSACS), one had Autosomal Recessive Ataxia type 8 (SCAR8), one had Multiple System Atrophy—Cerebellar type (MSA-C), and twenty-six had CA with no defined genetic analysis. The patients were 36 males and 51 females, aged between 22 and 76 years (mean age of 45.6 ± 13). We have also included three subjects with positive genetic analysis but clinically asymptomatic (two SCA1 and one MERRF), considered apart from ataxic and healthy subjects. One hundred and seventy healthy subjects were sex and age matched (85 males and 85 females, aged between 18 and 75 years, mean age of 41.36 ± 14.68). To detect any possible changes in the 15-WDACT execution time due to gender or age, we have considered the results based on male and females and based on two groups aged 18–45 and 46–75, respectively. To summarize and analyze the obtained data, the patients were divided into 18 groups according to the severity of the ataxia symptoms as assessed by SARA scale. The intra-rater reliability, the internal consistency and the accuracy (intra-subject variability) of the measurements obtained with the 15-WDACT were estimated via 5 consecutive measurements obtained for each hand and carried out on 48 patients. A 4-week test–retest was used to estimate the reliability over time in twenty-one patients. The time period between the first and the second administrations was 4 weeks, long enough to prevent learning, carry-over effects or recall. As 15-WDACT execution time was automatically measured by the application and not by the examiner, the inter-rater reliability was not estimated.

Clinical evaluation

The disease severity was clinically evaluated by SARA scale and CCFS score. The SARA scale consists of eight items, the sum of which determines a total score ranging from 0 (no ataxia) to 40 (most severe ataxia). The upper limb impairment was evaluated through items 5, 6 and 7 of the SARA scale, by the Finger-to-nose test, Finger-to-finger test and Fast alternating hand movements test, respectively [5–9]. Gait disturbances (item 1, score 0–8), posture disturbance (item 2, score 0–6 and item 3, score 0–4), speech disorders (item 4, score 0–6), down limb kinetic function (item 8, score 0–4) were evaluated too. To determine the CCFS score, the results obtained using the 9HPT and Click Test were combined according to the formula [14]

$$\mathrm{CCFS} = \log_{10}\left(7 + \frac{\mathrm{Z\ peg}}{10}\right) + \left(4 \times \frac{\mathrm{Z\ click}}{10}\right),$$

where Z peg (execution time of 9HPT with dominant hand) and Z click (execution time of Click Test with dominant hand) are calculated according to the following formulas, if age ≥ 20:

$$\mathrm{Z\ peg} = \mathrm{9HPT\ observed\ time} - \left(0.002 \times \mathrm{age}^2 - 0.16 \times \mathrm{age} + 13.4\right),$$

$$\mathrm{Z\ click} = \mathrm{Click\ observed\ time} - (0.05 \times \mathrm{age} + 8),$$

and according to the following, if age < 20:

$$\mathrm{Zpeg} = \mathrm{9HPT\ observed\ time} - \left(-0.08 \times \mathrm{age}^2 + 12.62\right),$$

$$\mathrm{Z\ click} = \mathrm{Click\ observed\ time} - \left(0.03 \times \mathrm{age}^2 - 1.14 \times \mathrm{age} + 18.89\right).$$

To perform the 9HPT, the Jamar 9-Hole Peg Test kit was used. The patients, who were seated, were asked to hold nine dowels (5 mm in diameter and 38 mm long) in the non-dominant hand and to place them randomly, one by one, with the dominant hand in a board with nine holes. Timing began

and stopped when the first and last pegs, respectively, were placed. To perform the Click Test, the patients were asked to press alternatively, using their index finger, the buttons placed on two mechanical counters fixed on a wooden board, 39 cm apart. Timing began when one button was pressed for the first time and stopped when both the counters reported the number ten. 9HPT and Click Test were executed once with the dominant hand.

Functional tests

The upper limb movement impairment was evaluated using the 9HPT (SCAFI rules) [12, 13], the Click Test (CCFS rules) [14] and the new 15-WDACT. To perform the 9HPT, the pegboard (Jamar 9-Hole Peg Test kit) was placed on the table facing the patient with the peg container in front of the hand that was going to be tested, while the other hand was used to stabilize the board. The timing of the 9HPT began when the patient touched the first peg and stopped when the last peg was removed and hit the container. If a peg was dropped during the test, the examiner stopped the timer and the patient restarted the test from the beginning. The test was performed for each hand. To perform the Click Test, a wooden board with two mechanical counters fixed 39 cm apart was placed on a table facing the patient. The patient, who was seated, was asked to press the buttons placed on the two counters ten times alternatively using their index finger. Timing began when the button in front of the hand that was going to be tested was pressed for the first time and stopped when the same button was pressed for the tenth time. Two trials of 9HPT and Click Test were performed with the dominant hand, followed by two trials of the non-dominant hand. For each test an average of the respective four trials was finally considered. The pause in between each trial was no longer than 5 min. Time was reported within 0.1 s rounded as needed. To execute the 15-WDACT a 10.1-inch-sized tablet-PC was used. It was placed on a table with the touch-screen facing the sitting patient. Patients were asked to touch with their index finger 15 white dots (2.5 cm in diameter) appearing consecutively and randomly on the screen at different positions, each one showing up after the previous one was correctly reached. The patient was not allowed to rest their forearm and elbow on the desk during the test. As illustrated in Online Resource 3 and shown in the animation (Online Resource 4) the test began after the patient pressed the start button, while time started to run when the first dot appeared on the screen and stopped when the last one was successfully reached. The dominant hand was tested first for two consecutive trials, followed by two trials of the non-dominant hand. All trials were separated by a 1-min pause. Scores from the two trials for each hand were averaged and then converted to the reciprocals of the mean times for each hand. Finally, the two reciprocals were averaged and used as

Table 1 Measurements obtained in one hundred and seventy healthy subjects (eighty-five females and eighty-five males) using the 15-WDACT, 9HPT and Click Test

FEMALE									MALE								
n	Age (range)	15-WDACT					9HPT	Click Test	*n*	Age (range)	15-WDACT					9HPT	Click Test
		D.H.	N.D.H.	Mean	Lowe	High					D.H.	N.D.H.	Mean	Lowe	High		
20	18–29	8.50	8.87	8.68	7.91	9.13	16.85	10.05	20	18–29	8.54	8.92	8.73	7.88	9.57	17.05	10.24
20	30–39	8.54	8.93	8.74	8.02	9.96	16.28	10.46	20	30–39	8.66	8.91	8.79	7.92	10.02	17.22	10.82
20	40–49	8.92	9.32	9.12	8.41	10.22	17.93	10.72	20	40–49	8.90	9.34	9.12	8.04	10.18	18.26	11.02
15	50–60	9.65	10.21	9.93	8.83	10.96	19.50	11.59	15	50–60	9.71	10.31	10.01	8.62	11.43	20.01	12.04
10	61–75	11.04	11.73	11.38	10.21	13.72	21.89	13.01	10	61–75	10.94	11.91	11.42	9.94	14.02	22.71	13.52
85	18–75	9.11	9.56	9.34	7.91	13.72	18.03	10.93	85	18–75	9.14	9.61	9.38	7.88	14.02	18.56	11.27

Test execution times are expressed in seconds

D.H. dominant hand, *N.D.H.* non-dominant hand

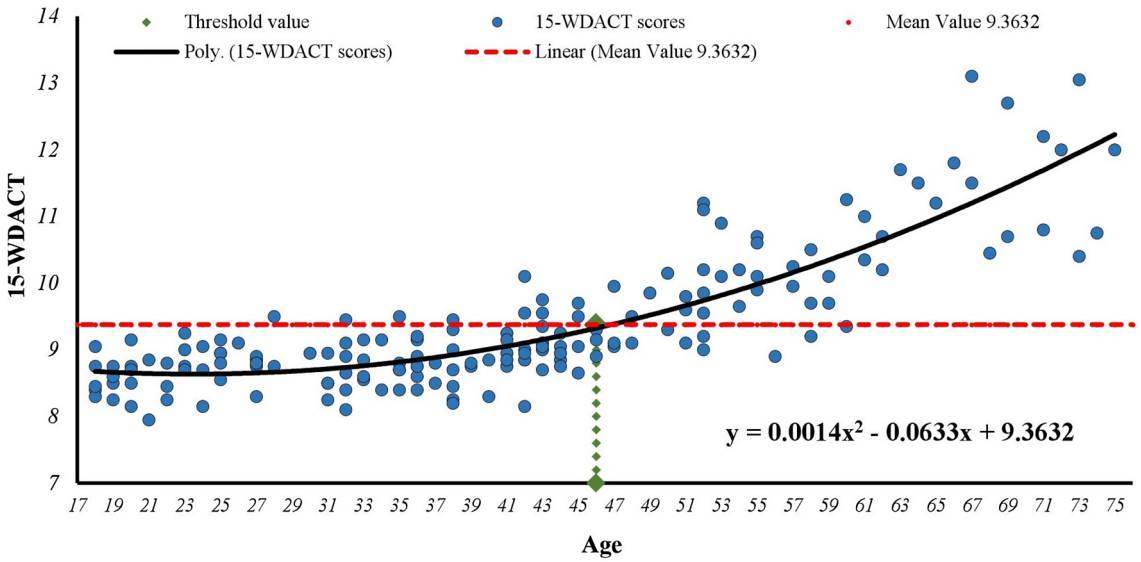

Fig. 1 Polynomial regression plot showing the increase of the 15-WDACT measurements due to age in one hundred and seventy healthy subjects. 15-WDACT scores are expressed in seconds and reported on the *y*-axis, while ages are reported on the *x*-axis

final score. The final score represents the time needed for the task execution, reported to within 0.01 s, rounded as needed.

Statistical analyses

The statistical analysis was performed using the IBM SPSS statistics base software Version 22 and Microsoft Excel 2007. To investigate the correlation between the scores obtained with the application and the severity of the ataxic symptoms assessed by SARA and CCFS, a linear regression modeling was used. The Pearson's correlation coefficient (R) was used to determine the correlation between the 15-WDACT scores with the 9HPT and Click Test scores, evaluated with a 95% confidence interval (CI), while the significance with p values (whose minimum levels were set at $p < 0.01$). Pearson's correlation was used to examine construct validity of the APP scores. Pearson's correlation exceeding 0.75 were considered good–excellent [19]. To detect any possible changes in the 15-WDACT execution time due to gender or age factor, the Student's t test for paired sample was used [19]. Polynomial regression, a second-degree polynomial function, was computed to test the influence of aging on the 15-WDACT. To compare the results obtained with the 15-WDACT and determine if the mean scores obtained in different groups of patients were significantly different from each other, we used the one-way analysis of variance (ANOVA) and calculated F statistic of Fisher–Snedecor (95% CI) [19]. To analyze the APP scores accuracy, we calculated the CV according to the formulas $\mathrm{CV} = \frac{\mathrm{SD}\times 100}{\mathrm{Mean}}$ and $\mathrm{SD} = \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - X)^2}{n}}$, where SD is the standard deviation, n is the number of the x_i elements and X is the mean of all elements. The CV is a measure of the variability of a data set and the closer to 0 the CV is the less variable the data are [19]. The internal consistency of the measurements obtained with the 15-WDACT was estimated by Cronbach's alpha while the intra-rater reliability and the reliability over time (test–retest reliability) were estimated by intraclass correlation coefficient ($\mathrm{ICC}_{2.1}$) with absolute agreement, single rater/measurement and a 95% CI [20]. Good reliability and consistency are indicated by an $\mathrm{ICC}_{2.1}$ and a Cronbach's alpha values greater than 0.75. An $\mathrm{ICC}_{2.1}$ or a Cronbach's alpha exceeding 0.90 were considered excellent [19]. Estimates of reliability (absolute estimates) were also calculated using the standard error of measurement (SEM) [21] and the minimal detectable change (MDC) [22]. The SEM, that is a measure of variability, indicates the stability of the score when a measure is repeated. The SEM that was used to evaluate consistency of scores from each trial for each participant was calculated using the following formula:

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

Smaller SEM numbers indicate a more reliable measurement.

The MDCn (n = CI), referred to as the smallest detectable difference, is an absolute measure of reliability (measurement error). The MDC_{95} was calculated, after having set a CI = 95%, with the following formula:

$$\mathrm{MDC}_{95} = \mathrm{SEM} \times 1.96 \times \sqrt{2},$$

Table 2 Clinical evaluation and functional tests (15-WDACT, 9HPT and Click Test) carried out on eighty-seven CA patients stratified by whole SARA scores (data sets reported on the left) and according to the sum of items 5, 6, 7 (Finger-to-nose test, Finger-to-finger test and Fast alternating hand movements test) of SARA (data sets reported on the right)

k	Patients					Functional tests										*k*	Patients				15-WDACT		
	Age			Clinical scales		15-WDACT						*n*	9HPT[b]	*n*	Click test[c]		Age			SARA items (5, 6, 7)			
	n	Mean value	Range	SARA total score (range)	CCFS	*n*	Dominant hand	Non-dominant hand	Observed time	Age factor	Final score[a]						*n*	Mean value	Range		Observed time	Age factor	Final score[a]
1	3	46.3	33–54	2–3	1.02	3	11.60	12.07	11.83	−0.36	11.47	3	27.3	3	15.5	1	6	49.3	33–54	0	11.95	−0.32	11.62
2	5	53.8	38–60	4–5	1.00	5	11.78	12.82	12.30	−0.78	11.52	5	26.4	5	14.3	2	5	58.4	46–70	1	13.04	−0.94	12.10
3	4	57.3	42–71	6–7	1.04	4	12.88	13.90	13.39	−1.22	12.17	4	28.6	4	16.1	3	9	50.2	38–63	2	15.40	−0.30	15.10
4	7	46.4	42–52	8–9	1.09	7	14.39	16.39	15.39	−0.18	15.21	7	40.5	7	19.0	4	14	47.8	31–68	3	16.50	−0.22	16.28
5	6	49.3	43–58	10–11	1.12	6	15.49	16.56	16.02	−0.32	15.70	6	42.3	6	20.4	5	12	49.6	25–76	4	19.70	−0.73	18.97
6	6	55.0	31–76	12–13	1.19	6	18.07	20.83	19.45	−1.11	18.34	6	46.9	6	26.1	6	8	46.5	26–65	5	21.60	−0.28	21.32
7	7	43.1	20–67	14–15	1.16	7	17.14	19.97	18.56	−0.39	18.17	7	54.9	7	23.2	7	7	38.4	24–67	6	30.06	−0.25	29.81
8	5	58.2	25–75	16–17	1.20	5	17.64	19.94	18.79	−1.59	17.20	5	44.0	5	26.3	8	8	36.7	28–53	7	31.38	−0.04	31.34
9	4	52.0	43–65	19–20	1.22	4	17.44	19.78	18.61	−0.59	18.02	4	52.4	4	31.6	9	8	45.5	33–68	8	32.21	−0.22	31.99
10	3	49.0	38–65	21–22	1.33	3	19.77	23.10	21.43	−0.58	20.85	3	75.5	3	33.3	10	4	33.2	22–49	9	42.38	−0.03	42.38
11	3	44.7	23–57	23–24	1.26	3	20.97	22.25	21.61	−0.51	21.10	3	59.6	3	31.0	11	1	45	45	10	N/D	N/D	N/D
12	5	48.2	26–58	25–26	1.44	5	25.88	32.88	29.38	−0.60	28.79	5	125.4	5	43.4	12	0	N/D	N/D	11	N/D	N/D	N/D
13	5	39.0	25–50	28–29	1.39	5	24.66	26.66	25.66	−0.08	25.58	5	104.2	5	37.9	13	5	38	33–45	12	N/D	N/D	N/D
14	9	33.3	22–45	30–31	1.52	9	28.89	31.33	30.11	0.00	30.11	8	137.6	9	48.3								
15	4	38.3	33–45	32–33	1.50	4	30.33	35.64	32.99	0.00	32.99	3	168.1	4	49.4								
16	5	35.8	29–49	34–35	N/D	5	38.32	53.30	45.81	−0.04	45.77	0	N/D	5	103.3								
17	2	45.3	42–45	36–38	N/D	0	N/D	N/D	N/D	N/D	N/D	0	N/D	0	N/D								
18	4	37.0	33–42	39–40	N/D	0	N/D	N/D	N/D	N/D	N/D	0	N/D	0	N/D								
	87	48.4 ± 13	20–76	19.87 ± 10.89	1.23 ± 0.17	81	20.66 ± 8.31	23.93 ± 11.21	22.29 ± 9.50		21.81 ± 9.68	74	72.7 ± 52.9	81	32.6 ± 20.6		87	48.4 ± 13	20–76		22.29 ± 9.50		21.81 ± 9.68

Data sets are expressed as mean values. Summered data are expressed as weighted average + standard deviation. Test execution times are expressed in seconds

k group identification number, *n* number of tested patients, *N/D* no data

[a]All values were adjusted according to age, subtracting from the 15-WDACT observed time the "Age factor = {(0.0014 × Age^2) − (0.0633 x Age)}" for patients older than 45 years

[b]SCAFI rules

[c]CCFS rules

where 1.96 is the z score associated with the 95% confidence level and $\sqrt{2}$ is used to account for the underlying extra uncertainty during measurement in two time points. MDC values are used to assist in interpreting results and determining whether a change between repeated tests is a random variation or a real change (improvement or deterioration) in performance [23]. An MDC% < 10% with MDC% calculated with the following formula

$$\mathrm{MDC\%} = \mathrm{MDC} / \mathrm{Mean} \times 100$$

was considered as excellent [23].

Results

Controls

All healthy subjects were tested with both 9HPT and Click Test, with data set stratified for age groups and divided for sex. Table 1 reports also the mean execution time of the 15-WDACT, both with the dominant (D.H.) and non-dominant hand (N.D.H.). As reported in Table 1 and shown in Fig. 1, the final average of the 15-WDACT scores of the control group was 9.36 ± 1.01 (9.34 for the females and 9.38 for males). The difference in the mean execution

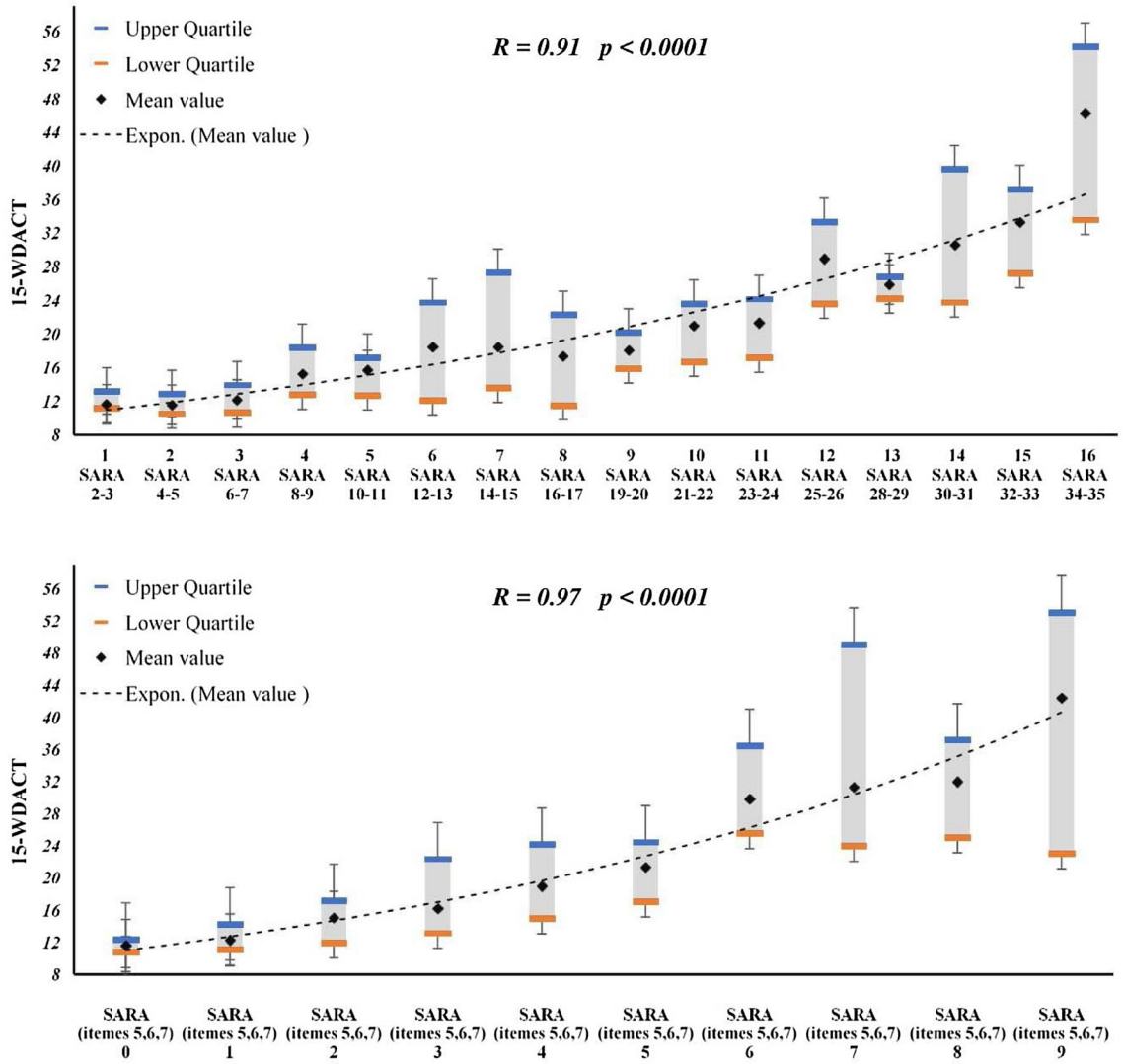

Fig. 2 Box plot showing progressively higher 15-WDACT scores in sixteen groups of patients, stratified by SARA whole scores (upper graph), and in ten groups of patients stratified by the sum of SARA items 5, 6, 7 (Finger-to-nose test, Finger-to-finger test and Fast alternating-movements-hand test). 15-WDACT measurements are expressed in seconds and reported on the y-axis, while groups with their SARA total score (first group) and the sum of SARA items 5, 6, 7 (second group) are reported on the x-axis of the upper and lower graphs, respectively. R is Pearson's correlation coefficient

time of the 15-WDACT based on gender was non-significant ($p > 0.05$). The executed T test disclosed that there were no differences based on sex (t value $t = -0.26936$ and a p value $p = 0.393993$), whereas the difference in the mean execution time of the 15-WDACT in two groups of healthy subject, aged 18–45 and 46–75, respectively, was significant ($p < 0.001$), as confirmed by a T test and a t value = − 13.6496 (additional data are given in Online Resource 5). As shown in Fig. 1, the 15-WDACT execution time tended to increase with age in subjects older than 45 years (threshold value), according to the following second-degree polynomial function:

$$Y = 9.3642 + \left[\left(0.0014 \times \text{age}^2\right) - (0.0633 \times \text{age})\right],$$

where 9.3642 is the mean value of 15-WDACT obtained in control group and $\left[\left(0.0014 \times \text{age}^2\right) - (0.0633 \times \text{age})\right]$ is the increase of the 15-WDACT score based on age at examination.

Patients

In the ataxic patients, the 15-WDACT execution time tended to increase with the disease severity but changes in the score due to age must be distinguished from those caused by the worsening of the disease. In subsequent analyses, the 15-WDACT scores were, therefore, adjusted for age > 45 (threshold value). We calculated the age factor according to the following formula:

$$\text{Age factor} = \left[\left(0.0014 \times \text{age}^2\right) - (0.0633 \times \text{age})\right].$$

Age factor was subtracted from the 15-WDACT observed time to get the adjusted score. The 87 patients were divided into 18 groups according to the severity of the ataxia symptoms, as assessed by SARA scale (Table 2). 9HPT measurements were obtained in seventy-four patients while Click Test and 15-WDACT were obtained in eighty-one patients. None of the six patients with SARA score ranging 36–40 were able to complete the 15-WDACT, the 9HPT and the Click Test, while seven patients with a SARA score ranging 31–35 were not able to conclude the 9HPT. The F value calculated with the one-way ANOVA test was $F = 25.0958$ (additional data are given in Online Resource 6). Therefore, as shown in Fig. 2, the mean values obtained with the 15-WDACT in the sixteen patient groups, stratified according to the whole SARA score, were significantly different ($p < 0.001$), with an increase in average execution time as the severity of the ataxic symptoms worsened ($R = 0.91$). The correlation between the 15-WDACT measurements and the sum of the scores of items 5, 6, 7 of the SARA scale (Fig. 2) was very strong too ($R = 0.97$). The correlation between the measurements obtained with the

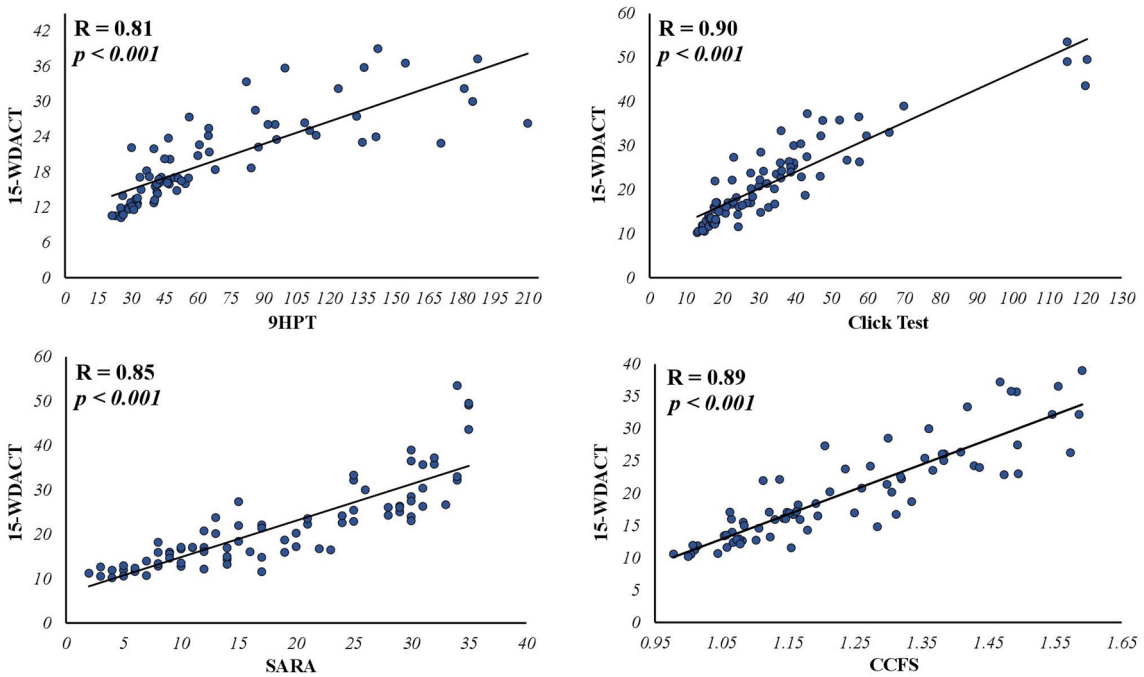

Fig. 3 Linear regression plots showing the correlation of the 15-WDACT measurements with the 9HPT scores of seventy-four patients (upper left graph), with the Click Test scores of eighty-one patients (upper right graph), and with the SARA (lower left graph) and CCFS scores (lower right graph) of eighty-one ataxic patients. Measurements obtained with the 15-WDACT (in seconds) are reported on the y-axes, while those obtained with the 9HPT and Click Test (all expressed in seconds) and with the SARA and the CCFS are reported on the x-axis of the upper left, upper right, lower left graph and lower right graph, respectively. R is Pearson's correlation coefficient

Table 3 Intra-rater reliability estimated via five consecutives trials carried out on forty-eight CA patients with a SARA mean value of 15.85

Patients	*n*	15-WDACT (execution time in seconds)					Mean (all trials)	CV (%)	SD	$ICC_{2.1}$ (95% CI)	SEM	MDC_{95} (%)	α	*p* value
		T_1	T_2	T_3	T_4	T_5								
Dominant	48	18.58	17.99	18.23	18.48	18.13	18.53	0.061 (6.1)	1.24	0.98 (0.97–0.99)	0.175	0.486 (2.6)	0.98	< 0.001
Non-dominant	48	20.82	20.39	20.82	20.70	20.30	21.30	0.055 (5.5)	1.22	0.98 (0.97–0.99)	0.172	0.478 (2.2)	0.98	< 0.001
Mean	48	19.70	19.19	19.52	19.59	19.21	19.91	0.058 (5.8)	1.23	0.98 (0.97–0.99)	0.173	0.482 (2.4)	0.98	< 0.001

Test execution times are expressed in seconds. There was a 1-min pause in between each of the five trials

n number of patients tested, T_n number of trials, *SD* standard deviation, *CV* coefficient of variation (SD/Mean), $ICC_{2.1}$ two-way random effects, absolute agreement, single rater/measurement (IC = 95%), *SEM* standard error of measurement, MDC_{95} minimal detectable change with interval of confidence (CI) 95%, MDC_{95}(%) MDC_{95}/mean, *α* Cronbach's alpha

15-WDACT and the scores obtained with the 9HPT and Click Test was high (Fig. 3). The performance obtained with the 15-WDACT by each of the 81 observed patients, showed also a strong correlation with the severity of the ataxic symptoms assessed by SARA scale and CCFS score (Fig. 3). As reported in Table 3, the intra-rater reliability ($ICC_{2.1}$) was excellent. The accuracy (CV) of the measurements was high. Internal consistency of the data obtained using the APP (Cronbach's alpha) was excellent too. As shown in Fig. 4, the intra-subject variability in the 5 performances obtained with the 15-WDACT by each of the 48 observed patients, represented for each patient by the gap between each of the five scores, tended to increase as the severity of the ataxia worsened [24]. As reported in Table 4, the reliability over time (rest–retest reliability) of the measurements obtained with the APP, expressed by $ICC_{2.1}$, was excellent. The correlation between the scores obtained in two different administrations was very strong ($R = 0.98$). The accuracy of the obtained data was high. As reported in Table 3 and in Table 4, the SEM was small and all MDC% were between 2.0% and 2.6%, representing an excellent random measurement error.

Finally, the measurements obtained in three asymptomatic subjects aged 40–49 showed a mean score to perform the 15-WDACT of 9"23, in accordance with those obtained from healthy subjects aged 40–49 (Table 1).

Discussion

The 9HPT is currently the most used method for quantitative evaluation of the upper limb movement impairment. The Click Test is another method used for the same purpose. However, the Click Test and the 9HPT are not easy to use for patients with an advanced state of ataxia [14, 17, 18]. In our study, patients with a SARA score > 30 have completed the 9HPT (CCFS and SCAFI rules) and the Click Test with many difficulties. Patients with a severe ataxic symptomatology dropped the pegs many times, forcing the examiner to stop the timer and ask them to repeat the test again, thus generating fatigue and distress in the patient. Seven patients with SARA scores ranging 31–35 have not been able to complete the 9HPT, but they have instead executed the 15-WDACT with relative ease. Measurements obtained in all patients show a mean performing time of 72.7 ± 52.9 s to execute the 9HPT, 32.6 ± 20.6 for the Click Test and 21.81 ± 9.68 for the 15-WDACT (Table 2). As compared with both Click Test and 9HPT, the App-Coo-Test proved to be faster and easier to perform. As reported in the controls, there were no differences based on sex, whereas the 15-WDACT execution time tended to increase in subjects older than 45 years old. Consequently, all values obtained in patients older than 45 years were adjusted automatically by the APP which subtracted the "Age Factor" from the observed execution time, allowing us to study the rate of progression of

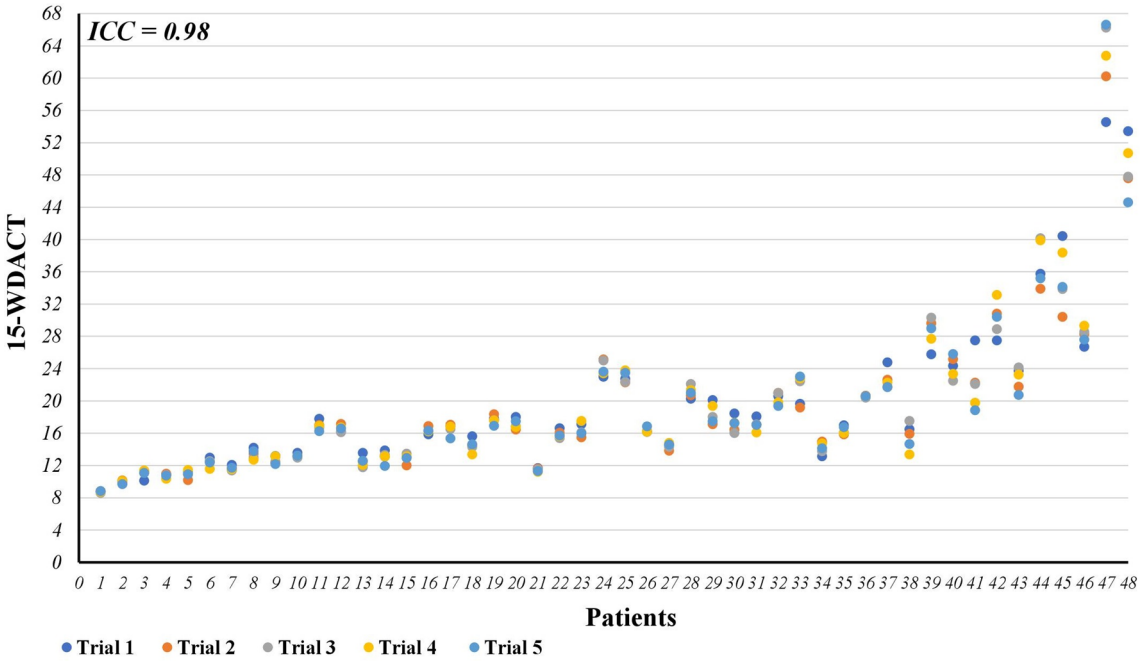

Fig. 4 A dot plot showing an excellent intraclass correlation ($ICC_{2.1}$) calculated via five repeated trials carried out on forty-eight CA patients. 15-WDACT scores are expressed in seconds and reported on the *y*-axis, while the five trials performed by each patient are reported on the *x*-axis. From left to right, patients are reported on the *x*-axis based on an increasing value of the SARA. The intra-subject variability, graphically represented by the gaps between the five dots (trials) for each patient, tends to increase as the severity of ataxia symptomatology worsens

the 15-WDACT scores independently from the age differences. The correlation between the 15-WDACT and the Click Test and 9HPT, respectively, was very strong. The ANOVA test explained that the 15-WDACT execution time was progressively higher as the severity of the ataxic symptoms increased (Fig. 2). The Test–Retest method showed a very high level of intra-rater and reliability over time ($ICC_{2.1} = 0.98$), while the inter-rater reliability was not estimated, as the tally of the test is automatic. The accuracy of the measurements, assessed by CV, was very high. We found an excellent level of MDC_{group} ranging from 0.48 to 0.35, corresponding to an average value of MDC% of 2.3%, which indicates that a change greater than just 2.3% could be viewed, with 95% certainty, as a real change in performance [23]. As we found that the intra-subject variability in the performances obtained with the 15-WDACT tended to increase as the severity of the ataxia worsened (Fig. 4), we believe that the $MDC_{individual}$ could be smaller for low SARA scores and tending to increase proportionally to the ataxia severity.

Conclusion

SARA and ICARS are semiquantitative scales that rely on subjective ratings by clinicians. Therefore, these scales yield only partially objective measurements of upper limb movement impairment [16, 17]. To develop new and effective drugs or to test the effectiveness of new rehabilitative protocols, it is necessary to develop widely applicable methods for the objective evaluation in CA. The scores obtained from the 15-WDACT showed a high correlation with those obtained from the 9HPT and Click Test; therefore, we think that the new touch-screen application may be a valid method able to objectively quantify the upper limb movement impairment regardless of the disease from which the disorder arises. The strong correlation between the 15-WDACT measurements and the sum of items 5, 6, 7 of the SARA, confirm that the APP may be an alternative method to assess the severity of upper limbs ataxia. As the most numerous subgroups of observed patients was the one with Friedreich's ataxia ($n = 36$), the data obtained suggest that the 15-WDAC can be certainly used to evaluate the rapid and coordinated upper limb movements, typical impairment in these patients. Furthermore, the high correlation between the scores measured with the 15-WDACT and those obtained both with the whole of SARA score and with the CCFS score in a heterogeneous group of CA patients, like the one selected in this study, seems to confirm that the APP may be an alternative method to assess the progression of the ataxic disorder, independently of the genetic results. The 15-WDACT is an easy-to-perform functional test compared to the

Table 4 Intra-rater reliability, reliability over time and minimal detectable change (MDC_{group}) estimated via five consecutives tests carried out in two different administrations (4 weeks from each other) on thirty-one CA patients with a SARA mean value of 12.04

Patients		Intra-rater reliability																			Reliability over time							
		First administration									Second administration										Test–retest		*R*	*p* value	$ICC_{2.1}$ (95% CI)	SD	SEM	MDC_{95} (%)
		15-WDACT (execution time in seconds)					SD	CV	$ICC_{2.1}$	α	15-WDACT (execution time in seconds)					SD	CV	$ICC_{2.1}$	α		First admin	Second admin						
	n	T_1	T_2	T_3	T_4	T_5					T_1	T_1	T_3	T_4	T_5						Mean	Mean						
D.H.	31	16.70	16.28	16.45	16.81	16.28	0.77	0.04	0.98	0.98	16.79	15.93	16.40	16.67	16.17	0.92	0.05	0.98	0.97		16.50	16.40	0.98	< 0.001	0.98 (0.96–0.99)	0.85	0.120	0.333 (2.0)
N.D.H.	31	17.90	17.43	17.62	17.54	17.31	1.03	0.06	0.98	0.98	17.72	17.73	17.55	17.85	17.44	0.88	0.05	0.98	0.97		17.56	17.66	0.98	< 0.001	0.98 (0.96–0.99)	0.95	0.134	0.366 (2.1)
Mean	31	17.53	16.86	17.03	17.17	16.79	0.90	0.05	0.98	0.98	17.26	16.83	16.97	17.26	16.81	0.90	0.05	0.98	0.97		17.03	17.03	0.98	< 0.001	0.98 (0.96–0.99)	0.90	0.127	0.352 (2.1)

Test execution times are expressed in seconds

D.H. dominant hand, *N.D.H.* non-dominant hand, *n* number of tested patients, T_n number of tests, *SD* standard deviation, *CV* coefficient of variation (SD/Mean), $ICC_{2,1}$ two-way random effects, absolute agreement, single rater/measurement, α Cronbach's alpha, *R* Pearson's coefficient of correlation, *SEM* standard error of measurement, MDC_{95} minimal detectable change with interval of confidence (CI) 95%, $MDC_{95}(\%)$ MDC_{95}/mean

9HPT, so it seems to be a reliable method to assess ataxic patients, especially those with a high severity of the disease (SARA > 30) who are often excluded from clinical trials, being unable to perform the quantitative measurement tests currently used. Regarding a possible scoring affected by the involvement of other neurological system like distal upper limb weakness, which is present in many patients with hereditary ataxias, it would be interesting to administer the 15-WDACT to patients with a pure ataxia (autosomal dominant cerebellar ataxia type III), to determine the effect caused by upper limb disability from involvement of other systems. Considering that the autosomal dominant cerebellar ataxia (ADCA) of type III as SCA5, SCA6 (the most common form of this group with an estimated prevalence of less than 1/100,000), SCA11, SCA26, SCA30, and SCA31 (the second most common form that is found mainly in Japan) are very rare forms of ataxias [25–27], a multicenter study would be necessary for this purpose. Moreover, as the 9HPT is used in clinical trials that involve patients with different disease, i.e., parkinsonism, multiple sclerosis, etc., in the next future it would be interesting to administer the 15-WDACT to patients with other motor coordination disorders. Finally, like SARA scale, that assesses upper and lower limb, we are developing other APP-Coo-Test functions able to evaluate lower limb performance to generate a composite score.

Compliance with ethical standards

Conflicts of interest The manuscript has been approved by all the authors. The authors declare that they have no financial kind of relationships that might lead to a conflict of interest.

Ethical approval This study was conducted in the Department of Medical and Surgical Sciences and Biotechnologies, "Sapienza" University of Rome, at "Marco Pasquali" Institute—ICOT (Latina), from October 2016 to September 2018. All patients who participated to the study were informed of the use of measurements for research purposes and gave their written informed consent. All the procedures performed were in accordance with the ethical standards of the institutional and national research committee. The regulations of our institution concerning intellectual property has been respected. The ethical standards, laid down in 1964 Declaration of Helsinki and its later amendments, were respected.

References

1. Bodranghien F, Bastian A, Casali C, Hallett M, Louis ED, Manto M et al (2016) Consensus paper: revisiting the symptoms and signs of cerebellar syndrome. Cerebellum 15(3):369–391
2. Patel M, Isaacs CJ, Seyer L, Brigatti K, Gelbard S, Strawer C et al (2016) Progression of Friedreich ataxia: quantitative characterization over 5 years. Ann Clin Transl Neurol 3:684–694
3. Reetz K, Dogan I, Hilgers RD, Giunti P, Mariotti C, Durr A et al (2016) Biological and clinical characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) cohort: a cross-sectional analysis of baseline data. Lancet Neurol 15:1346–1354
4. Diallo A, Jacobi H, Cook A, Labrum R, Durr A, Brice A et al (2018) Survival in patients with spinocerebellar ataxias types 1,2,3 and 6 (EUROSCA): a longitudinal cohort study. Lancet Neurol 17:327–334
5. Yabe I, Matsushima M (2008) Usefulness of the scale for the assessment and rating of ataxia (SARA). J Neurol Sci 266:164–166
6. Schmitz-Hubsch T et al (2006) Scale for the assessment of ataxia: development of a new clinical scale. Neurology 66:1717–1720
7. Weyer A, Abele M, Schmitz-Hubsch T (2007) Reliability and validity of the scale for the assessment and rating of ataxia: a study in 64 ataxia patients. Mov Disord 22:1633–1637
8. Bürk K, Sival DA (2018) Scales for clinical evaluation of cerebellar disorders. Hand Clin Neurol 154:329–339
9. Tanguy Melac A, Mariotti C, Filipovic Pierucci A, Giunti P, Arpa J, Boesch S et al (2018) Friedreich and dominant ataxias: quantitative differences in cerebellar dysfunction measurements. J Neurol Neurosurg Psychiatry 89:559–565
10. Schmitz-Hübsch T, Tezenas-Du-Montcel S, Baliko L, Boesch S, Bonato S, Fancellu R et al (2005) Reliability and validity of the international cooperative ataxia rating scale: a study in 156 spinocerebellar ataxia patients. Mov Disord 21:699–704
11. Tison F, Yekhlef F (2002) Application of the international cooperative ataxia scale rating in multiple system atrophy. Mov Disord 17(6):1248–1254
12. Schmitz-Hübsch T, Giunti P, Stephenson DA, Globas C, Baliko L, Saccà F et al (2008) SCA Functional Index—usefulness of a compound performance measure in spinocerebellar ataxia patients. Neurology 71:486–492
13. Schmitz-Hübsch T, Fimmers R, Rakowicz M, Rola R, Zdzienicka E, Fancellu R et al (2010) Responsiveness of different rating instruments in spinocerebellar ataxia patients. Neurology 74:678–684
14. du Montcel ST, Charles P, Ribai P, Goizet C, Le Bayon A, Labauge P et al (2008) Composite cerebellar functional severity score: validation of a quantitative score of cerebellar impairment. Brain 131:1352–1361
15. Mathiowetz V, Weber K, Kashman N, Volland G (1985) Adult norms for the nine hole peg test of finger dexterity. Occup Ther J Res 5(1):24–38
16. Oxford Grice K, Vogel KA, Le V, Mitchell A, Muniz S, Vollmer MA (2003) Adult norms for a commercially available nine hole peg test for finger dexterity. Am J Occup Ther 57(5):570–573
17. Paap BK, Roeske S, Durr A, Schöls L, Ashizawa T, Boesch S, Bunn LM et al (2016) Standardized assessment of hereditary ataxia patients in clinical studies. Mov Disord Clin Pract 3(3):230–240
18. Tanguy Melac A, Mariotti C, Filipovic Pierucci A, Giunti P, Arpa J, Boesch S, Klopstock T, Müller Vom Hagen J, Klockgether T, Bürk K, Schulz JB, Reetz K, Pandolfo M, Durr A, Tezenas du Montcel S (2018) Friedreich and dominant ataxias: quantitative differences in cerebellar dysfunction measurements. J Neurol Neurosurg Psychiatry 89(6):559–565
19. Portney LG, Watkins MP (2009) Foundations of clinical research: applications to practice. Pearson/Prentice Hall, Upper Saddle River
20. Baumgartner TA, Chung H (2001) Confidence limits for intraclass reliability coefficients. Meas Phys Educ Exerc Sci 5:179–188
21. Harvill L (1991) Standard error of measurement. Educ Meas 10:33–41
22. Donoghue D, Stokes EK (2009) How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. J Rehabil Med 41:343–346

23. Haley SM, Fragala-Pinkham MA (2006) Interpreting change scores of tests and measures used in physical therapy. Phys Ther 86:735–743
24. Serrao M, Pierelli F, Ranavolo A, Draicchio F, Conte C, Casali C et al (2012) Gait pattern in inherited cerebellar ataxias. Cerebellum 11(1):194–211
25. Fujioka S, Sundal C, Zbigniew K (2013) Autosomal dominant cerebellar ataxia type III: a review of the phenotypic and genotypic characteristics. Orphanet J Rare Dis 18(8):14
26. Schols L, Bauer P, Schmidt T, Schulte T, Riess O (2004) Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. Lancet Neurol 3:291–304
27. Basri R, Yabe I, Soma H, Sasaki H (2007) Spectrum and prevalence of autosomal dominant spinocerebellar ataxia in Hokkaido, the northern island of Japan: a study of 113 Japanese families. J Hum Genet 52:848–855